Article:
Mosammam, AM and Kent, JT (2016) Estimation and testing for covariance-spectral spatial-temporal models. Environmental and Ecological Statistics, 23 (1). pp. 43-64. ISSN 1352-8505

https://doi.org/10.1007/s10651-015-0322-y

Environmental and Ecological Statistics manuscript No.
(will be inserted by the editor)

Estimation and Testing for Covariance-Spectral Spatial-Temporal models

Ali M. Mosammam · John T. Kent

Received:14 Nov. 2014 / Accepted: date

Abstract In this paper we explore a *covariance-spectral modelling* strategy for spatial-temporal processes which involves a spectral approach for time but a covariance approach for space. It facilitates the analysis of coherence between the temporal frequency components at different spatial sites. Stein (2005) developed a semi-parametric model within this framework. The purpose of this paper is to give a deeper insight into the properties of his model and to develop simpler and more intuitive methods of estimation and testing. A very neat estimation for drift direction is proposed while Stein assumes it is known. An example is given using the Irish wind speed data. Stein constructed various plot to assess the goodness of fit of the model, we use similar plots to estimates the parameters.

Keywords Space-time model · Covariance-spectral model · Coherence function · Asymmetry

1 Introduction

There is a need for tractable yet flexible spatial-temporal models in applications such as environmental modelling. Two natural starting points are models for purely spatial or purely temporal data. For example, one may consider time as an extra spatial dimension; then spatial statistics techniques (Cressie, 1993) can be applied. However, this approach ignores the fundamental differences between space and time such as coherence, which arises, e.g., if a wind is

Ali M. Mosammam
Department of Statistics, University of Zanjan, Zanjan, Iran
Tel.: +98- 24 -33054081
E-mail: a.m.mosammam@znu.ac.ir

John T. Kent
Department of Statistics, University of Leeds, Leeds LS2 9JT, UK
E-mail: j.t.kent@leeds.ac.uk

blowing across a spatial region. On the other hand, starting from a time series perspective, one way to think of a spatial-temporal process is as a multiple time series (Priestley, 1981) where the spatial locations of the data index the components of the time series. However, this approach ignores the regularity in space and does not allow inferences about the process at sites where data are not observed.

In this paper we focus on a *covariance-spectral* modelling strategy which intertwines the roles of space and time in a deeper way. Consider a real-valued stationary spatial-temporal process $Z(\boldsymbol{s},t)$ defined on $\mathbb{R}^d \times \mathbb{R}$ with covariance function $C(\boldsymbol{h},u)$, where $\boldsymbol{h} \in \mathbb{R}^d$ represents a spatial lag in d dimensions, and $u \in \mathbb{R}$ represents a temporal lag. The covariance function has a spectral representation

$$C(\boldsymbol{h},u) = FT_{ST}\{f(\boldsymbol{\omega},\tau)\} = \iint e^{i(\boldsymbol{h}'\boldsymbol{\omega}+u\tau)} f(\boldsymbol{\omega},\tau) d\boldsymbol{\omega} d\tau, \tag{1}$$

where for simplicity we usually assume the spectral measure has a density $f(\boldsymbol{\omega},\tau)$ for $\boldsymbol{\omega} \in \mathbb{R}^d$, $\tau \in \mathbb{R}$. The subscripts "S" and "T" denote Fourier transforms with respect to space and time respectively. Taking a "half Fourier transform" of f over the spatial frequency yields an intermediate function

$$H(\boldsymbol{h},\tau) = FT_S\{f(\boldsymbol{\omega},\tau)\} = \int e^{i\boldsymbol{h}'\boldsymbol{\omega}} f(\boldsymbol{\omega},\tau) d\boldsymbol{\omega}, \tag{2}$$

so that

$$C(\boldsymbol{h},u) = FT_T\{H(\boldsymbol{h},\tau)\} = \int e^{iu\tau} H(\boldsymbol{h},\tau) d\tau.$$

We shall call H a "covariance-spectral function" since it depends on the spatial lag $\boldsymbol{h}$ and the temporal frequency τ. Our modelling strategy will be to look for tractable and flexible choices for H.

This paper is organized as follows. In Section 2, general properties and special cases of H are discussed, including Stein's model. A simulation study is done in Section 4. An exploratory analysis of the Irish wind data is carried out in Section 5; this data set provides a test case for the estimation and testing methods developed in Section 3.

2 Stationary spatial-temporal models

2.1 General properties

As described in the Introduction a stationary spatial-temporal covariance structure can be represented equivalently in terms of a covariance function $C(\boldsymbol{h},u)$, a spectral density $f(\boldsymbol{\omega},\tau)$ or a covariance-spectral function $H(\boldsymbol{h},\tau)$. In this section we investigate the relationships between these representations, and explore H in more detail.

The following proposition based on standard Fourier analysis sets out the properties possessed by each of these representations.

Proposition 1 *For an integrable real-valued function $f(\boldsymbol{\omega}, \tau)$ on $\mathbb{R}^d \times \mathbb{R}$, let $C(\boldsymbol{h}, u) = FT_{ST}\{f(\boldsymbol{\omega}, \tau)\}$ and $H(\boldsymbol{h}, \tau) = FT_S\{f(\boldsymbol{\omega}, \tau)\}$. Then the following are equivalent.*

(a) C is an even ($C(\boldsymbol{h}, u) = C(-\boldsymbol{h}, -u)$), real-valued positive semi-definite (p.s.d) function.
(b) f is an even ($f(\boldsymbol{\omega}, \tau) = f(-\boldsymbol{\omega}, -\tau)$) nonnegative function.
(c) $H(\boldsymbol{h}, \tau)$ is an even ($H(\boldsymbol{h}, \tau) = H(-\boldsymbol{h}, -\tau)$) complex-valued function and $H(\boldsymbol{h}, \tau) = \bar{H}(-\boldsymbol{h}, \tau)$. $H(\boldsymbol{h}, \tau)$ is p.s.d. as a function of $\boldsymbol{h}$ for every τ.

Statistical modelling strategies can be based on looking for tractable choices in terms of either C, f or H. In this paper we focus on H. We need to find choices for H satisfying (c) together with an integrability condition $\int H(\mathbf{0}, \tau)d\tau < \infty$.

Recall that for two stationary time series, the coherence function $\rho(\tau)$ gives the complex correlation in the spectral domain between the two series at frequency τ. It is often convenient to express $\rho(\tau) = |\rho(\tau)| \exp\{i\text{Arg}\rho(\tau)\}$ in terms of the absolute coherence function $|\rho(\tau)|$ and the phase $\text{Arg}\rho(\tau)$. The phase of the coherence function determines the extent to which one process leads or lags the other process.

In the spatial-temporal setting, the coherence function also depends on the spatial lag h. It takes a convenient form in terms of the covariance-spectral function,

$$\rho(\boldsymbol{h}, \tau) = \frac{H(\boldsymbol{h}, \tau)}{\sqrt{H(\mathbf{0}, \tau)H(\mathbf{0}, \tau)}} = \frac{H(\boldsymbol{h}, \tau)}{k(\tau)}.$$

For a fixed spatial site $\boldsymbol{s}$, $k(\tau) = H(\mathbf{0}, \tau)$ is the spectral density of the stationary time series $\{Z(\boldsymbol{s}, t),\ t \in \mathbb{R}\}$. For any fixed spatial lag $\boldsymbol{h}$, $\rho(\boldsymbol{h}, \tau)$ is the coherence function of the two stationary time series $\{Z(\boldsymbol{s}, t),\ t \in \mathbb{R}\}$ and $\{Z(\boldsymbol{s} + \boldsymbol{h}, t),\ t \in \mathbb{R}\}$, each with spectral density $k(\tau)$. For each $\tau \in \mathbb{R}$, $\boldsymbol{h} \neq 0$, we have $|\rho(\boldsymbol{h}, \tau)| \leq 1$ and $\rho(\mathbf{0}, \tau) = 1$.

2.2 Special Cases

Separable models. If any of the following three equivalent conditions holds,

$$C(\boldsymbol{h}, u) = C_S(\boldsymbol{h})C_T(u),\ f(\boldsymbol{\omega}, \tau) = f_S(\boldsymbol{\omega})f_T(\tau),\ H(\boldsymbol{h}, \tau) = C_S(\boldsymbol{h})f_T(\tau),$$

then the covariance structure is said to be *separable*. In this case the coherence function takes the form

$$\rho(\boldsymbol{h}, \tau) = \rho_S(\boldsymbol{h}) = C_S(\boldsymbol{h})/C_S(\mathbf{0}),$$

which is real-valued and does not depend on the frequency τ. Further $k(\tau) = C_S(\mathbf{0})f_T(\tau)$. Separability is a convenient mathematical assumption but is usually far too stringent for practical applications.

Fully symmetric models. The assumption of full symmetry is less restrictive than separability, but still imposes strong constraints on the covariance structure. The property of full symmetry can be expressed in three equivalent ways:

$$C(\boldsymbol{h}, u) = C(-\boldsymbol{h}, u),\ f(\boldsymbol{\omega}, \tau) = f(-\boldsymbol{\omega}, \tau),\ H(\boldsymbol{h}, \tau) = H(-\boldsymbol{h}, \tau).$$

Hence, full symmetry is equivalent to the condition that the coherence function $\rho(\boldsymbol{h}, \tau)$ is real-valued. Cressie and Huang (1999) and Gneiting (2002) have chosen real-valued $H(\boldsymbol{h}, \tau)$ for the covariance-spectral representation of spatial-temporal covariance functions; hence they have obtained fully symmetric covariance functions. Subba Rao et al. (2014) investigate the use of the half-spectral representation to carry out approximate maximum likelihood inference under the assumption of full symmetry and isotropy.

Temporal frozen field models. In some sense the opposite of full symmetry is the frozen field model, in which a single time series $Z_T(t)$ with covariance $C_T(u)$ and spectral density $f_T(\tau)$ is observed at each spatial site, but subject to a suitable temporal lag,

$$Z(\boldsymbol{s}, t) = Z_T(t + \boldsymbol{v}'\boldsymbol{s})$$

where $\boldsymbol{v} \in \mathbb{R}^d$ represents a spatial "drift". The covariance, generalized spectral density and covariance-spectral functions of the resulting process take the forms

$$\begin{aligned} C(\boldsymbol{h}, u) &= C_T(u + \boldsymbol{v}'\boldsymbol{h}), \\ f(\boldsymbol{\omega}, \tau) &= f_T(\tau)\delta(\boldsymbol{\omega} - \boldsymbol{v}\tau), \\ H(\boldsymbol{h}, \tau) &= f_T(\tau)e^{i\boldsymbol{v}'\boldsymbol{h}\tau}, \end{aligned} \tag{3}$$

where δ is the Dirac delta function, viewed here as a generalized function of $\boldsymbol{\omega}$ for each τ. In this case the coherence function $\rho(\boldsymbol{h}, \tau) = e^{i\boldsymbol{v}'\boldsymbol{h}\tau}$ has absolute value one for all $\boldsymbol{h}$ and τ, reflecting the coherent dependence of the two time series $\{Z_T(t),\ t \in \mathbb{R}\}$ and $\{Z_T(t + \boldsymbol{v}'\boldsymbol{s}),\ t \in \mathbb{R}\}$. There are two other types of frozen field model, not considered here, which start with a spatial process $Z_S(\boldsymbol{s})$ or a spatial-temporal process $Z_0(\boldsymbol{s}, t)$, respectively; see e.g. Cox and Isham (1988) and Ma (2003).

2.3 Stein's Covariance-Spectral Model

The temporal frozen field model (3) is too rigid in practice, so we consider a more general model due to Stein (2005)

$$H(\boldsymbol{h}, \tau) = k(\tau)D(\boldsymbol{h}\gamma(\tau))e^{i\theta(\tau)\boldsymbol{v}'\boldsymbol{h}}, \tag{4}$$

where various components have the following interpretations.

(a) The temporal spectral density $k(\tau) = k(-\tau)$ is a symmetric nonnegative integrable function on $\mathbb{R}$. It is identical to $f_T(\tau)$ in (3).

(b) The latent spatial covariance function $D(\boldsymbol{h}) = D(-\boldsymbol{h})$ is a real-valued positive definite function on $\mathbb{R}^d$ with $D(\mathbf{0}) = 1$. Note that the coherence function of (4) is

$$\rho(\boldsymbol{h}, \tau) = D(\boldsymbol{h}\gamma(\tau))e^{i\theta(\tau)\boldsymbol{v}'\boldsymbol{h}}, \tag{5}$$

with absolute value $|\rho(\boldsymbol{h}, \tau)| = D(\boldsymbol{h}\gamma(\tau)) \leq 1$ for $\boldsymbol{h} \neq \mathbf{0}$. Thus D governs how the absolute coherence decays with increasing spatial lag up to a factor depending on temporal frequency. For practical work we follow Stein (2005) and assume D takes the specific form

$$D(\boldsymbol{h}) = e^{-|\boldsymbol{h}|^p}, \ 0 < p \leq 2, \tag{6}$$

where p is generally unknown, in order to simplify the estimation of the remaining parts of the model.

(c) The temporal decay rate function $\gamma(\tau) = \gamma(-\tau)$ is a positive even function of $\tau \in \mathbb{R}$. It governs how the rate of decay of absolute coherence in $\boldsymbol{h}$ depends on the temporal frequency τ. The simplest choice is the constant function $\gamma(\tau) = \text{const}$.

(d) The temporal phase rate function $\theta(\tau) = -\theta(-\tau)$ is an odd function on $\mathbb{R}$. It governs how phase of the coherence function, $\text{Arg}\rho(\boldsymbol{h}, \tau) = \theta(\tau)\boldsymbol{v}'\boldsymbol{h}$ depends on temporal frequency. The simplest choice is the linear function $\theta(\tau) = b\tau$.

(e) Finally the unit vector $\boldsymbol{v}$ specifies a direction for the spatial-temporal asymmetry.

The corresponding spectral density of (4) is

$$f(\boldsymbol{\omega}, \tau) = \frac{k(\tau)}{\gamma(\tau)} f_S\left(\frac{\boldsymbol{\omega} - \boldsymbol{v}\theta(\tau)}{\gamma(\tau)}\right), \tag{7}$$

where $f_S(\boldsymbol{\omega})$ is the purely spatial spectral density of the spatial covariance function $D(\boldsymbol{h})$. The vector $\boldsymbol{v}\theta(\tau)$ in the phase shift of the coherence function (4) appears as the location shift of the spectral density (7) and the scaling function $\gamma(\tau)$ appears as a scale shift.

One way to motivate a simple special case of Stein's model is through a combination of a separable model and a temporally frozen model. If $D(\boldsymbol{h}) = C_S(\boldsymbol{h})$, $\gamma(\tau) = 1$, $k(\tau) = f_T(\tau)$ and $\theta(\tau) = \tau$ in (4), then

$$\begin{aligned} C(\boldsymbol{h}, u) &= C_S(\boldsymbol{h})C_T(u + \boldsymbol{v}'\boldsymbol{h}), \\ f(\boldsymbol{\omega}, \tau) &= f_S(\boldsymbol{\omega} - \boldsymbol{v}\tau)f_T(\tau), \\ H(\boldsymbol{h}, \tau) &= f_T(\tau)C_S(\boldsymbol{h})e^{i\tau\boldsymbol{v}'\boldsymbol{h}}. \end{aligned} \tag{8}$$

However, Stein's approach in (4) allows a greater degree of flexibility by allowing more choices for $\gamma(\tau)$ and $\theta(\tau)$.

Many statistical tests for separability have been proposed recently based on parametric models, likelihood ratio tests and spectral methods, e.g. Fuentes (2006) and Mitchell et al. (2006). In the purely spatial context, Scaccia and

Martin (2005) and Lu and Zimmerman (2002) developed tests for axial symmetry and diagonal symmetry. These tests are valid only under a full symmetry assumption. A lack of full symmetry in Stein's model can be carried out by examining whether $\theta(\tau) = 0$. If $\theta(\tau) = 0$, then the resulting spatial-temporal covariance function is fully symmetric. Furthermore, if $\theta(\tau) = 0$ and $\gamma(\tau) = \text{const.}$, then the model is separable.

3 Inference for Stein's covariance-spectral model

In this section we investigate methods of inference for Stein's covariance-spectral model (4). The parameters are $p,\ 0 < p \leq 2$ in the latent spatial covariance function (5) and three functional parameters $k(\tau)$, $\gamma(\tau)$ and $\theta(\tau)$. The goals are to develop methods for parameter estimation, goodness of fit assessment and interpretation.

The estimation procedure has several steps, so it is helpful to set out the general strategy and notation before the details are given.

(a) Starting from the data, construct the empirical covariance-spectral function $\tilde{H}(\boldsymbol{h}, \tau)$, as a raw summary statistic of the data.
(b) Carry out preliminary nonparametric smoothing of $\tilde{H}(\boldsymbol{h}, \tau)$ over τ to get a smoothed empirical covariance-spectral function $\tilde{\tilde{H}}(\boldsymbol{h}, \tau)$.
(c) Transform the H function to be linear in the unknown parameters, and use regression analysis (with the transformed $\tilde{\tilde{H}}(\boldsymbol{h}, \tau)$ playing the role of the response variable) to estimate the parameters. This strategy is used first to estimate $k(\tau)$ and second to estimate jointly p and $\gamma(\tau)$. The procedure to estimate $\theta(\tau)$ follows the same general principles, but involves a preliminary estimate of $\boldsymbol{v}$ based on maximizing a certain ratio of quadratic forms.
In other words the basic strategy is to estimate $k(\tau)$, $\gamma(\tau)$ and $\theta(\tau)$ is to match the smoothed empirical covariance-spectral function $\tilde{\tilde{H}}(\boldsymbol{h}, \tau)$ to its theoretical value in (4) using regression methods, after transforming (4) to linearize the dependence of $k(\tau)$, $\gamma(\tau)$ and $\theta(\tau)$ in turn, on τ. The regression can be either parametric or non-parametric. Stein (2005) has suggested parametric forms based on trigonometric polynomials for $k(\tau)$, $\gamma(\tau)$ and $\theta(\tau)$. Maximizing the likelihood numerically is a computationally intensive procedure due to the need to invert large matrices. Stein (2005) considered approximate likelihoods based on a multivariate version of the Whittle likelihood. Our method can be seen as simpler and more graphical, enabling visual judgments to be made about the model. Stein (2005) constructed various plots to assess the goodness of fit of the model. We use similar plots to estimate the parameters through a regression analysis. Both parametric and nonparametric regression models are accommodated by this methodology. The effect and importance of each parameter then can be seen directly in the appropriate plot.
(d) Finally, it is necessary to estimate standard errors of the parameters.

3.1 Initial data processing

Suppose the data $\{Z(\boldsymbol{s}_i, t),\ t = 1, \ldots, T\}$ are given at irregular spatial locations $\boldsymbol{s}_i,\ i = 1, \ldots, S$ and equally-spaced integer times $t = 1, \ldots, T$. For notational convenience suppose the data have already been centered to have mean 0. The first step is to construct the half-Fourier transform of the data

$$J(\boldsymbol{s}_i, \tau) = \sum_{t=1}^{T} Z(\boldsymbol{s}_i, t) e^{-2\pi i t \tau},\ \tau = 1/T, \ldots, [T/2]/T.$$

To distinguish the empirical and smoothed versions of various quantities, we use " ˜ " and "$\tilde{}$" to denote the empirical and the initial smoothed versions, respectively.

The sample covariance-spectral function is defined to be

$$\tilde{H}(\boldsymbol{h}, \tau) = \frac{1}{T} J(\boldsymbol{s}_i, \tau) \bar{J}(\boldsymbol{s}_j, \tau),\ \boldsymbol{h} = \boldsymbol{s}_i - \boldsymbol{s}_j \neq \boldsymbol{0},$$
$$\tilde{H}(\boldsymbol{0}, \tau) = \frac{1}{S} \sum_{i=1}^{S} \tilde{H}_i(\boldsymbol{0}, \tau),$$

where $\bar{J}$ is the complex conjugate of J and $\tilde{H}_i(\boldsymbol{0}, \tau) = \frac{1}{T}|J(\boldsymbol{s}_i, \tau)|^2$. Here $\boldsymbol{h}$ ranges through the set of spatial lags $\boldsymbol{s}_i - \boldsymbol{s}_j$, assumed for simplicity to have no replication except for $\boldsymbol{h} = \boldsymbol{0}$. The sample temporal spectral density at site $\boldsymbol{s}_i$ is defined by $\tilde{k}_i(\tau) = \tilde{H}_i(\boldsymbol{0}, \tau)$ and the sample overall temporal spectral density by

$$\tilde{k}(\tau) = \tilde{H}(\boldsymbol{0}, \tau). \tag{9}$$

Define the sample coherence and the sample phase for the process at two sites separated by spatial lag $\boldsymbol{h} = \boldsymbol{s}_i - \boldsymbol{s}_j \neq 0$ by

$$\tilde{\rho}(\boldsymbol{h}, \tau) = \tilde{H}(\boldsymbol{h}, \tau) / \sqrt{\tilde{H}_i(\boldsymbol{0}, \tau) \tilde{H}_j(\boldsymbol{0}, \tau)},\ \boldsymbol{h} = \boldsymbol{s}_i - \boldsymbol{s}_j \neq \boldsymbol{0} \tag{10}$$
$$\tilde{D}(\boldsymbol{h}, \tau) = |\tilde{\rho}(\boldsymbol{h}, \tau)|, \qquad \tilde{g}(\boldsymbol{h}, \tau) = \mathrm{Arg}(\tilde{\rho}(\boldsymbol{h}, \tau)), \tag{11}$$

so $\tilde{\rho}(\boldsymbol{h}, \tau) = \tilde{D}(\boldsymbol{h}, \tau) \tilde{g}(\boldsymbol{h}, \tau)$.

3.2 Initial smoothing

Note that $\tilde{D}(\boldsymbol{h}, \tau) = |\tilde{\rho}(\boldsymbol{h}, \tau)| = 1$ for all $\boldsymbol{h} = \boldsymbol{s}_i - \boldsymbol{s}_j \neq \boldsymbol{0}$, making it useless as it stands for the estimation of $D(\boldsymbol{h}, \tau)$. To fix this problem, we propose that some initial smoothing of $\tilde{H}(\boldsymbol{h}, \tau)$ with respect to the time frequency be carried out. Denote the resulting "initially smoothed" sample covariance-spectral function by $\tilde{\tilde{H}}(\boldsymbol{h}, \tau)$. Similarly, the corresponding initially smoothed empirical temporal spectral density $\tilde{\tilde{k}}(\tau)$, absolute phase $\tilde{\tilde{D}}(\boldsymbol{h}, \tau)$ and argument of phase $\tilde{\tilde{g}}(\boldsymbol{h}, \tau)$ are obtained by using $\tilde{\tilde{H}}(\boldsymbol{h}, \tau)$ instead of $\tilde{H}(\boldsymbol{h}, \tau)$ in equations (9)–(11).

One way to carry out the initial smoothing is to use the R function, `spec.pgram`, which smooths $\tilde{H}(\boldsymbol{h},\tau)$ with a series of modified Daniell smoothers. The exact amount of initial smoothing is not critical here. Enough smoothing is needed to make $\tilde{\tilde{D}}(\boldsymbol{h},\tau)$ suitable for estimation purposes. At the same time we do not want to mask any broad patterns in the data which will be fitted later using parametric or nonparametric models. An explicit illustration of a suitable amount of smoothing is given in Section 5 for the Irish wind data. It is also possible to include tapering over time but for simplicity we have not done so here.

3.3 Estimation of $k(\tau)$

The initially smoothed temporal spectral density $\tilde{\tilde{k}}(\tau)$ is a crude estimate of $k(\tau)$. This estimate can be refined in two ways, depending on whether we carry out nonparametric or parametric modelling.

(i) (nonparametric) A simple way to estimate the function $k(\tau)$ is simply by regressing $\tilde{\tilde{k}}(\tau)$ on τ nonparametrically, e.g. using the Nadaraya-Watson (Nadaraya, 1964; Watson, 1964) estimator. A convenient implementation in R is given by the functions `dpill` and `locpoly` in the package `KernSmooth` (Ruppert et al., 1995). Let $\hat{k}_{\mathrm{np}}(\tau)$ define the fitted nonparametric estimate.

(ii) (parametric) Stein (2005) suggested a parametric fractional exponential model

$$\log k(\tau) = -\beta \log \sin(|\pi\tau|) + \sum_{k=0}^{K_1} c_k \cos(2\pi k\tau), \qquad \tau = 1/T, \ldots, [T/2]/T \tag{12}$$

where the condition $0 \le \beta < 1$ guarantees the integrability of $k(\tau)$ and allows for long-range dependence (Bloomfield, 1973; Beran, 1994, Ch. 6). Here K_1 is assumed known for the moment. Assume $\log \tilde{\tilde{k}}(\tau)$ equals the right hand side (RHS) of (12) plus independently and identically distributed (i.i.d.) errors. Then OLS regression yields estimates $\hat{\beta}, \hat{c}_0, \ldots, \hat{c}_{K_1}$, which define a fitted spectral density $\hat{k}_{\mathrm{par}}(\tau)$.

3.4 Estimation of p and $\gamma(\tau)$

Under assumptions (4) and (6) the following transformation linearizes the dependence of $D(\boldsymbol{h},\tau) = |\rho(\boldsymbol{h},\tau)|$ on the parameters p and $\gamma(\tau)$:

$$\log(-\log(D(\boldsymbol{h},\tau))) = p\log(|\boldsymbol{h}|) + p\log\gamma(\tau). \tag{13}$$

In terms of the data, we shall treat $\log(-\log(\tilde{\tilde{D}}(\boldsymbol{h},\tau)))$ as the dependent variable in a regression on the righthand side of (13) with i.i.d. normal errors. Estimation can take two forms, depending on whether we carry out nonparametric or parametric modelling.

(i) (nonparametric) We propose estimation in two stages. Initially treat (13) as a parallel-lines regression model on $\log(|\boldsymbol{h}|)$, with common slope p and with intercepts $p \log \gamma(\tau)$ depending on τ. Fit the parameters by OLS and denote the resulting estimate of $\gamma(\tau)$ by $\hat{\gamma}_{\text{init}}(\tau)$. For the second stage, regress $\hat{\gamma}_{\text{init}}(\tau)$ on τ nonparametrically to get $\hat{\gamma}_{\text{np}}(\tau)$ and $\hat{D}_{\text{np}}(\boldsymbol{h}, \tau)$.

(ii) (parametric) Following Stein (2005), one way to model the even nonnegative function $\gamma(\tau)$ is with trigonometric polynomials,

$$\log \gamma(\tau) = \sum_{k=0}^{K_2} a_k \cos(2\pi k \tau), \tag{14}$$

where K_2 is a pre-specified number of terms. Then the log-log transformation linearizes the dependence of $D(\boldsymbol{h}, \tau)$ on the parameters p and $a_0, \ldots, a_{K_2}$.

$$\log(-\log(D(\boldsymbol{h}, \tau))) = p \log(|\boldsymbol{h}|) + p \sum_{k=0}^{K_2} a_k \cos(2\pi k \tau). \tag{15}$$

Here (15) depends linearly on $\log(|\boldsymbol{h}|)$ and $\cos(2\pi k \tau),\ k = 0, \ldots K_2$, where the slope p is the same for all τ. Fitting this model by OLS leads to estimates $\hat{p}$ and $\hat{a}_0, \ldots, \hat{a}_{K_2}$ which define $\hat{\gamma}_{\text{par}}(\tau)$ and $\hat{D}_{\text{par}}(\boldsymbol{h}, \tau)$.

3.5 Estimation of "drift" direction $\boldsymbol{v}$ and the phase $\theta(\tau)$

In (4), recall

$$g(\boldsymbol{h}, \tau) = \text{Arg}(H(\boldsymbol{h}, \tau)) = \theta(\tau) \boldsymbol{v}' \boldsymbol{h}. \tag{16}$$

We propose estimating the parameters by regressing the smoothed empirical phase function $\tilde{\tilde{g}}(\boldsymbol{h}, \tau)$ on $\boldsymbol{h}$ and τ using ordinary least squares. However, there are two complications: $g(\boldsymbol{h}, \tau)$ is an angle, not a number, and the regression is nonlinear.

First we deal with the angular problem; that is, $g(\boldsymbol{h}, \tau)$ is an angular variable for each $\boldsymbol{h}$ and τ, and hence defined only up to an integer multiple of 2π. But since $g(\boldsymbol{h}, \tau)$ is a continuous function of τ for each $\boldsymbol{h}$, it can also be regarded as a real-valued function initialized by $\tilde{g}(\boldsymbol{h}, 0) = 0$. That is, for each $\boldsymbol{h}$ and τ an unambiguous choice for the winding number can be found. Let $g_R(\boldsymbol{h}, \tau)$ denote this real-valued extension of $g(\boldsymbol{h}, \tau)$. Similarly, provided the noise is not too large, the empirical phase $\tilde{g}(\boldsymbol{h}, \tau)$ and its smoothed version $\tilde{\tilde{g}}(\boldsymbol{h}, \tau)$ can be unambiguously unwound to give real-valued extensions $\tilde{g}_R(\boldsymbol{h}, \tau)$ and $\tilde{\tilde{g}}_R(\boldsymbol{h}, \tau)$.

Next we regress $\tilde{\tilde{g}}_R(\boldsymbol{h}, \tau)$ on $\theta(\tau) \boldsymbol{v}' \boldsymbol{h}$ on $\boldsymbol{v}$ and $\theta(\tau)$. Since the regression is nonlinear we proceed in two stages. The first stage produces an estimate of $\boldsymbol{v}$ and an initial estimate of $\theta(\tau)$. The second stage produces a more refined estimate of $\theta(\tau)$.

Here are the details. Ordinary least squares estimation involves minimizing the sum of squares

$$SSE = \sum_{\tau} \sum_{\boldsymbol{h} \neq \boldsymbol{0}} \left(\tilde{\tilde{g}}_R(\boldsymbol{h}, \tau) - \theta(\tau) \boldsymbol{v}' \boldsymbol{h} \right)^2 . \tag{17}$$

If $\boldsymbol{v}$ is known, then for each fixed τ, the OLS estimate of $\theta(\tau)$ is given by

$$\hat{\theta}_{\text{init}}(\tau; \boldsymbol{v}) = \frac{\boldsymbol{v}' \sum_{\boldsymbol{h} \neq \boldsymbol{0}} \tilde{\tilde{g}}_R(\boldsymbol{h}, \tau) \boldsymbol{h}}{\boldsymbol{v}' A \boldsymbol{v}}, \tag{18}$$

where $A = \sum_{\boldsymbol{h} \neq \boldsymbol{0}} \boldsymbol{h}\boldsymbol{h}'$ is a $d \times d$ matrix. Inserting (18) into (17) yields the reduced sum of squares

$$SSE(\boldsymbol{v}) = \sum_{\tau} \sum_{\boldsymbol{h} \neq \boldsymbol{0}} \tilde{\tilde{g}}_R(\boldsymbol{h}, \tau)^2 - \frac{\boldsymbol{v}' B \boldsymbol{v}}{\boldsymbol{v}' A \boldsymbol{v}}, \tag{19}$$

where $B = \sum_{\tau} \boldsymbol{\beta}(\tau) \boldsymbol{\beta}'(\tau)$ is a $d \times d$ matrix defined in term of the d-dimensional vector $\boldsymbol{\beta}(\tau) = \sum_{\boldsymbol{h}} \tilde{\tilde{g}}_R(\boldsymbol{h}, \tau) \boldsymbol{h}$.

Minimizing (19) now reduces to an optimization problem for a ratio of quadratic forms. The optimal $\boldsymbol{v}$ is given by the eigenvector corresponding to the largest eigenvalue of $A^{-1}B$ (e.g. Mardia et al. (1979, p. 479)). Let $\hat{\boldsymbol{v}}$ denote the result. Once $\boldsymbol{v}$ has been estimated, then the regression equation reduces to

$$g_R(\boldsymbol{h}, \tau) = \theta(\tau) \hat{\boldsymbol{v}}' \boldsymbol{h}$$

and the initial estimate of $\theta(\tau)$ becomes

$$\hat{\theta}_{\text{init}}(\tau) = \frac{\hat{\boldsymbol{v}}' \sum_{\boldsymbol{h} \neq \boldsymbol{0}} \tilde{\tilde{g}}_R(\boldsymbol{h}, \tau) \boldsymbol{h}}{\hat{\boldsymbol{v}}' A \hat{\boldsymbol{v}}}.$$

We can get a refined estimate of $\theta(\tau)$ as follows.

(i) (nonparametric) Regress $\hat{\theta}_{\text{init}}(\tau)$ on τ nonparametrically to get $\hat{\theta}_{\text{np}}(\tau)$.

(ii) (parametric) Regress $\hat{\theta}_{\text{init}}(\tau)$ on τ parametrically to get $\hat{\theta}_{\text{par}}(\tau)$. Following Stein (2005), one way to model the odd function $\theta(\tau)$ is with trigonometric polynomials,

$$\theta(\tau) = \sum_{k=1}^{K_3} b_k \sin(2\pi k \tau), \tag{20}$$

for some fixed K_3.

3.6 Estimation of Standard Errors

In estimating both functional parameters $\gamma(\tau)$ and $\theta(\tau)$ we must do initial smoothing. In the first case we do initial smoothing to ensure that the absolute coherence is less than one and in the latter case to ensure the winding of the phase angle varies smoothly with τ. Although there is no need for initial smoothing of the empirical spectral density, we have initially smoothed the empirical spectral density to unify our estimation procedures.

But initial smoothing leads to autocorrelated errors, underestimated standard errors and minor shift in the intercepts. Although ignoring correlation usually introduces little bias in the estimates of regression coefficients, it can introduce substantial bias in the estimates of standard errors and this may lead to incorrect inferences about OLS estimates. To resolve this problem, we assume an approximate model for the autocorrelation of the initially smoothed residuals. The general procedure can be described as follows. Consider a general linear regression model $\boldsymbol{Y} = X\boldsymbol{\beta}+\boldsymbol{\varepsilon}, E(\boldsymbol{\varepsilon}) = \mathbf{0}, \mathrm{cov}(\boldsymbol{Y}) = \Sigma$, where $\boldsymbol{Y}$ is an $SF\times 1$ random vector where $F = [T/2]$ and $\boldsymbol{\beta}$ is a vector of unknown regression parameters. Under the separability assumption we have $\Sigma = \Sigma_S \otimes \Sigma_F$, where $\Sigma_S = (\sigma_{i,j})_{i,j=1}^{S}$ and $\Sigma_F = (\delta_{i,j})_{i,j=1}^{F}$ are covariance matrices and $\otimes$ denotes the Kronecker product. The OLS estimates are given by $\hat{\boldsymbol{\beta}} = (X'X)^{-1}X'\boldsymbol{Y}$.

Under the separability assumption we have $\mathrm{cov}(Y_{i_1,j_1}, Y_{i_2,j_2}) = \sigma_{i_1,i_2}\delta_{j_2-j_1}$. Let $\hat{\boldsymbol{\varepsilon}}$ be the $SF \times 1$ vector of the corresponding OLS residuals. The corresponding estimates are given by

$$\hat{\delta}_u = \frac{1}{SF}\sum_i \sum_{j=1}^{F-u} \hat{\varepsilon}_{i,j}\hat{\varepsilon}_{i,j+u}/\hat{\sigma}_{i,i},$$
$$\hat{\sigma}_{i_1,i_2} = \frac{1}{F}\sum_j \hat{\varepsilon}_{i_1,j}\hat{\varepsilon}_{i_2,j},$$

where without loss of generality we scale the covariance matrices so that $\hat{\delta}_0 = 1$. The variance of regression coefficients are given by

$$\mathrm{var}(\hat{\boldsymbol{\beta}}) = (X'X)^{-1}X'(\hat{\Sigma}_S \otimes \hat{\Sigma}_F)X(X'X)^{-1}.$$

Note that for large values of S and F calculating the Kronecker product $\hat{\Sigma}_S \otimes \hat{\Sigma}_F$ first and then multiplying it by X is not only computationally inefficient, it is also not feasible due to memory problems in R. To resolve this problem we use the fact that $\hat{\Sigma}_S \otimes \hat{\Sigma}_F = (\hat{\Sigma}_S \otimes I_F)(I_S \otimes \hat{\Sigma}_F)$. These identity matrices make the calculation easier because it is easy to see that sub-blocks of X are multiplied by $\hat{\Sigma}_F$ so we multiply each block separately. Also, for numerical stability we approximate $\hat{\Sigma}_F$ by a Toeplitz matrix based on a stationary AR(1) process.

3.7 General Considerations

- When fitting trigonometric polynomials for functional parameters $\kappa(\tau)$, $\gamma(\tau)$ and $\theta(\tau)$, it is necessary to choose values for K_1, K_2, K_3. This choice can be made subjectively or by some model selection criteria such as AIC or BIC. In our example we will use AIC.
- We can construct pointwise confidence intervals for $\kappa(\tau)$, $\gamma(\tau)$ and $\theta(\tau)$ using asymptotic properties of the spectral density, coherence and phase functions (Bloomfield, 1976).

4 A Simple Example: Bivariate Stationary Processes

In order to develop some intuition for the coherence function under Stein's model we look at a simplified model involving just 2 sites and discrete time. Suppose that $\{y_t\}$ is a latent stationary Gaussian process with covariance function

$$C_{yy}(u) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f_{yy}(\tau)\cos(u\tau)d\tau,\ u \in \mathbb{Z}$$

where $f_{yy}(\tau)$ is the spectral density function of the latent process $\{y_t\}$. Suppose that $\{z_t^1\}$ and $\{z_t^2\}, t = 1, \ldots, n$ are observed processes

$$z_t^1 = y_t + \varepsilon_t^1,\ z_t^2 = y_{t+\ell} + \varepsilon_t^2, \tag{21}$$

where $\{\varepsilon_t^1\}$ and $\{\varepsilon_t^2\}$ are independent white noise processes with variance σ^2, and the integer ℓ is a delay parameter. The cross-covariance function of the two processes is $C_{zz}^{12}(u) = E(z_t^1 z_{t+u}^2) = C_{yy}(u-\ell)$. The spectral and the cross-spectral densities for the observed processes are

$$\begin{aligned} f_{zz}^1(\tau) &= f_{zz}^2(\tau) = f_{yy}(\tau) + \sigma^2, \\ f_{zz}^{12}(\tau) &= \exp\{i\ell\tau\} f_{yy}(\tau), \end{aligned} \tag{22}$$

and so the coherence function is

$$\rho(\tau) = \frac{f_{zz}^{12}(\tau)}{\sqrt{f_{zz}^1(\tau) f_{zz}^2(\tau)}} = \frac{\exp\{i\ell\tau\}}{1 + \sigma^2/f_{yy}(\tau)}. \tag{23}$$

Note how the addition of noise to the bivariate time series produces an absolute coherence less than one. The phase is $g(\ell, \tau) = \ell\tau$, where the slope ℓ corresponds to the delay between two processes.

To illustrate these functions and their estimates from data, let $\{y_t\}$ be an AR(1) process, with $C_{yy}(u) = \sigma_0^2 e^{-\alpha|u|}$ and spectral density $f_{yy}(\tau) = \sigma_0^2\{1 - 2\alpha\cos(\tau) + \alpha^2\}^{-1}$. Figure 1(a and b) show one realization of the model defined by (21) with $\alpha = 0.5$, $\sigma_0^2 = \sigma^2 = 1$ and $\ell = 4$ observed at $n = 1000$ times. The theoretical spectrum and the smoothed periodograms of $\{z_t^1\}$ and $\{z_t^2\}$ are shown in Figure 1(c and d). The Nadaraya-Watson estimate (Nadaraya, 1964; Watson, 1964) is used to smooth the periodograms. The smoothing parameter

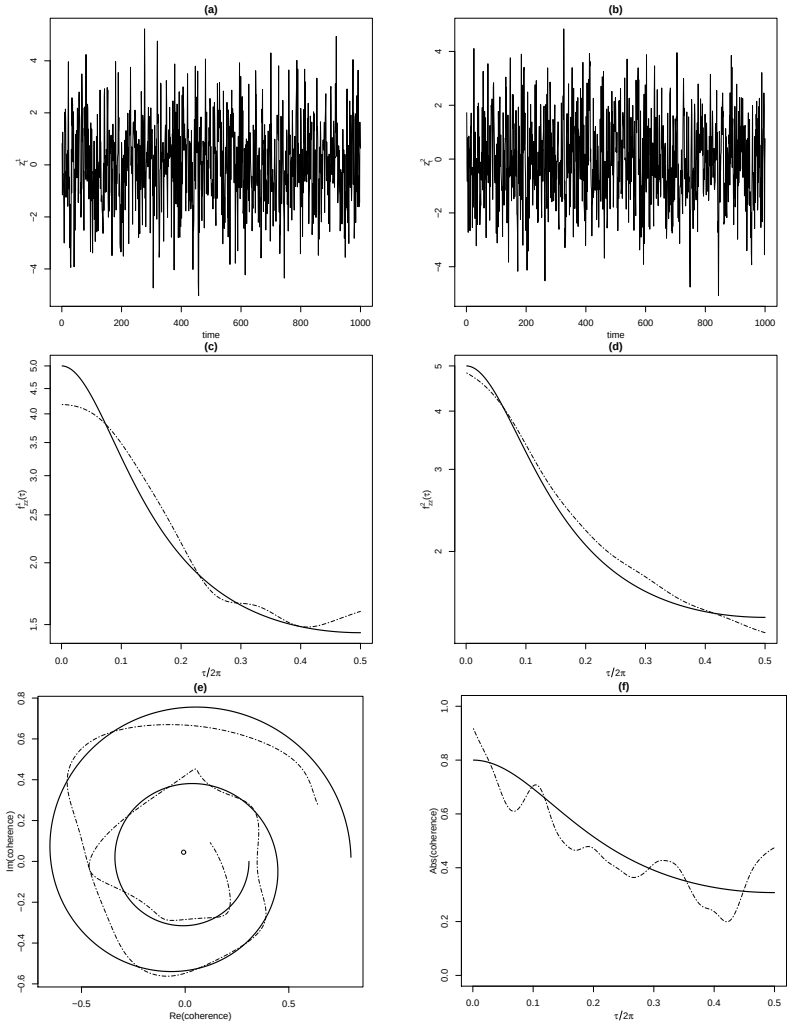

Fig. 1 Simulation of the bivariate time series (21) with $\alpha = 0.5$, $\sigma_0^2 = \sigma^2 = 1$ and $\ell = 4$. (a) and (b): Time series $\{z_t^1\}$ and $\{z_t^2\}$; (c) and (d): Theoretical spectral density (solid lines) defined by (22) and smoothed periodograms (dot-dash lines) of $\{z_t^1\}$ and $\{z_t^2\}$; (e) and (f): Theoretical (solid lines) and smoothed complex (dot-dash lines) coherence and theoretical (solid lines) and absolute coherence (dot-dash lines).

is obtained by the function `dpill` in R. This function computes a direct plug-in estimator of the bandwidth, as described by Ruppert et al. (1995) which tries to minimize MSE.

Figure 1 (e and f), show the smoothed sample coherence between $\{z_t^1\}$ and $\{z_t^2\}$ together with the theoretical coherence derived from (23). Note that the coherence is strong at low frequencies; its absolute values decreases and the phase spirals anti-clockwise towards zero as the frequency increases.

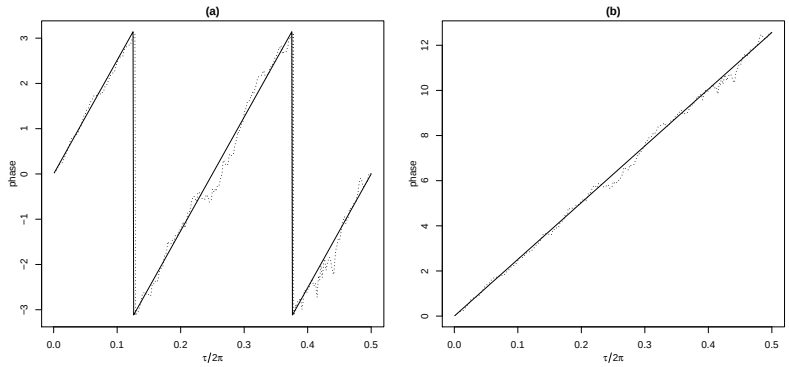

Fig. 2 Phase spectrum of the simulated bivariate time series (21) with $\alpha = 0.5,\ a = 1,\ b = 1,\ \sigma_0^2 = \sigma_1^2 = \sigma_2^2 = 1$ and $\ell = 4$. (a): the phase angle is plotted naively as a number in $(-\pi, \pi)$; (b): by taking winding into account, a real-valued continuous version of the phase has been plotted. The solid line represents the theoretical phase spectrum of $\{z_t^1\}$ and $\{z_t^2\}$. The dotted line represent the smoothed phase.

In Figure 2(a), the phase angle is plotted naively as a number in $(-\pi, \pi)$ and hence shows discontinuity as a function of τ. By taking winding into account, a version of the phase function which is continuous across the boundaries $\pm\pi$ can be defined and estimated. This has been presented in Figure 2(b) and shows that the underlying phase is linear in τ. This example illustrates the types of pattern to look for in real data.

5 Application to the Irish Wind Data

The Irish wind data set is used here to provide a test case for the estimation and testing methods developed in Section 3. The data consist of average daily wind speeds (meters per second) measured at 11 synoptic meteorological stations located in the Republic of Ireland during the period 1961-78, with 6,574 observations per location. Following Haslett and Raftery (1989) we take a square root transformation to stabilize the variance over both stations and time periods for each day of the year and subtract the seasonal effect from the data.

Gneiting (2002), by plotting spatial-temporal correlations for different spatial and temporal lags, has shown that the wind speeds measured at different stations are highly correlated and the correlations decay substantially as spatial or temporal lag increases. De Luna and Genton (2005) by plotting the correlation function in different directions concluded that the data have an isotropic correlation structure.

Gneiting (2002) used this data set to illustrate the lack of the separability and full symmetry assumptions. Indeed winds in Ireland are predominantly westerly; hence for different temporal lags the west-to-east correlation of wind speed of stations will be higher than the east-to-west correlation; that is $C(\boldsymbol{h}, u) > C(-\boldsymbol{h}, u);\ \boldsymbol{h}' = (h_1, 0),\ h_1 > 0, u > 0.$

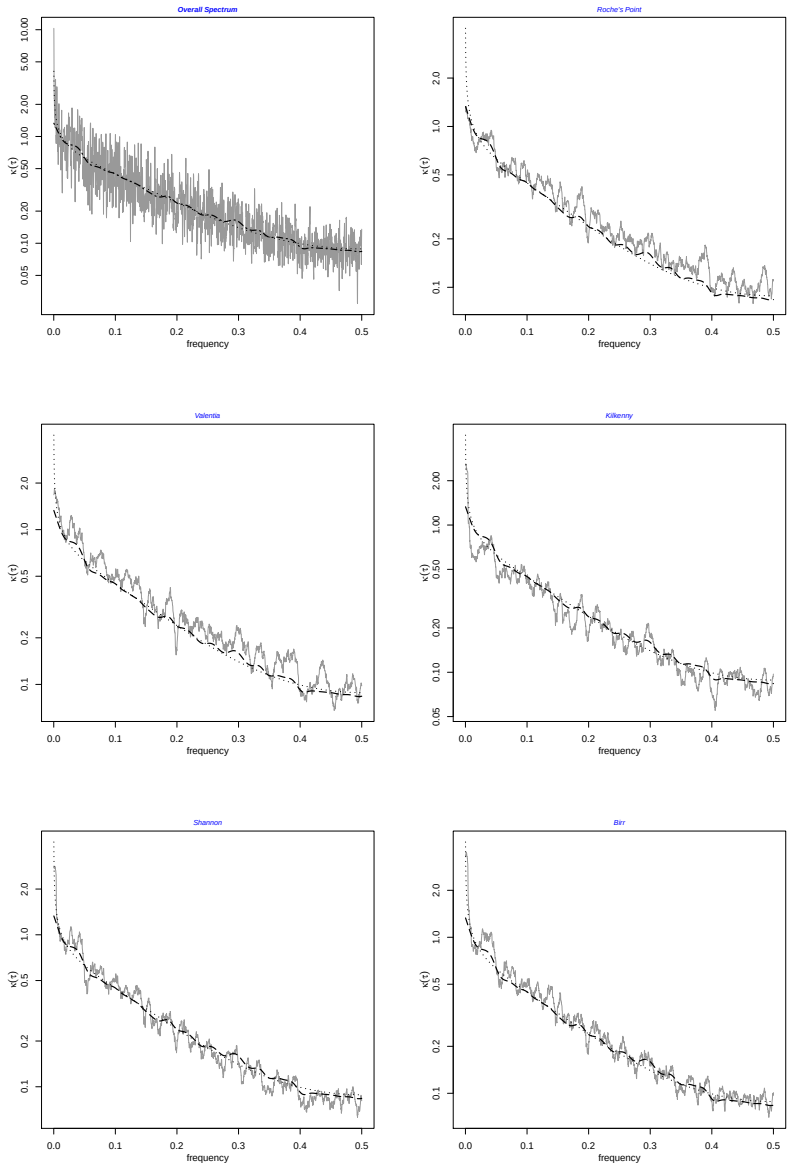

Fig. 3 Smoothed empirical $\tilde{\tilde{k}}_i(\tau)$ (gray curves), parametric estimate $\hat{k}_{\rm par}(\tau)$ (dotted curves), and nonparametric Nadaraya-Watson estimate $\hat{k}_{\rm np}(\tau)$ (long-dash curves) versus frequency for the 11 individual stations and their average marginal spectral, $\tilde{\tilde{k}}(\tau)$ (first plot), by the standard program `spec.pgram` in `R` with the span set to 5 for average marginal spectra and 55 for individual marginal spectra.

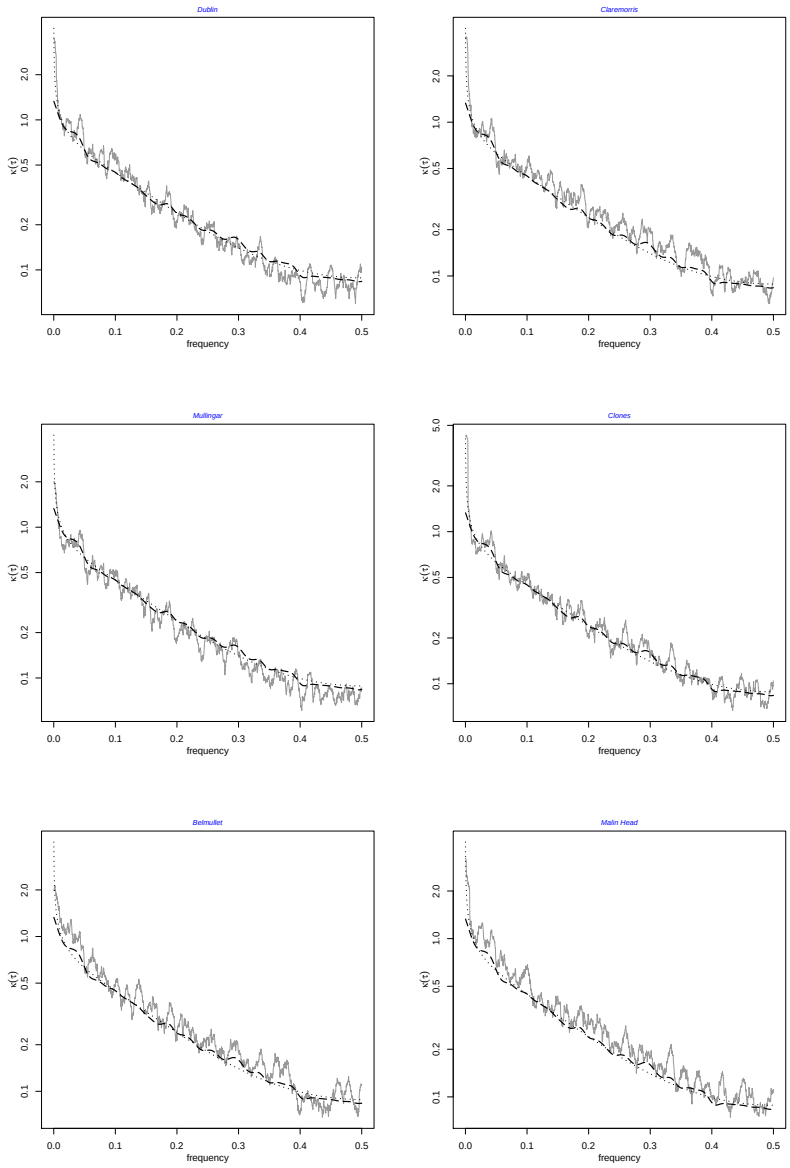

Fig. 4 Smoothed empirical $\tilde{\tilde{k}}_i(\tau)$ (gray curves), parametric estimate $\hat{k}_{\mathrm{par}}(\tau)$ (dotted curves), and nonparametric Nadaraya-Watson estimate $\hat{k}_{\mathrm{np}}(\tau)$ (long-dash curves) versus frequency for the 11 individual stations and their average marginal spectral, $\tilde{\tilde{k}}(\tau)$ (first plot), by the standard program `spec.pgram` in `R` with the span set to 5 for average marginal spectra and 55 for individual marginal spectra.

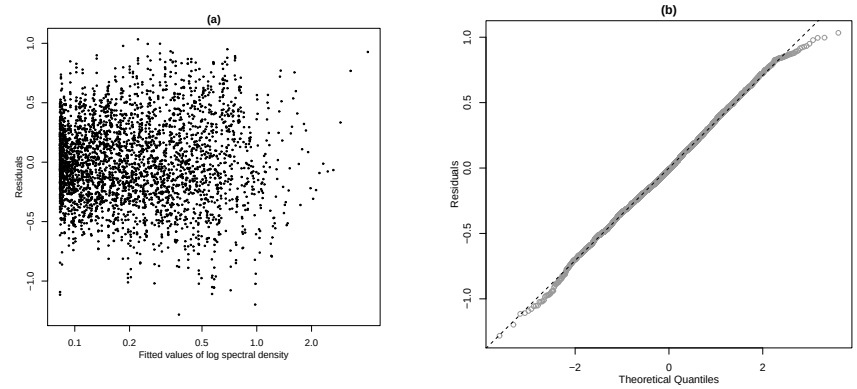

Fig. 5 (a): Residuals versus logarithm of fitted values of spectral density, $\log \hat{k}_{\mathrm{par}}(\tau)$. (b) normal Q-Q plot of residuals.

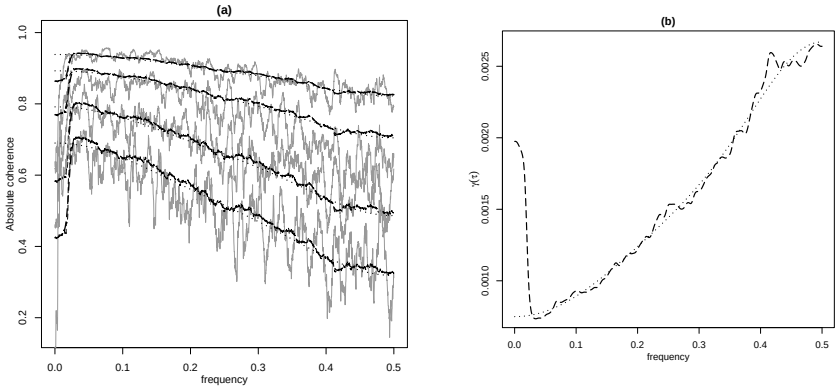

Fig. 6 (a): plot of empirical and estimated absolute coherence for different stations; from top: Birr-Mullingar: $|\boldsymbol{h}| = 60.68$ km, Birr-Dublin: $|\boldsymbol{h}| = 115.40$ km, Birr-Malin Head: $|\boldsymbol{h}| = 256.40$ km and Valentia-Malin Head: $|\boldsymbol{h}| = 427.34$ km. (b): plot of estimated $\gamma(\tau)$. Empirical estimate (gray curve), Nadaraya-Watson estimate (long-dash curve) and trigonometric regression estimation (dotted curve). Empirical estimate calculated by the standard program `spec.pgram` in `R` with the span set to 255.

We now fit the Stein's asymmetric model (4), with unknown parameters $k(\tau)$, $\gamma(\tau)$, $\theta(\tau)$, $\boldsymbol{v}$ and p to the Irish wind data. The initially smoothed version $\tilde{\tilde{H}}(\boldsymbol{h}, \tau)$ of the empirical covariance-spectral function $\tilde{H}(\boldsymbol{h}, \tau)$ has been constructed using the standard program `spec.pgram` in `R` with the span chosen subjectively. For estimation purposes as well as plotting figures the common choice span =255 has been used, though other choices have also been used below for exploratory purposes.

Estimation of $k(\tau)$: First we estimate $k(\tau)$. A stationary process with spectral density having a pole at zero frequency is called a stationary process with long-range dependence (Beran, 1994, Ch.2). Figures 3 and 4 show smoothed

empirical spectral densities for the 11 individual stations, $\tilde{\tilde{k}}_i(\tau)$, along with their overall spectral density $\tilde{\tilde{k}}(\tau)$. Since the averaging operation will tend to smooth out the noise in individual periodograms, we have chosen a smaller smoothing parameter for the averaged spectrum rather than for the individual periodograms. The plots of the marginal spectral densities show the spectral densities have roughly the same form at all stations with an apparent pole at zero frequency suggesting the existence of long-range dependence; Therefore Stein's parametric fractional exponential model (12) seems to be an appropriate model for the long-range dependence spectral density $k(\tau)$. The model selection criterium AIC suggests $K_1 = 3$ is appropriate choice in (12). For $K_1 = 3$, letting $\boldsymbol{Y} = \left(\log \tilde{\tilde{k}}(\tau)\right)$, $X = [\mathbf{1}, -\log\sin(|\pi\tau|), \cos(2\pi\tau), \cos(4\pi\tau), \cos(6\pi\tau)]$, $\boldsymbol{\beta} = (c_0, \beta, c_1, c_2, c_3)'$ and regressing $\log \tilde{\tilde{k}}(\tau)$ on the right hand side of (12) by OLS yields estimates $\hat{\beta} = 0.315 \pm 0.115, \hat{c}_0 = -1.769 \pm 0.092, \hat{c}_1 = 0.710 \pm 0.132, \hat{c}_2 = 0.022 \pm 0.086, \hat{c}_3 = 0.033 \pm 0.074$. Here the intervals are based on 2 standard errors fitted by the procedure developed in the previous section taking into account the possibly correlated errors.

The fitted parametric spectral density $\hat{k}_{\text{par}}(\tau)$ and the fitted nonparametric Nadaraya-Watson estimate $\hat{k}_{\text{np}}(\tau)$ are plotted in Figure 3. As can be seen in Figure 5, it seems that the normality and homoscedasticity assumptions on the residuals are sufficiently satisfied to justify the parametric OLS estimation.

Estimation of coherence: Next we consider empirical coherence plots for various pairs of sites. Figure 6(a) shows a plot of the smoothed absolute coherence versus time frequency for a subset of spatial lags including the biggest and smallest spatial lags. Our investigation indicates that the optimal smoothing parameters of coherence do not differ significantly from each other and so we use a common smoothing parameter for different spatial lags. This plot shows that the coherence decays exponentially with the decay parameter depending on the spatial lag. Therefore we choose the parametric power exponential function (6). The low coherence at low frequencies seems to be due to the long-range dependence in time. We omitted the first 300 frequencies in our estimation procedure.

Estimation of $\gamma(\tau)$: In the next step we estimate $\gamma(\tau)$ and p. First in (15) we assess the validity of common slope model by fitting a regression model with a separate slope for each τ and investigate how the fitted slopes behave. Figure 7 indicates that the fitted value of p are nearly constant over the time frequency, so a parallel-lines regression is appropriate. Thus we use common slope model (15).

The estimated common slope is $\hat{p} = 0.905 \pm 0.005$. For $K_2 = 3$, the estimated coefficients are given by $\hat{a}_0 = -6.551 \pm 0.019$, $\hat{a}_1 = -0.594 \pm 0.028$, $\hat{a}_2 = 0.010 \pm 0.027$, $\hat{a}_3 = -0.042 \pm 0.026$ which define $\hat{\gamma}_{\text{par}}(\tau)$. The parametric and Nadaraya-Watson estimates are presented in Figure 6(b). By model (14), estimates of p and $\gamma(\tau)$ define estimate of the absolute coherence for any spatial lag i.e. see Figure 6(a). The agreement with the empirical absolute coherence looks reasonable. Figures 8 depicts residuals for the empir-

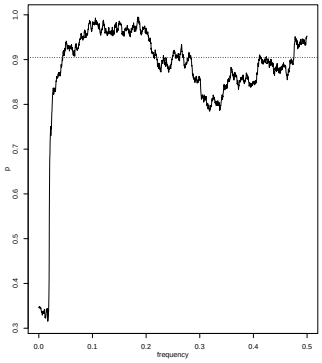

Fig. 7 Least squares estimate of p from separate slope model (15) for different time frequencies. The dotted line is the estimated common slope

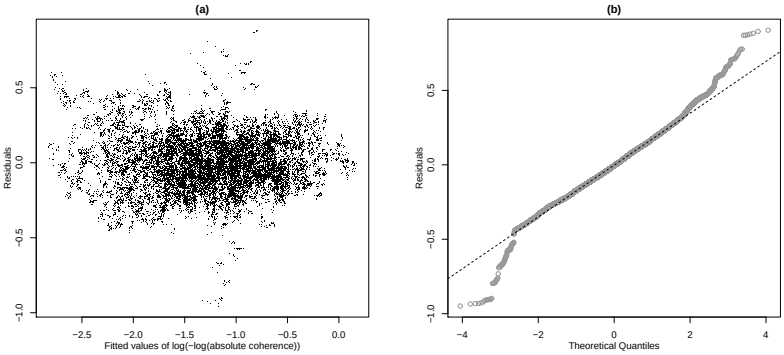

Fig. 8 (a): plot of residuals versus fitted $\log(-\log($ absolute coherence$))$. (b): normal Q-Q plot of residuals.

ical version of (15) and indicates that the homoscedasticity assumption of the $\log(-\log($ absolute coherence$))$ transform is plausible.

Estimation of $\theta(\tau)$:

Figure 9 shows a plot of the empirical phase $\tilde{\tilde{g}}_R(\boldsymbol{h}, \tau)$ versus time frequency τ for all different spatial lags. All the curves lie above the horizontal axis, but to improve visibility, half of them have been plotted below the axis. For each curve, the phase seems to increase linearly in τ for small τ, but is pulled back to 0 at $\tau = 0.5$ due to the periodic boundary conditions. Note that the maximum absolute value on the vertical axis is well below $\pi = 3.14$ so there is no distinction between the angular variable $\tilde{\tilde{g}}(\boldsymbol{h}, \tau)$ and its real-valued extension $\tilde{\tilde{g}}_R(\boldsymbol{h}, \tau)$ for this dataset. As illustrated in Figure 10(a), the slope of each curve depends on the spatial lag $\boldsymbol{h}$.

Using the estimation procedure in Section 3, the optimally estimated wind direction is given by $\hat{\boldsymbol{v}} = (0.999, 0.038)'$, i.e. winds in Ireland are predominantly westerly. Once $\boldsymbol{v}$ has been estimated, the initial estimate $\hat{\theta}_{\text{init}}(\tau)$ is

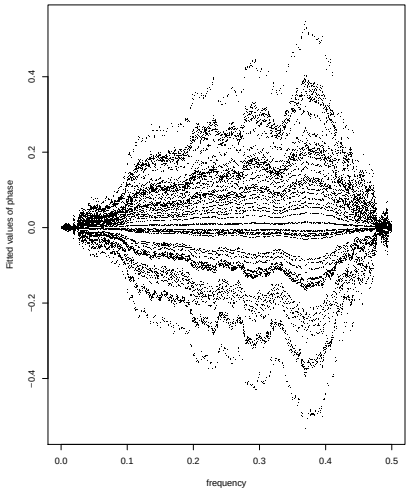

Fig. 9 plot of smoothed empirical phase for all different pairs of stations.

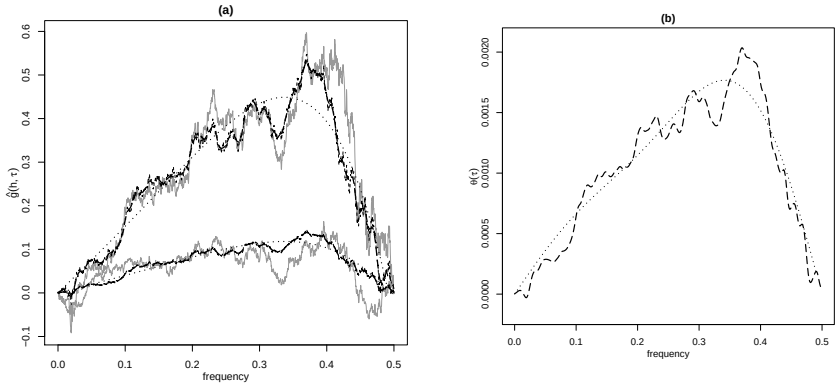

Fig. 10 (a): plot of empirical and estimated phase for two different pairs of stations; top: Valentia-Dublin with distance 316.99 km and bottom: Clones-Dublin with distance 105.47 km. (b): plot of estimated $\theta(\tau)$. Empirical estimate (gray curve), Nadaraya-Watson estimate (long-dash curve) and trigonometric regression estimation (dotted curve).

given by (18). Adopting a similar modelling strategy to that used for $\gamma(\tau)$, we model the function $\theta(\tau)$ with a trigonometric polynomial (20). After regressing $\hat{\theta}_{\text{init}}(\tau)$ on τ, the optimal order is found to be $K_3 = 2$ by the AIC criterion, with estimated coefficients $\hat{b}_1 = 0.00159 \pm 0.05021$, $\hat{b}_2 = -0.00045 \pm 0.04022$ which define $\hat{\theta}_{\text{par}}(\tau)$. The result is displayed in Figure 10(b) along with the nonparametric estimate. The parametric curve, the empirical curve, and the Nadaraya-Watson curve all fit the phase reasonably well. Figure 10(a) shows a plot of the estimated phase versus time frequency for two different spatial lags. Figure 11, shows that that the normality and homoscedasticity assumptions of residuals are broadly satisfied for parametric estimation.

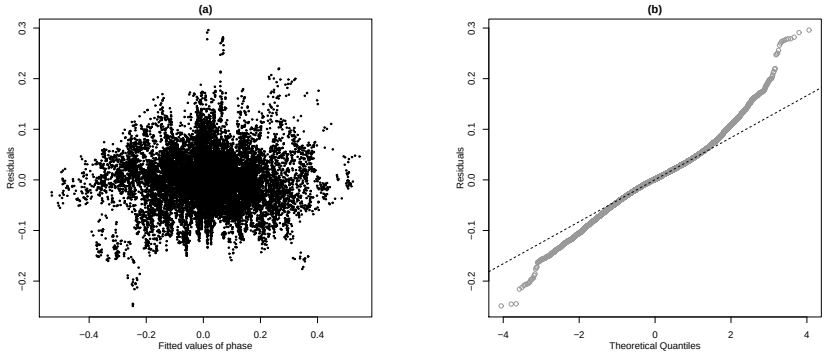

Fig. 11 (a): Residuals versus fitted values of $\theta(\tau)$. (b): Normal Q-Q plot of residuals.

6 Conclusions

In this paper we have given a deeper insight into the covariance-spectral modelling strategy of Stein (2005) and its properties. We proposed a simple transformation on the covariance-spectral function to make it linear in the unknown parameters, which facilitate the use of standard methods from regression analysis. Hence the method of estimation proposed here is more intuitive and easier to use than Stein (2005).

The effect of initial smoothing on the fitted function estimates and their standard errors has been explored. A elegant method to estimate the drift direction has been proposed; Stein (2005) assumed it is known. Phase winding is another important issue which has been explored clearly in this paper.

Stein constructed various plots to assess the goodness of fit of the model; we use similar plots to estimate the parameters. Our method can be seen as more graphical, enabling visual judgments to be made about the suitability of the proposed models for the functions $\gamma(\tau)$ and $\theta(\tau)$, and for the parameter p. In addition residual plots are constructed for goodness of fit assessment.

In general the behaviour of our estimates matches Steins estimates well for the Irish wind data. However, since Stein treats the data as lying on a sphere rather than on a plane, and since he uses a different amount of smoothing, numerical comparison with our estimates is not straightforward.

In our approach and that of Stein (2005), parametric and nonparametric models have been constructed on the log $\gamma(\tau)$ scale. In more recent work, Stein (2009) suggested an alternative parameterization in which such models are constructed directly for $1/\gamma(\tau) = \delta(\tau)$, say. This approach seemed to reduce the problem of singularities near $\tau = 0$ and gives an interesting avenue for future research.

Acknowledgements The authors would like to thank two referees and the associate editor, whose comments have been very helpful in improving the manuscript.

References

Beran, J. (1994). *Statistics for Long Memory Processes*. Chapman and Hall, New York.

Bloomfield, P. (1973). An exponential model for the spectrum of a scalar time series. *Biometrika*, 60:217–226.

Bloomfield, P. (1976). *Fourier Analysis of Time Series: An Introduction*. Wiley, New York.

Cox, D. R. and Isham, V. (1988). A simple spatial-temporal model of rainfall. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 415(1849):317–328.

Cressie, N. A. C. (1993). *Statistics for Spatial Data (Wiley Series in Probability and Statistics)*. Wiley-Interscience.

Cressie, N. A. C. and Huang, H. (1999). Classes of non-separable, spatio-temporal stationary covariance functions. *Journal of the American Statistical Association*, 94:1330–1340.

De Luna, X. and Genton, M. (2005). Predictive spatio-temporal models for spatially sparse environmental data. *Statistica Sinica*, 15:547–568.

Fuentes, M. (2006). Testing for separability of spatial-temporal covariance functions. *Journal of Statistical Planning and Inference*, 136(2):447–466.

Gneiting, T. (2002). Non-separable, stationary covariance functions for space-time data. *Journal of the American Statistical Association*, 97(458):590–600.

Haslett, J. and Raftery, A. E. (1989). Space-time modelling with long-memory dependence: Assessing Ireland's wind power resource. *Journal of Applied Statistics*, 38(1):1–50.

Lu, N. and Zimmerman, D. L. (2002). Testing for directional symmetry in spatial dependence using the periodogram. *Journal of Statistical Planning and Inference*, 129(1-2):369–385.

Ma, C. (2003). Families of spatio-temporal stationary covariance models. *Journal of Statistical Planning and Inference*, 116(2):489–501.

Mardia, K. V., Kent, J. T., and Bibby, J. M. (1979). *Multivariate Analysis*. Academic Press.

Mitchell, M. W., Genton, M. G., and Gumpertz, M. L. (2006). A likelihood ratio test for separability of covariances. *Journal of Multivariate Analysis*, 97(5):1025–1043.

Nadaraya, E. A. (1964). On estimating regression. *Theory of probability and its applications*, 10:186–190.

Priestley, M. B. (1981). *Spectral Analysis and Time Series*, volume I and II. Academic Press, London.

Ruppert, D., Sheather, S. J., and Wand, M. P. (1995). An effective bandwidth selector for local least squares regression. *Journal of the American Statistical Association*, 90:1257–1270.

Scaccia, L. and Martin, R. J. (2005). Testing axial symmetry and separability of lattice processes. *Journal of Statistical Planning and Inference*, 131(1):19–39.

Stein, M. L. (2005). Statistical methods for regular monitoring data. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67:667–687.

Stein, M. L. (2009). Spatial interpolation of high frequency monitoring data. *Annals of Applied Statistics*,3, 272–291.

Subba Rao, T., Das, S., and Boshnakov, G. N. (2014). A frequency domain approach for the estimation of parameters of spatio-temporal stationary random processes. *Journal of Time Series Analysis*, 35(4):357–377.

Watson, G. S. (1964). Smooth regression analysis. *Sankhya*, A26:359–372.