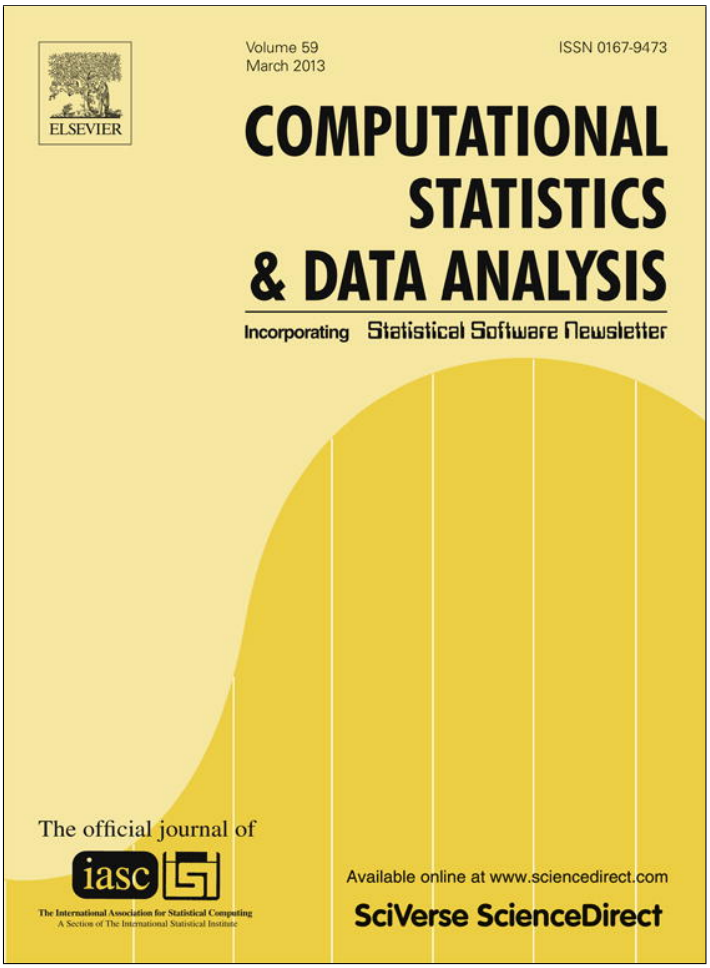

Computational Statistics and Data Analysis 59 (2013) 82–94

Contents lists available at SciVerse ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

Nonparametric inference in small data sets of spatially indexed curves with application to ionospheric trend determination

Oleksandr Gromenko [a,*], Piotr Kokoszka [b]

[a] Department of Mathematics and Statistics, Utah State University, Logan, UT 84322, United States
[b] Department of Statistics, Colorado State University, Fort Collins, CO 80523, United States

ARTICLE INFO

Article history:
Received 18 April 2012
Received in revised form 26 July 2012
Accepted 27 September 2012
Available online 4 October 2012

Keywords:
Functional data
Ionosphere
Long-term trend
Nonparametric inference
Spatial statistics

ABSTRACT

This paper is concerned with estimation and testing in data sets consisting of a small number (about 20–30) of curves observed at unevenly distributed spatial locations. Such data structures may be referred to as spatially indexed functional data. Motivated by an important space physics problem, we model such data as a mean function plus spatially dependent error functions. Given a small number of spatial locations, the parametric methods for the estimation of the spatial covariance structure of the error functions are not satisfactory. We propose a fully nonparametric estimator for the mean function. We also derive a test to determine the significance of the regression coefficients if the mean function is a linear combination of known covariate functions. In particular, we develop methodology for the estimation a trend in spatially indexed functional data, and for assessing its statistical significance. We apply the new tools to global ionosonde records to test the hypothesis of ionospheric cooling. Nonparametric modeling of the space–time covariances is surprisingly simple, much faster than those previously proposed, and less sensitive to computational errors. In simulated data, the new estimator and test uniformly dominate those based on parametric modeling.

1. Introduction

Models for data which exhibit both space and time dependence have attracted increasing attention in geophysical and environmental research. This is a fast growing branch of statistics, for a general overview see Cressie and Wikle (2011) and Sherman (2011), for a fast, accessible introduction, we recommend (Gneiting et al., 2007). Space–time data could be roughly separated into several categories according to the amount of information contained, respectively, in their spatial and temporal components. One category is the data which have a very rich spatial component and relatively limited temporal component. Such data usually come from satellites, see e.g. Jun and Stein (2009), Cressie et al. (2010) and Katzfuss and Cressie (2011), among many others. Another category is data which have a rich temporal component and a relatively limited spatial component. Such data come typically as collections of long time series recorded at different spatial locations by ground based instruments. For example, the Irish wind data studied by Haslett and Raftery (1989) and consequently used in many other papers, the Canadian weather data extensively used in Ramsay and Silverman (2005) and Ramsay et al. (2009), pollution data studied by Bowman et al. (2009), and many others.

In this paper, we propose a flexible, fully nonparametric methodology for data of the latter type. It includes estimation of the mean function and is applied to testing the statistical significance of a linear trend. Our methodology builds on the theory of Hall et al. (1994) and Hall and Patil (1994) by (1) developing a practically applicable tool set for the estimation and

* Corresponding author. Tel.: +1 3152625416.
E-mail addresses: agromenko@gmail.com (O. Gromenko), Piotr.Kokoszka@colostate.edu (P. Kokoszka).

0167-9473/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.csda.2012.09.016

testing in the spatial context with few data locations, (2) extending it to the framework of spatially indexed functional data, (3) developing suitable confidence bounds, and (4) applying it to an important space physics problem. The work presented in this paper is a direct result of our attempts to solve this important space physics problem in a fairly conclusive manner that would be satisfactory to the space physics community. Since the problem concerns the detection of a long term (many decades) trend, we hope that out methodology is general and useful enough to be applicable to other similar data sets and problems. Spatially indexed functional data have been the focus of several recent studies, see Delicado et al. (2010); Giraldo et al. (2011), Nerini et al. (2010), Gromenko et al. (2012) and Gromenko and Kokoszka (forthcoming). Existing approaches however often fail when the number of spatial locations is small because in such cases the numerical optimization required to fit a parametric spatial model may fail, or the fit may be poor. The research we report is, to a large extent, a result of computational difficulties we encountered with standard approaches. The resulting new methodology is computationally faster and the algorithms never fail to converge (in our data sets and simulations).

The paper is organized as follows. In Section 2, we develop a nonparametric covariance estimation procedure for scalar data. Next, in Section 3, a statistical model for spatially indexed functional data is introduced. Section 4 presents the estimation procedure for this model. In Section 5, we derive a test for assessing the significance of regression coefficients when the mean function is a linear combination of known covariate functions. The application of this test to the assessment of a long term cooling trend in the ionosphere is presented in Section 6. Section 7 presents the results of simulation studies that validate the methodology we propose and its application to the ionosonde data.

2. Description of the method for scalar data

In this section, we assume that ζ is a mean zero stationary and isotropic *scalar* random field observed at locations $\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_N$, and $\boldsymbol{\Gamma}$ is the $N \times N$ matrix of covariances

$$\gamma(d_{k\ell}) = \mathrm{Cov}(\zeta(\mathbf{s}_k), \zeta(\mathbf{s}_\ell)) = E[\zeta(\mathbf{s}_k)\zeta(\mathbf{s}_\ell)],$$

where $d_{k\ell}$ is the distance between $\mathbf{s}_k$ and $\mathbf{s}_\ell$. Estimation of $\boldsymbol{\Gamma}$ is not trivial for small samples. A standard variogram based estimator for small spatial data sets is generally unstable, and the the optimization often fails to converge. It is recommended that every lag interval should contain at least 30 distinct distances, but for small sample sizes, it is difficult to meet this condition without reducing the number of intervals to a level which makes fitting a parametric model difficult. We therefore develop nonparametric methodology, based on the work of Hall et al. (1994) and Hall and Patil (1994), which is suitable for small data sets. It forms the basis of the estimation and testing procedures for functional spatially indexed data, but can also be used for different spatio-temporal models, as illustrated in Example 7.1.

Recall that $d_{k\ell}$ is the distance between $\mathbf{s}_k$ and $\mathbf{s}_\ell$, and consider the preliminary estimator

$$\tilde{\gamma}(d_{k\ell}) = \zeta(\mathbf{s}_k)\zeta(\mathbf{s}_\ell). \tag{1}$$

It is possible that for some distances there exist several distinct estimators $\tilde{\gamma}(d_{k\ell})$, in fact for $d_{k\ell} = 0$ there are always N different preliminary estimators. The estimated covariances are ordered so that the corresponding distances do not decrease: denoting the $d_{k\ell}$ by d_i, we thus have $d_i \le d_{i+1}$, $1 \le i \le N(N+1)/2$. The resulting sequence $\{\tilde{\gamma}(d_i) : 1 \le i \le N(N+1)/2\}$ is very noisy and must be smoothed. We use local linear regression, see Fan and Gijbels (1996), rather than the kernel smoother suggested by Hall et al. (1994). The reason for using the local linear regression is that it introduces a slightly smaller bias for small and large distances d_i. Let $\kappa(x)$ be a compactly supported symmetric probability density function. The smoothed value of $\gamma(d)$ is thus estimated by $\hat{m}(d)$ computed by minimizing

$$(\hat{m}(d), \hat{m}_1) = \arg\min_{m,m_1} \sum_{i=1}^{N(N+1)/2} \kappa\left(\frac{d-d_i}{h}\right)\{\tilde{\gamma}(d_i) - m(d) - m_1(d-d_i)\}^2. \tag{2}$$

We performed simulations using several popular kernels (triangular, quadratic, Epanechnikov, triweight, tricube), and found that they produce practically the same estimates. The results reported in this paper are based on the Epanechnikov kernel. As with all problems of this type, the most difficult issue is the selection of the bandwidth h; Hall et al. (1994) do not recommend any specific procedure. They developed an interactive software which allows the user to choose the bandwidth and visually compare the resulting estimates. We describe our method of bandwidth selection in the Appendix.

To construct a positive definite covariance function, we use Bochner's theorem: We compute the Fourier transform of $\hat{m}$ and delete all negative lobes. The inverse Fourier transform is then our final estimator $\hat{\gamma}(d)$. We enhance the idea of Hall et al. (1994) by providing a procedure to construct functional confidence intervals for $\hat{\gamma}(\cdot)$, see the Appendix. The application of the procedure to simulated data is illustrated in Fig. 1.

Hall et al. (1994) showed that to achieve consistency in the estimation of $\gamma(d)$, the distance between $\min(d_i)$ and $\max(d_i)$ (the range) must grow much slower than the number of the d_i. This condition is naturally satisfied in the spatial setting because adding one more $\mathbf{s}_k$ roughly increases the range at most by a unit, but increases the number of the d_i by N.

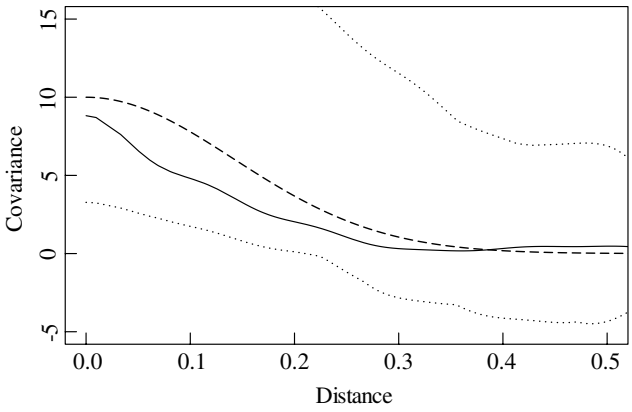

Fig. 1. Illustration of the estimation procedure for scalar data. The true covariance function (dashed line), its estimate (solid line) and the 95% confidence region (dotted lines).

3. Statistical model for spatially indexed functional data

The methodology developed in this paper is motivated by the problem of the estimation and modeling of the mean function $\mu(\cdot)$ in the model

$$X(\mathbf{s}; t) = \mu(t) + \varepsilon(\mathbf{s}; t). \tag{3}$$

The data are curves $X(\mathbf{s}_k; t)$, $t \in [0, T]$, observed at spatial locations $\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_N$. Such functional data structures are quite common: examples are discussed in Delicado et al. (2010), Nerini et al. (2010), Giraldo et al. (2011), Hörmann and Kokoszka (forthcoming), Gromenko et al. (2012), Gromenko and Kokoszka (forthcoming) and Chapters 17 and 18 of Horváth and Kokoszka (2012). In model (3), the error curves $\varepsilon(\mathbf{s})$ are assumed to form a mean zero strictly stationary and isotropic spatial random field taking values in the Hilbert space of square integrable functions with the usual inner product, see e.g. Chapter 2 of Horváth and Kokoszka (2012). We assume that

$$E\|\varepsilon(\mathbf{s})\|^2 = E\int \varepsilon^2(\mathbf{s}; t)dt < \infty.$$

If the above assumptions are satisfied, the error term can be represented using the Karhunen–Loève expansion $\varepsilon(\mathbf{s}; t) = \sum_{j=1}^{\infty} \zeta_j(\mathbf{s})v_j(t)$, where v_j is the jth functional principal component (FPC), and $\zeta_j(\mathbf{s}) = \langle X(\mathbf{s}) - \mu, v_j\rangle$ is the score of $\varepsilon(\mathbf{s})$ with respect to it. Recall that the v_i are the eigenfunctions of the covariance operator $E[\langle X(\mathbf{s}) - \mu, \cdot\rangle (X(\mathbf{s}) - \mu)]$. This leads to the model

$$X(\mathbf{s}; t) = \mu(t) + \sum_{j=1}^{\infty} \zeta_j(\mathbf{s})v_j(t). \tag{4}$$

In the above model, the mean function $\mu(\cdot)$ does not depend on the spatial location. It represents the mean temporal evolution of the spatially distributed curves. The inference for $\mu(\cdot)$, when then number of the spatial locations $\mathbf{s}_k$ is small, is the focus of this paper.

The fields $\zeta_j(\cdot)$ are mean zero purely spatial random fields. Set

$$\boldsymbol{\zeta}_j = [\zeta_j(\mathbf{s}_1), \ldots, \zeta_j(\mathbf{s}_N)]^T, \qquad \boldsymbol{\Gamma}_j = \mathrm{Var}[\boldsymbol{\zeta}_j].$$

The matrix $\boldsymbol{\Gamma}_j$ is a positive definite $N \times N$ matrix with elements

$$\gamma_j(\mathbf{s}_k - \mathbf{s}_\ell) = E\left[\zeta_j(\mathbf{s}_k)\zeta_j(\mathbf{s}_\ell)\right].$$

Our modeling framework requires that for every k, l

$$E[\zeta_j(\mathbf{s}_k)\zeta_{j'}(\mathbf{s}_\ell)] = 0 \quad \text{if } j \neq j'. \tag{5}$$

Note that (5) is always true for $k = \ell$, but for $k \neq \ell$ it does not follow from any mathematical argument. One can show that the separability of the spatio-temporal covariance function implies (5), and we need assumption (5) to ensure that the spatio-temporal covariance function is positive definite. Observe that under (5) the covariances are given by

$$\mathrm{Cov}(X(\mathbf{s}_k; t), X(\mathbf{s}_\ell; t')) = c(\mathbf{s}_k, \mathbf{s}_\ell; t, t') = \sum_{j=1}^{\infty} \gamma_j(\mathbf{s}_k - \mathbf{s}_\ell)v_j(t)v_j(t'). \tag{6}$$

It is not difficult to verify that if each $\boldsymbol{\Gamma}_j$ is positive definite, then

$$\sum_{k,\ell=1}^{N} \sum_{t,t'} a_k(t) a_\ell(t') c(\mathbf{s}_k, \mathbf{s}_\ell; t, t') \geq 0.$$

Model (6) is obviously nonseparable and enjoys the property of full symmetry. Its *spatial* component is (strictly) stationary and isotropic and temporal component is nonstationary. Model (6) is thus more general than those proposed in Cressie and Huang (1999) and Gneiting (2002).

4. Estimation of the mean function

In this section, we put together the developments of Sections 2 and 3, and propose a complete nonparametric methodology for the estimation of the mean in model (4). Gromenko et al. (2012) considered several approaches to the estimation of the function μ and found that the smallest integrated mean square and absolute errors are obtained by using a weighted sum of functional observations:

$$\hat{\mu}(t) = \sum_{k=1}^{N} w_k X(\mathbf{s}_k; t), \quad \sum_{k=1}^{N} w_k = 1. \tag{7}$$

The weights w_k are found by minimizing the expected value of the L^2 norm of the difference $\hat{\mu}(t) - \mu(t)$, and are given by

$$\mathbf{w} = \mathbf{C}^{-1}\mathbf{1}/(\mathbf{1}^T\mathbf{C}^{-1}\mathbf{1}), \tag{8}$$

where

$$\mathbf{C} = E\left[\langle \boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}^T \rangle\right] = \sum_{j=1}^{\infty} \boldsymbol{\Gamma}_j. \tag{9}$$

The estimation of μ is summarized in the following algorithm.

(a) Compute the sample mean $N^{-1}\sum_{k=1}^{N} X(\mathbf{s}_k; t)$. This yields the first estimate $\hat{\mu}(t)$, which will be improved in subsequent iterations.
(b) Calculate the functional version of (1), i.e.

$$\tilde{\gamma}(d_{k\ell}) = \langle X(\mathbf{s}_k) - \hat{\mu}, X(\mathbf{s}_\ell) - \hat{\mu} \rangle = \int \left(X(\mathbf{s}_k; t) - \hat{\mu}(t)\right)\left(X(\mathbf{s}_\ell; t) - \hat{\mu}(t)\right) dt$$

and estimate the covariance matrix $\mathbf{C}$ as described in Section 2.
(c) Compute the weights using (8), and the mean function, $\hat{\mu}(t)$, using (7).
(d) Repeat steps (b)–(c) until suitable convergence is reached.

The convergence of the algorithm is evaluated by means of the quantity

$$R_i = \int_0^1 |\mu^{(i+1)}(t) - \mu^{(i)}(t)| dt,$$

where the index i denotes the number of iteration. When the R_i do not decrease with i, we stop the algorithm. The graphs of the R_i for real data are shown in Fig. 6.

A similar estimation procedure for model (4) was considered in Gromenko et al. (2012) (Method M2) and Gromenko and Kokoszka (forthcoming), but the nonparametric covariance estimation was not used. We note that we do not estimate the covariances $\boldsymbol{\Gamma}_j$ of the scores processes separately. This is a very computationally expensive process, and our approach avoids it.

A simulation study that validates the above method and shows its superiority (in small samples) relative to current approaches is presented in Section 7.

5. Significance of regression coefficients

In this section, we assume that the mean function $\mu(\cdot)$ is a linear combination of q known functions, so that model (4) takes the form

$$X(\mathbf{s}; t) = \sum_{i=1}^{q} \beta_i z_i(t) + \sum_{j=1}^{\infty} \zeta_j(\mathbf{s}) v_j(t). \tag{10}$$

Model (10) is motivated by the application to ionosonde data described in Section 6, where there is a dominant explanatory function $z(t)$ which quantifies the solar activity. Model (10) can include a linear trend $z(t) = t$, and can be used to test

the significance of the coefficient of this trend. Problems related to testing the significance of long term trends abound in geophysical and ecological sciences. We now describe how this can be done if the relevant time series are measured at several spatial locations.

Introduce the following vectors

$$\boldsymbol{\beta} = [\beta_1, \ldots, \beta_q]^T, \qquad \mathbf{z}(t) = [z_1(t), \ldots, z_q(t)]^T,$$

and matrices

$$\mathbf{Q} = \left[\langle z_i, z_{i'}\rangle,\ 1 \le i, i' \le q\right], \qquad \boldsymbol{\Omega} = \left[\langle z_i, v_j\rangle,\ 1 \le i \le q, 1 \le j \le p\right].$$

The number p of the FPC's is typically selected using the cumulative variance criterion, see e.g. Ramsay and Silverman (2005) or Horváth and Kokoszka (2012). A general recommendation is to use p such that the first p components explain about 85%–90% of the variance. It is often useful to perform the inference for several values of p. If the conclusions do not depend on p, we can place more confidence in them.

To estimate the parameter vector $\boldsymbol{\beta}$, we minimize

$$\left\| \sum_{n=1}^{N} w_n X(\mathbf{s}_n) - \sum_{i=1}^{q} \beta_i z_i \right\|^2, \tag{11}$$

which lead to the solution

$$\hat{\boldsymbol{\beta}} = \mathbf{Q}^{-1} \left\langle \mathbf{z}, \mathbf{w}^T \mathbf{X} \right\rangle. \tag{12}$$

The quantity $\left\langle \mathbf{z}, \mathbf{w}^T \mathbf{X} \right\rangle$ is a $q \times 1$ vector with the ith entry $\left\langle z_i, \sum_{k=1}^{N} w_k X(\mathbf{s}_k) \right\rangle$. Since $\mathbf{w}^T \mathbf{X}$ is an estimate of the mean function $\mu(\cdot)$, (12) can be written as

$$\hat{\boldsymbol{\beta}} = \mathbf{Q}^{-1} \left\langle \mathbf{z}, \hat{\mu} \right\rangle. \tag{13}$$

It is clear now that the estimation of the regression coefficients is a two step procedure: first we estimate $\mu(\cdot)$ using the methodology of Section 4, then we use Eq. (13).

The variance of (12), $\mathrm{Var}[\hat{\boldsymbol{\beta}}]$, can be calculated in two different ways. The first way is by substituting the bootstrap sample of $\hat{\mu}$ into (13) and using the sample variance. This approach is computationally very expensive and we do not recommend it. Long run times are due to the spatial estimation and the estimation of the FPC's. Instead, we propose an approach based on the following calculations. Observe that

$$\begin{aligned} \left\langle \mathbf{z}, \mathbf{w}^T \boldsymbol{\varepsilon} \right\rangle &= \left[\sum_{k=1}^{N} \sum_{j=1}^{p} w_k \zeta_j(\mathbf{s}_k) \left\langle v_j, z_1 \right\rangle, \ldots, \sum_{k=1}^{N} \sum_{j=1}^{p} w_k \zeta_j(\mathbf{s}_k) \left\langle v_j, z_q \right\rangle \right]^T \\ &= \left[\sum_{j=1}^{p} \tilde{\zeta}_j \left\langle v_j, z_1 \right\rangle, \ldots, \sum_{j=1}^{p} \tilde{\zeta}_j \left\langle v_j, z_q \right\rangle \right]^T, \end{aligned} \tag{14}$$

where $\tilde{\zeta}_j$ is the weighted sum of the scores: $\tilde{\zeta}_j = \sum_{k=1}^{N} w_k \zeta_j(\mathbf{s}_k)$. Let $\tilde{\boldsymbol{\zeta}} = [\tilde{\zeta}_1, \ldots, \tilde{\zeta}_p]^T$ then $\left\langle \mathbf{z}, \mathbf{w}^T \boldsymbol{\varepsilon} \right\rangle = \boldsymbol{\Omega} \tilde{\boldsymbol{\zeta}}$ and hence

$$\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} = \mathbf{Q}^{-1} \boldsymbol{\Omega} \tilde{\boldsymbol{\zeta}}.$$

Thus

$$\mathrm{Var}[\hat{\boldsymbol{\beta}}] = \mathbf{Q}^{-1} \boldsymbol{\Omega} \mathrm{Var}[\tilde{\boldsymbol{\zeta}}] \boldsymbol{\Omega}^T \mathbf{Q}^{-1}. \tag{15}$$

Assuming that the weights are known constants, we have

$$\mathrm{Var}[\tilde{\boldsymbol{\zeta}}] = \mathrm{diag}\left(\mathbf{w}^T \boldsymbol{\Gamma}_1 \mathbf{w}, \ldots, \mathbf{w}^T \boldsymbol{\Gamma}_p \mathbf{w}\right), \tag{16}$$

which is a $p \times p$ diagonal matrix. The diagonal form of $\mathrm{Var}[\tilde{\boldsymbol{\zeta}}]$ is a consequence of assumption (5). All quantities appearing in (15) and (16) can be estimated using the methodology presented in the previous sections.

If the functions $X(\mathbf{s}_k)$ are normally distributed, then, by (12), the estimator $\hat{\boldsymbol{\beta}}$ is approximately normal. Thus to test $\beta_i = 0$, for a fixed i, we assume that the statistic $\hat{\beta}_i / \sqrt{\mathrm{Var}[\hat{\beta}_i]}$ has the standard normal distribution. We note that the spatial quasi-bootstrap method also requires assumptions on the distribution of the spatial processes ξ_j, and the only practical assumption is that these processes are normal (we must use the Cholesky decomposition). In Section 6 we verify that the assumption of normality is approximately satisfied by the ionosonde data.

A simulation study that validates our method is presented in Section 7.

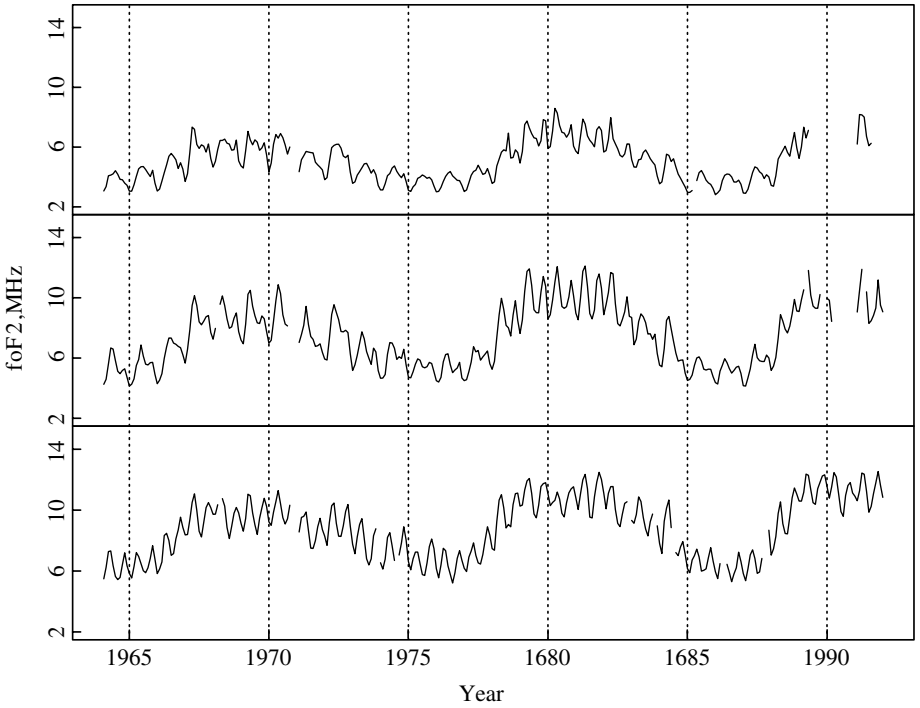

Fig. 2. F2-layer critical frequency curves at three locations. Top to bottom (latitude in parentheses): Yakutsk (62.0), Yamagawa (31.2), Manila (14.7). The functions exhibit a latitudinal trend in amplitude.

6. Application to ionosonde data

We first provide some background and motivation. The problem we study is related to the the hypothesis of Roble and Dickinson (1989) who argued that the increasing amounts of greenhouse gases in the upper atmosphere should lead to its global cooling because these gases radiate heat into space. Rishbeth (1990) pointed out that such a cooling would result in a thermal contraction and the global lowering of the highest or peak electron density. The ionospheric layer which contain the peak electron density is known as the F2 region. The peak electron density above any location on the Earth can be measured indirectly. Data obtained from a type of radar called the ionosonde allow to compute the a critical frequency, denoted foF2. This frequency is related to the peak electron density by an equation based partially on laws of physics and partially on empirical corrections. There are, in fact, several versions of this equation, but the main point is that a lowering in the peak electron density corresponds to a decrease in the foF2 frequency. There has consequently been extensive space physics research aimed at determining if a decreasing temporal trend in the foF2 frequency indeed exists. An interested reader is referred to Lastovicka et al. (2008), as a starting point. We note that a long term change in the upper atmosphere can impact space-based navigation, short-wave (3–30 MHz) radio communication and the operation of low orbit satellites. But perhaps even more importantly, long-term changes in the upper atmosphere and in the lower atmosphere (troposphere) can be governed by the same factors such as solar activity, and change the Earth's magnetic field and greenhouse gases concentration. Consequently, understanding of the influence of these factors and their combinations in the upper atmosphere can provide additional information on long-term changes in the troposphere.

Long-term changes in the upper atmosphere are usually described using a linear approximation which is called the ionospheric trend. The main problem in its determination is the separation of the solar activity and other factors, like the long term changes in the internal magnetic field of the Earth, see Clilverd et al. (2003). The solar cycle however clearly dominates the shape of the foF2 curves, as shown in Fig. 2. A comprehensive overview of statistical methods proposed in the space physics community is given in Lastovicka et al. (2006). The main problem from which they suffer is their inability to combine the information from many spatial locations. The model and the testing approach we propose is an attempt to overcome this difficulty. We consider the following specialization of model (10):

$$\text{foF2}(\mathbf{s};t) = \beta_1 + \beta_2 t + \beta_3 \text{SRF}(t) + \sum_{j=1}^{p} \zeta_j(\mathbf{s}) v_j(t). \tag{17}$$

The covariate SRF is the solar radio flux measured in W/m^2 Hz. It is a proxy for the solar activity. Another possible proxy is the sunspot number. Our primary interest in this section is in testing the hypothesis $H_0 : \beta_2 = 0$.

In this paper, we use data from 28 ionosonde stations, see Fig. 3, located in the mid-latitude region from 30° to 60° in the magnetic coordinate system. To study the solar influence on trends we split data into Night data from 22 to 2 LT, Noon

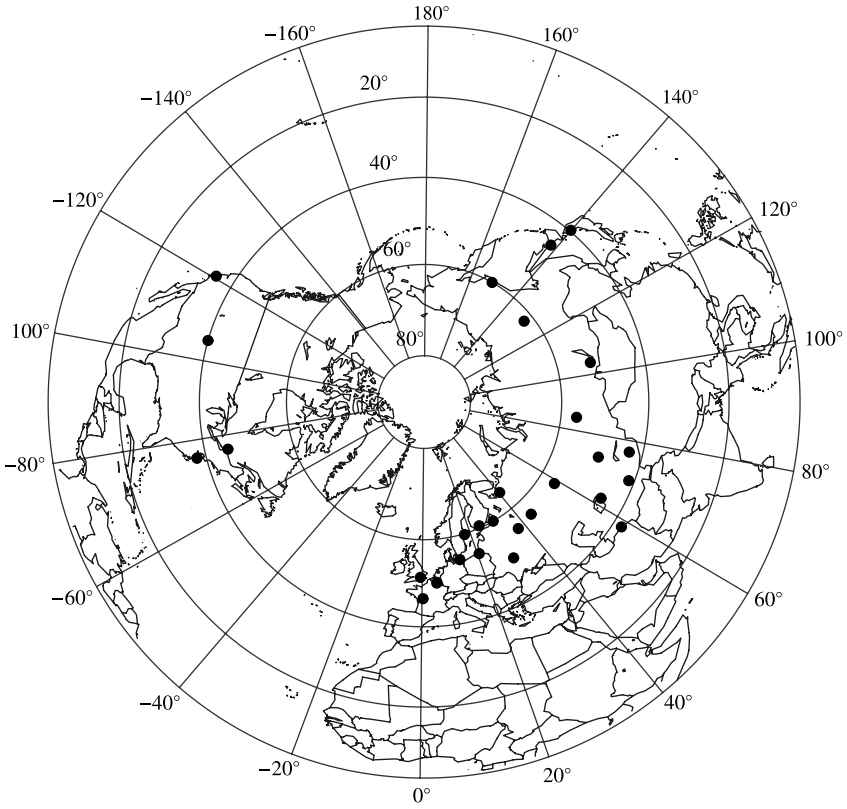

Fig. 3. Locations of 28 ionosonde stations in the northern hemisphere.

Table 1
P-values for the trend parameter as a function of the number of the FPC's.

Target CV, %	Day			Noon			Night		
	Final p	Final CV, %	P-value	Final p	Final CV, %	P-value	Final p	Final CV, %	P-value
80	2	80.81	$3.3 \cdot 10^{-03}$	2	84.08	$5.58 \cdot 10^{-3}$	2	81.82	0.3025
85	–	–	–	–	–	–	3	89.08	0.3197
90	3	90.81	0.0101	3	92.32	$5.79 \cdot 10^{-3}$	4	92.40	0.3302
–	4	93.10	0.0102	4	94.70	$5.79 \cdot 10^{-3}$	–	–	–
–	5	94.62	0.0103	–	–	–	–	–	–
95	6	95.87	0.0103	5	96.26	$6.04 \cdot 10^{-3}$	5	95.24	0.3304

Data from 10 to 14 LT and Day data (no time filter is applied). Here and below LT means local (latitudinal) time. The data are available from 1967-08-01 to 1989-08-01. This interval covers 22 years or 2 solar eleven year cycles. We do not discuss the details of creating the functional objects from the raw data, as these are fairly complex. An interested reader is referred to Gromenko et al. (2012) and Gromenko and Kokoszka (forthcoming). The distance between the locations on the globe is measured using the chordal distance.

Table 1 shows the P-values obtained for several values of p. These values were selected by targeting a specific percentage of variance explained by the first p FPC's. It is seen that the trend coefficient is significant for the Day and Noon data, but not for the Night data. A more comprehensive discussion of this finding will require a more detailed space physics research, but we note that our result is consistent with discussions published in space physics literature. Due to different physical processes, the behavior of the upper atmosphere is different at different times of a day. Our finding, in a sense, confirms that the problem of trend determination is complex, and a clear cut answer may not be available. The problem must be formulated in a more precise way. One might clearly wonder if the acceptance of $H_0 : \beta_2 = 0$ is not a type I error. As seen in Fig. 5, the test has the power of about 80%, so a type one error is a possibility. One way to increase power, would be to consider more ionosonde locations. There are however two problems that must first be solved. (1) As seen in Fig. 3, the amplitude of the curves depends on the latitude. Thus extending the latitudinal coverage would violate the assumption of stationarity that underlines our methodology. (2) Most stations outside the northern hemisphere have incomplete records. These are not a few missing observations, but missing stretches of data of length 5–20 years. A suitable methodology to accommodate such incomplete records would need to be developed.

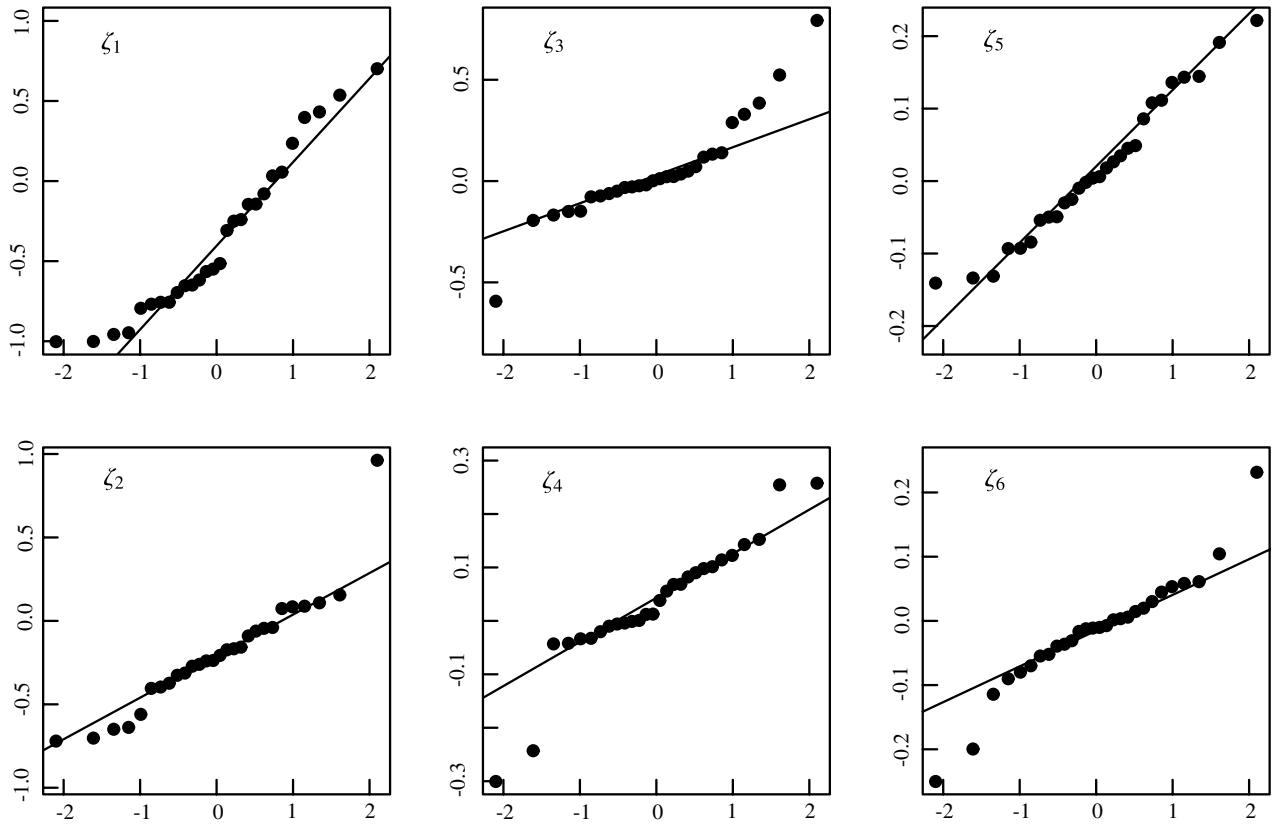

Fig. 4. Normal *QQ*-plots plots for the estimated scores, ζ_i, $1 \leq i \leq 6$ for the Day data.

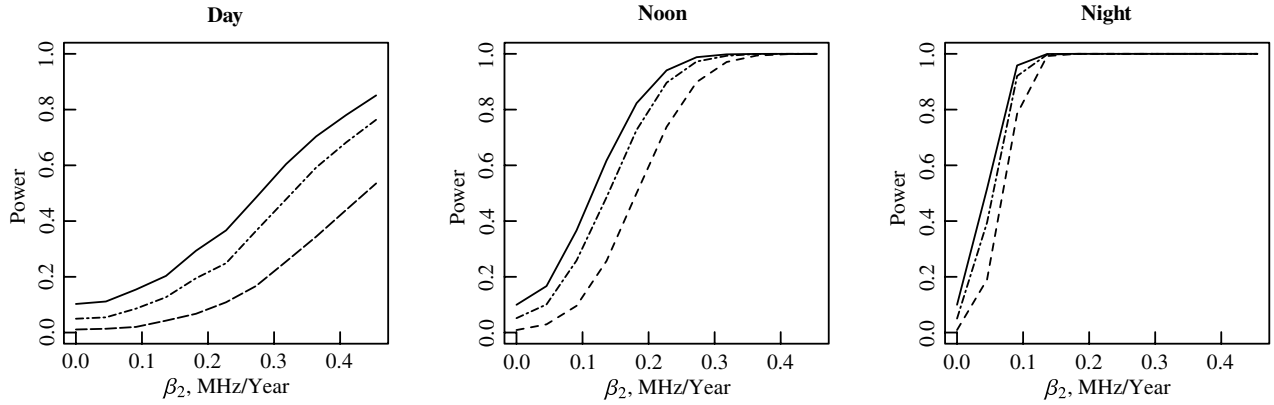

Fig. 5. Empirical power. Solid line—empirical power for $\alpha = 10\%$, dash-dotted line—empirical power for $\alpha = 5\%$, dashed line—empirical power for $\alpha = 1\%$.

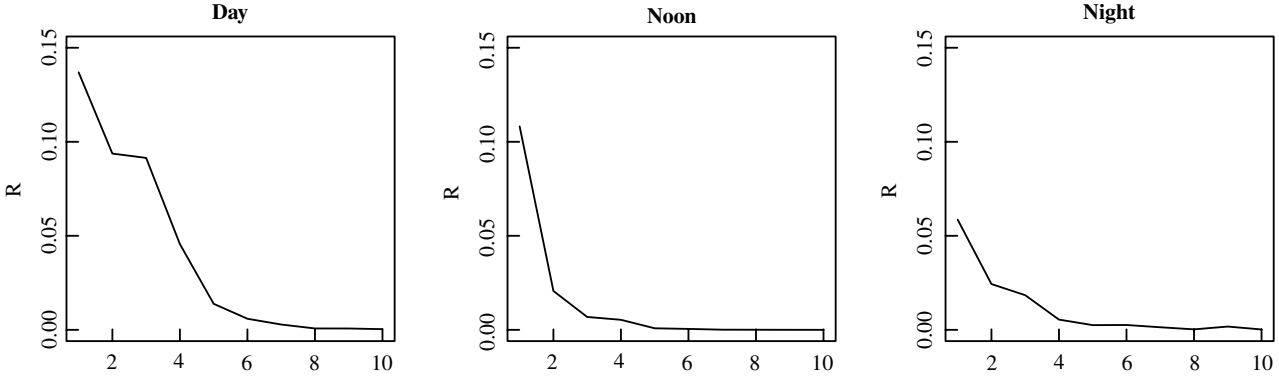

Fig. 6. Convergence of the iterative method for the estimation of the means R_i as a function of iteration i.

Table 2
Average L_2 distance between the estimated and true mean functions for different estimation methods; the second number represents the standard error. The entries are based on 10^4 replications.

Sample	S	T	P	N
15	0.164 ± 0.006	0.151 ± 0.005	0.145 ± 0.006	0.069 ± 0.004
20	0.137 ± 0.005	0.123 ± 0.004	0.116 ± 0.005	0.061 ± 0.003
25	0.122 ± 0.005	0.108 ± 0.004	0.098 ± 0.004	0.059 ± 0.004
30	0.115 ± 0.004	0.105 ± 0.004	0.098 ± 0.004	0.060 ± 0.004
35	0.109 ± 0.004	0.096 ± 0.004	0.089 ± 0.004	0.060 ± 0.004
40	0.104 ± 0.005	0.089 ± 0.004	0.082 ± 0.004	0.059 ± 0.005

7. Validation of the methodology

In this section, we present the results of several simulation studies that validate the estimation on testing methods introduced in Sections 4 and 5.

Methodology of Section 4. To demonstrate the superior performance (in small samples) of the new nonparametric estimation procedure, we designed the following simulation study. We generate data using model (4) with two FPC's, i.e.

$$X(\mathbf{s}; t) = \mu(t) + \zeta_1(\mathbf{s})v_1(t) + \zeta_2(\mathbf{s})v_2(t). \tag{18}$$

We set

$$v_1(t) = \sin(2\pi t \cdot 7) + \sin(2\pi t \cdot 2), \qquad v_2(t) = \sqrt{2}\sin(2\pi t \cdot 3);$$
$$\boldsymbol{\zeta}_1 \sim N(\mathbf{0}, \boldsymbol{\Gamma}_1), \qquad \boldsymbol{\zeta}_2 \sim N(\mathbf{0}, \boldsymbol{\Gamma}_2),$$

where the elements of the covariance matrices have the following parametric form:

$$\gamma_1(d_{k\ell}) = \exp(-d_{k\ell}/0.1), \qquad \gamma_2(d_{k\ell}) = 0.2\exp(-(d_{k\ell}/0.3)^2).$$

The spatial locations $\mathbf{s}_k$ are uniformly distributed on the unit square (different locations for every MC replication).

We compare four estimation procedures:

S Simple average, which totally ignores the spatio-temporal dependence;
T The infeasible method which uses the true covariance matrix $\mathbf{C} = \boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}_2$;
P The method that uses a *misspecified* parametric spatio-temporal covariance function $\sigma^2 C(\mathbf{h}; u)$, with $C(\mathbf{h}; u)$ given by

$$C(\mathbf{h}; u) = \frac{1}{1 + a|u|^{2\alpha}} \exp\left(-\frac{c\|\mathbf{h}\|}{(1 + a|u|^{2\alpha})^{\beta/2}}\right). \tag{19}$$

(This covariance function is discussed in detail in Example 7.1 below in this section.)
N The new nonparametric method.

Method *S* corresponds to using only step (a) in the algorithm presented earlier in this section. It would be the default method if standard *R* or Matlab software were to be used. The remaining methods are iterative, and differ in the way in which the covariance matrix in step (b) is estimated. The parametric model used in method *P* is discussed in greater detail in Example 7.1.

Table 2 compares the performance of the four methods for sample sizes ranging from 15 to 40. We report Monte Carlo averages and standard deviations of the L_2 distance defined as

$$L_2 = \left\{\int \left(\hat{\mu}_i(t) - \mu(t)\right)^2 dt\right\}^{1/2}. \tag{20}$$

The most important conclusion is that method *N* is significantly better than all other methods at the 5% confidence level (a difference of more than two standard deviations, which will be the benchmark in the discussion that follows). It is even better than the parametric method *T* which assumes the known true covariances. This illustrates a relatively well-known fact that a flexible nonparametric model can approximate the stochastic structure of the data better than a true parametric model, when the number of data points is small. Method *P*, based on a more flexible parametric family is significantly worse than method *N*, but "almost" significantly better than Method *T*: the differences between *N* and *P* are significant at about a 10% level. The standard method *S* is significantly worse than any other methods. This means that taking into account spatial dependence of the curves even in a suboptimal way significantly improves the estimates. We emphasize that for method *P* only the cases in which the variogram optimization converged were considered; it did not converge in over 10% of replications. The conclusions reached from the analysis of Table 2 do not change if the data generating process (18) is modified by adding more principal components which however do not account for more than 10% of variability, as is the case for the ionosonde data.

Table 3
Average L_2 distance between the estimated and true means; the second number represents the standard error. Entries are based on 10^4 replications.

Parameters	Sample	S	T	P	N
$\beta = 0, a = 0.1,$	15	0.201 ± 0.002	0.185 ± 0.003	0.192 ± 0.004	0.181 ± 0.001
	20	0.188 ± 0.002	0.167 ± 0.002	0.183 ± 0.007	0.166 ± 0.002
	25	0.178 ± 0.002	0.158 ± 0.003	0.158 ± 0.004	0.155 ± 0.002
	30	0.173 ± 0.003	0.154 ± 0.003	0.158 ± 0.006	0.150 ± 0.003
$\beta = 0, a = 1,$	15	0.199 ± 0.002	0.183 ± 0.002	0.191 ± 0.003	0.181 ± 0.002
	20	0.189 ± 0.002	0.171 ± 0.003	0.178 ± 0.006	0.166 ± 0.002
	25	0.179 ± 0.002	0.159 ± 0.002	0.163 ± 0.003	0.156 ± 0.002
	30	0.172 ± 0.002	0.150 ± 0.002	0.154 ± 0.003	0.149 ± 0.001
$\beta = 0.5, a = 0.1,$	15	0.200 ± 0.002	0.182 ± 0.002	0.196 ± 0.004	0.182 ± 0.002
	20	0.189 ± 0.002	0.166 ± 0.002	0.171 ± 0.002	0.166 ± 0.002
	25	0.179 ± 0.003	0.157 ± 0.005	0.167 ± 0.006	0.154 ± 0.003
	30	0.173 ± 0.002	0.152 ± 0.002	0.158 ± 0.003	0.151 ± 0.002
$\beta = 0.5, a = 1,$	15	0.199 ± 0.002	0.181 ± 0.002	0.188 ± 0.002	0.180 ± 0.002
	20	0.182 ± 0.002	0.161 ± 0.002	0.171 ± 0.004	0.161 ± 0.002
	25	0.179 ± 0.002	0.157 ± 0.002	0.167 ± 0.004	0.156 ± 0.002
	30	0.174 ± 0.002	0.162 ± 0.006	0.168 ± 0.005	0.151 ± 0.001
$\beta = 1, a = 0.1,$	15	0.199 ± 0.001	0.185 ± 0.003	0.196 ± 0.004	0.181 ± 0.002
	20	0.189 ± 0.002	0.171 ± 0.003	0.180 ± 0.007	0.166 ± 0.002
	25	0.177 ± 0.002	0.157 ± 0.002	0.166 ± 0.004	0.156 ± 0.002
	30	0.172 ± 0.002	0.150 ± 0.001	0.158 ± 0.003	0.148 ± 0.001
$\beta = 1, a = 1,$	15	0.201 ± 0.001	0.185 ± 0.002	0.189 ± 0.002	0.186 ± 0.002
	20	0.187 ± 0.002	0.172 ± 0.003	0.176 ± 0.003	0.166 ± 0.002
	25	0.181 ± 0.002	0.159 ± 0.002	0.165 ± 0.003	0.157 ± 0.002
	30	0.172 ± 0.002	0.150 ± 0.001	0.152 ± 0.003	0.149 ± 0.001

The results reported in Table 2 reveal the good performance of Method P based on the misspecified covariances (19). The example below considers this spatio-temporal model a bit closer. It is not a model for spatially indexed functional data that drives our methodology, so it serves to underline the flexibility and extendability of our approach, which proposes to estimate the spatial components nonparametrically, when the number of spatial locations is small. It turns out that our nonparametric approach can improve on the current estimation methodology for model (19) as well.

Example 7.1. We now consider the model $Z(\mathbf{s}; t) = \mu + \varepsilon(\mathbf{s}; t)$, with the true mean $\mu = 0$. The spatial locations $\mathbf{s}_k$ are uniformly distributed on the unit square (different locations for every MC replication). The errors are normal with the space–time correlation function given by (19), i.e. by Eq. 14 in Gneiting (2002). The scale parameters a and c are nonnegative, the smoothness parameter α, and the space–time interaction parameter β take values in [0, 1]. The goal is to study the performance of the nonparametric method in three regimes: no space–time interaction $\beta = 0$, moderate space–time interaction $\beta = 0.5$, and strong space–time interaction $\beta = 1$. The parameter a is also of importance; if it is small, it induces long range temporal correlation, if it is large the temporal correlation decays fast. The space scale parameter and the smoothness parameter are fixed ($c = 5$, $\alpha = 0.7$). Again, to estimate the mean, we must estimate the covariance structure. The details are outlined at the end of this example. The methods we study are:

- S Simple average, which totally ignores the spatio-temporal dependence;
- T The infeasible method which uses the *correct* model and *true* parameter values;
- P The parametric method that uses the *correct* model and *estimated* parameters;
- N The new nonparametric method.

The results of our simulations are displayed in Table 3. In all but four cases, method N is significantly better than method P at 5% level of significance. In the remaining four cases, the average L_2 distance for method N is smaller, but the difference is smaller than two standard deviations. The difference is much more significant than 5% for $N = 15$ and $N = 20$. The marginally insignificant result for $N = 15$ and $\beta = 1, a = 1$ could be a type II error. In all cases considered the infeasible method (T) is not significantly better than N at the 5% level. Somewhat surprisingly, in a few cases with $N = 15$ and $N = 20$, method P is not significantly better than the trivial method S.

We conclude this example by outlining the covariance estimation procedure for model (19), see Section 4 in Gneiting (2002) for more details. By plugging in (19) into (9) one can see that for mean estimation only a purely spatial covariance is needed:

$$C(\mathbf{h}) = \int_0^1 C(\mathbf{h}; 0)dt = \exp(-c\|\mathbf{h}\|).$$

After determining the parameter c via nonlinear fitting we calculate the weights using (8). For method N, we estimate the covariance function nonparametrically, as described in Section 2.

Table 4
Empirical size of the test of $H_0 : \beta_2 = 0$.

p	Day			Noon			Night		
	10%	5%	1%	10%	5%	1%	10%	5%	1%
2	9.06	4.24	0.90	10.28	5.52	1.14	10.47	5.34	1.17
3	9.94	4.75	1.23	9.82	5.15	1.01	10.51	5.24	1.04
4	9.80	4.76	1.03	10.01	5.08	1.00	9.78	4.77	1.07
5	10.21	4.93	1.16	9.52	4.91	1.06	9.95	4.88	1.01

Table 5
Empirical size for the "simple" method.

p	Day			Noon			Night		
	10%	5%	1%	10%	5%	1%	10%	5%	1%
2	65.97	60.82	49.91	27.34	19.21	8.35	54.70	47.59	34.79
3	61.95	55.42	43.23	69.34	64.02	53.98	55.51	47.27	35.65
4	60.22	53.49	41.40	68.20	63.73	54.63	55.52	48.73	35.94
5	62.18	55.31	43.51	70.51	62.15	53.24	54.75	47.08	34.24

Methodology of Section 5. There are many reasonable data generating processes that could be used to evaluate the finite sample performance of the normal test based on $\hat{\beta}_i/\sqrt{\text{Var}[\hat{\beta}_i]}$. A number of spatial designs, shapes of the FPC's and the covariance functions for the scores could be employed. To avoid producing a large number of tables, we focus on a simulation study relevant to the science problem we consider in Section 6. The data generating processes are designed to resample true data.

To evaluate the empirical size and power, we generate the data using model (17) with $p = 3$ because the first three estimated FPC's explain about 91%–92% of the variance for each of the three types of data. The coefficients β_1 and β_3 are equal to these coefficients estimated from the real data. To evaluate the size, we set $\beta_2 = 0$, to study the power, we consider $0 < \beta_2 \leq 0.5$. The FPC's v_j are equal to those estimated from the real data. The vectors $\boldsymbol{\zeta}_j$ are $N(\mathbf{0}, \hat{\boldsymbol{\Gamma}}_j)$ with the $\hat{\boldsymbol{\Gamma}}_j$ equal to the covariance matrices estimated from the data (using the nonparametric method). Monte Carlo replications are generated by repeated simulations of the vectors $\boldsymbol{\zeta}_j$. The assumption of normality holds to a reasonable approximation as shown in Fig. 4 for the Day data. The plots for the Noon and Night data look similar. There is one outlying point in the QQ-plot of the $\zeta_2(\mathbf{s}_k)$, and the plot for the $\zeta_3(\mathbf{s}_k)$ indicates some departure from normality. The third FPC contributes however less than 10% to the variance, so its impact on our conclusions is small. In fact, the after removing this point, the trend parameter practically did not change.

The empirical size of the test developed in Section 5 is reported in Table 4 as a function of p. It is remarkably close to the nominal size and does not depend on p, as long as it remains in a reasonable range. This remains true if the scores are not normal, but severe departures from normality may distort the size for some values of p. For comparison, Table 5 shows the sizes for the method which can be termed "simple". It also relies on (13) and (15), but it uses the standard estimation procedure implemented in software packages discussed in Ramsay et al. (2009). (In the "simple" method, we set $w_k = 1/N$ and the v_j in $\boldsymbol{\Omega}$ are estimated as the eigenfunctions of the usual empirical covariance operator.) The "simple" method does not take into account the spatial dependence of the curves. This method severely overrejects. The empirical power of the new method is displayed in Fig. 5. The power curves for the Day data rise less steeply than those for the Noon and Night data. We will return to these curves when we discuss the results of the application of the test to the real data.

8. Summary and conclusions

The research reported in the paper is the outcome of our attempts to solve the hypothesis of global ionospheric cooling in a manner that would survive a rigorous scientific scrutiny. Our initial approaches failed because the standard parametric spatial model fitting produces very poor results if a small sample of spatial locations is available. To address this issue, we built on the nonparametric approach and developed a set of tools that can be used with confidence in small samples of spatially distributed curves. The main ingredients of the new methodology are the following. (1) Nonparametric estimation procedure for the mean function which produces significantly smaller mean squared errors than any of the existing procedures. (2) Estimation procedure for the mean function expressed as a linear combination of known functions. (3) A test to determine the significance of the coefficient of any of the known functions in (2). As explained in the Introduction, we hope that our methodology will be used in other problems of inference for functional data available at a small number of spatial locations.

Acknowledgment

Research partially supported by the NSF grant DMS/ATM 0931948.

Appendix. Bandwidth selection and the construction of the functional confidence intervals

In this section, we describe the procedures for the selection of the bandwidth h and for the construction of the functional confidence intervals, the two important ingredients of the methodology introduced in Section 2.

Regarding the choice of h, we performed a very extensive simulation study to evaluate the performance of several potential methods. It is important to note that the nonparametric covariance function estimation described in Section 2 is only an ingredient of a broader methodology for the estimation and testing in the functional spatio-temporal framework. The choice of the bandwidth must thus be tailored to the problems we want to solve. These problems revolve around the estimation of the mean function $\mu(t)$ in model (4). The procedure employed in all numerical work reported in this paper is the following: We first center the functions by their average (their possible spatial dependence is not taken into account). This step transforms the functions to a set of *approximately* mean zero functions. We then estimate the covariance functions using several choices of h, the same h for every functional principal component. We select h for which the estimated mean function is closest to zero. In other words, we select h which minimizes

$$\|\hat{\mu}_h - 0\|^2 = \int \hat{\mu}_h^2(t)dt$$

where $\hat{\mu}_h$ is the final estimated mean function (of the centered curves) obtained using bandwidth h.

The other approaches we experimented with included: (1) cross-validation to minimize the integrated mean squared error of $\hat{m}$ given by (2) or of $\hat{\gamma}$ described in the next paragraph. (2) Cross-validation to minimize the integrated mean squared error of the estimated function $\hat{\mu}_h$. (3) Spatial versions of the cross validations in (1) and (2), in which the removed observation is replaced by a spatial prediction obtained using kriging with various values of h. None of these approaches yielded uniformly satisfactory results.

We now turn to the construction of functional confidence bounds for the covariance function $\gamma(\cdot)$. The idea is as follows: First, using the quasi-bootstrap procedure proposed by Solow (1985) and further improved by Clark and Allingham (2011), we produce a collection of M independent covariance curves, $\hat{\gamma}_i(d)$, $1 \le i \le M$. Then using the concept of a functional depth we construct the confidence bounds. The concept of functional depth has been extensively used lately, see Fraiman and Muniz (2001), Febrero et al. (2008), López-Pintado and Romo (2009) and Sun and Genton (2011), but not in the context of covariances of spatial data.

We start with outlining the quasi-bootstrap procedure. The estimated covariance matrix $\hat{\boldsymbol{\Gamma}}$ is decomposed using the Cholesky decomposition as $\hat{\boldsymbol{\Gamma}} = \hat{\mathbf{L}}\hat{\mathbf{L}}^T$, where $\hat{\mathbf{L}}$ is the lower triangular matrix. Using $\hat{\mathbf{L}}$, the spatial field $\boldsymbol{\zeta}$ can be decorrelated as $\boldsymbol{\zeta}_0 = \hat{\mathbf{L}}^{-1}\boldsymbol{\zeta}$. Next, $\boldsymbol{\zeta}_0$ is resampled with replacement and recorrelated $\boldsymbol{\zeta}^i = \hat{\mathbf{L}}\boldsymbol{\zeta}_0$. The superscript $1 \le i \le M$ refers to the iteration step. Based on the bootstrap sample $\boldsymbol{\zeta}^i$, the correlation is estimated by $\hat{\gamma}_i(d)$ using the nonparametric method. These steps are repeated sufficiently many, say $M = 1000$, times. This leads to a collection of independent covariance curves $\hat{\gamma}_i(d)$.

A functional depth can be defined in many different ways. Here we use the definition presented in Chapter 1 of Horváth and Kokoszka (2012). Let

$$F_{N,d}(\gamma) = \frac{1}{N}\sum_{i=1}^{N} \mathbf{I}\left\{\hat{\gamma}_i(d) \le \gamma\right\},$$

be the empirical distribution function at point d. We define the functional depth (FD) of the curve $\hat{\gamma}_i$ by integrating the univariate depth:

$$\mathrm{FD}_N(\hat{\gamma}_i) = \int \left(1 - |1/2 - F_{N,d}(\hat{\gamma}(s))|\right) ds.$$

Next, we select K curves with the largest depth. Then, for each d we find $\min(\hat{\gamma}_i(d))$ and $\max(\hat{\gamma}_i(d))$ from among these K curves. This produces the lower and upper bounds which form the functional analogue of the univariate $1 - \alpha = K/M$ confidence interval. An example of the functional confidence bounds is shown in Fig. 1.

References

Bowman, A.W., Giannitrapani, M., Scott, E.M., 2009. Spatiotemporal smoothing and sulphur dioxide trends over Europe. Journal of the Royal Statistical Society (C) 58, 737–752.

Clark, R.G., Allingham, S., 2011. Robust resampling confidence intervals for empirical variograms. Mathematical Geosciences 43, 243–259.

Clilverd, M., Ulich, T., Jarvis, M.J., 2003. Residual solar cycle influence on trends in ionospheric F2-layer peak height. Journal of Geophysical Research 108, 1450.

Cressie, N., Huang, H.-C., 1999. Classes of nonseparable, spatio-temporal stationary covariance functions. Journal of the American Statistical Association 94, 1330–1340.

Cressie, N., Shi, T., Kang, E.L., 2010. Fixed rank filtering for spatio-temporal data. Journal of Computational and Graphical Statistics 19, 724–745.

Cressie, N., Wikle, C., 2011. Statistics for Spatio-Temporal Data. Wiley.

Delicado, P., Giraldo, R., Comas, C., Mateu, J., 2010. Statistics for spatial functional data: some recent contributions. Environmetrics 21, 224–239.

Fan, J., Gijbels, I., 1996. Local Polynomial Modelling and its Applications. Chapman & Hall/CRC.

Febrero, M., Galeano, P., González-Manteiga, W., 2008. Outlier detection in functional data by depth measures with application to identify abnormal NO_x levels. Environmetrics 19, 331–345. http://dx.doi.org/10.1002/env.878.

Fraiman, R., Muniz, G., 2001. Trimmed means for functional data. Test 10, 419–440.
Giraldo, R., Delicado, P., Mateu, J., 2011. Ordinary kriging for function-valued spatial data. Environmental and Ecological Statistics 18, 411–426.
Gneiting, T., 2002. Nonseparable, stationary covariance functions for space–time data. Journal of the American Statistical Association 97, 590–600.
Gneiting, T., Genton, M.G., Guttorp, P., 2007. Geostatistical space–time models, stationarity, separability and full symmetry. In: Finkenstaedt, B., Held, L., Isham, V. (Eds.), Statistics of Spatio-Temporal Systems. Chapman & Hall/CRC Press, pp. 151–175.
Gromenko, O., Kokoszka, P., 2012. Testing the equality of mean functions of spatially distributed curves. Journal of the Royal Statistical Society (C) 61, 715–731.
Gromenko, O., Kokoszka, P., Zhu, L., Sojka, J., 2012. Estimation and testing for spatially indexed curves with application to ionospheric and magnetic field trends. The Annals of Applied Statistics 6, 669–696.
Hall, P., Fisher, N.I., Hoffmann, B., 1994. On the nonparametric estimation of covariance functions. The Annals of Statistics 22, 2115–2134.
Hall, P., Patil, P., 1994. Properties of nonparametric estimators of autocovariance for stationary random fields. Probability Theory and Related Fields 99, 399–424.
Haslett, J., Raftery, A.E., 1989. Space–time modelling with long-memory dependence: assessing Ireland's wind power resource. Applied Statistics 38 (1), 1–50.
Horváth, L., Kokoszka, P., 2012. Inference for Functional Data with Applications. Springer.
Hörmann, S., Kokoszka, P., 2013. Consistency of the mean and the principal components of spatially distributed functional data. Bernoulli, 00, 0000–0000 (forthcoming).
Jun, M., Stein, M.L., 2009. Nonstationary covariance models for global data. The Annals of Applied Statistics 2, 1271–1289.
Katzfuss, M., Cressie, N., 2011. Spatio-temporal smoothing and EM estimation for massive remote-sensing data sets. Journal of Time Series Analysis 32, 430–446.
Lastovicka, J., Akmaev, R.A., Beig, G., Bremer, J., Emmert, J.T., Jacobi, C., Jarvis, J.M., Nedoluha, G., Portnyagin, Y.I., Ulich, T., 2008. Emerging pattern of global change in the upper atmosphere and ionosphere. Annales Geophysicae 26, 1255–1268.
Lastovicka, J., Mikhailov, A.V., Ulich, T., Bremer, J., Elias, A.G., Ortiz de Adler, N., Jara, V., Abarca del Rio, R., Foppiano, A.J., Ovalle, E., Danilov, A.D., 2006. Long term trends in foF2: a comparison of various methods. Journal of Atmospheric and Solar-Terrestrial Physics 68, 1854–1870.
López-Pintado, S., Romo, J., 2009. On the concept of depth for functional data. Journal of the American Statistical Association 104, 718–734.
Nerini, D., Monestiez, P., Mantéa, C., 2010. Cokriging for spatial functional data. Journal of Multivariate Analysis 101, 409–418.
Ramsay, J., Hooker, G., Graves, S., 2009. Functional Data Analysis with R and MATLAB. Springer.
Ramsay, J.O., Silverman, B.W., 2005. Functional Data Analysis. Springer.
Rishbeth, H., 1990. A greenhouse effect in the ionosphere? Planetary and Space Science 38, 945–948.
Roble, R.G., Dickinson, R.E., 1989. How will changes in carbon dioxide and methane modify the mean structure of the mesosphere and thermosphere? Geophysical Research Letters 16, 1441–1444.
Sherman, M., 2011. Spatial Statistics and Spatio-Temporal Data: Covariance Functions and Directional Properties. In: Wiley Series in Probability and Statistics, Wiley.
Solow, A., 1985. Bootstrapping correlated data. Mathematical Geosciences 17, 769–775.
Sun, Y., Genton, M.G., 2011. Functional boxplots. Journal of Computational and Graphical Statistics 20, 316–334.