

Supplement to “Consistency of the mean and the principal components of spatially distributed functional data”

Siegfried Hörmann* and Piotr Kokoszka†

S.1 Examples of spatially distributed functional data

As indicated in the paper, the Canadian temperature and precipitation data used in Ramsay and Silverman (2005) serve as good example for functional data sampled at spatial locations. We provide below a small selection of some other examples which have either been studied in the literature or which we have been motivated by our own research.

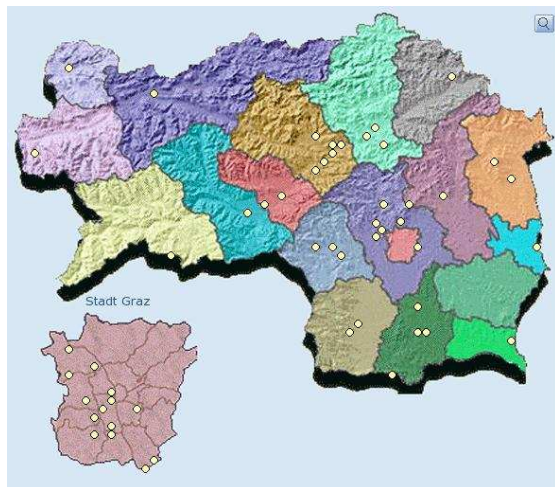
1. The Australian rainfall data set, has been recently used by Delaigle and Hall (2010). It consists of daily rainfall measurements from 1840 to 1990 at 191 Australian weather stations.

2. Pollution curves: $X(\mathbf{s}_k; t)$ is the concentration of a pollutant at time t at location \mathbf{s}_k . Data of this type were studied by Kaiser *et al.* (2002). A functional framework might be convenient because such data are typically available only at sparsely distributed time points t_j which can be different at different locations. Since pollution data usually come in addition with measurement errors, the data have to be firstly preprocessed and transformed in functional form using techniques as e.g. described in Ramsay and Silverman (2005). Often this is done for each observations separately, by using a so-called basis functions approach in connection with some penalized least-squares criterion. It is an interesting open problem how the (spatial) dependence could be used to improve the preprocessing step. Figure S.1.1 shows the monitoring network for different pollutants in the province of Styria (Austria).

*Département de Mathématique, Université Libre de Bruxelles, CP 210, Bd du Triomphe, B-1050 Brussels - Belgium. Email: shormann@ulb.ac.be

†Department of Statistics, Colorado State University, Fort Collins, CO 80523-1877, USA. Email: Piotr.Kokoszka@colostate.edu

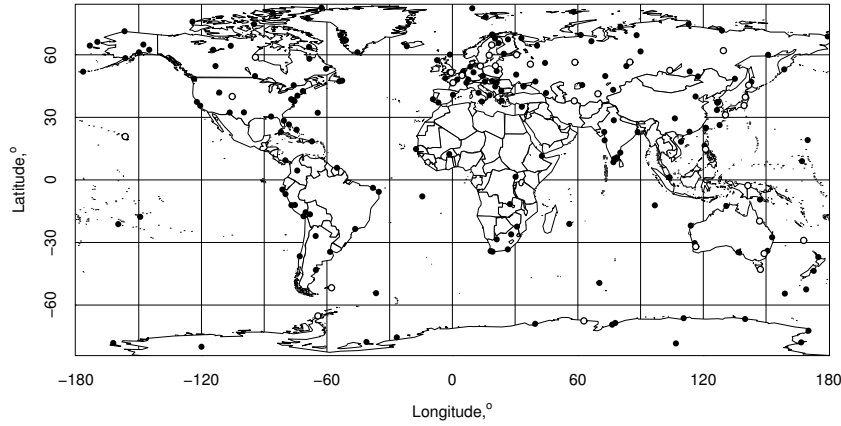
FIGURE S.1.1 Circles indicate the monitoring network for different pollutants in the province of Styria (Austria).



3. Snow water curves are measured at several dozen locations in every state over many decades. Such data have been studied in the spatial framework, e.g. Carroll *et al.* (1995) and Carroll and Cressie (1996), but useful insights can be gained by studying the whole curves reflecting the temporal dynamics. In many studies, $X(\mathbf{s}_k; t)$ is the count at time t of disease cases, where \mathbf{s}_k represents an average location in an areal model.

4. The data set that most directly motivated this research consists of the curves of the so-called F2-layer critical frequency $foF2$. In principle, $foF2$ curves are available at close to 200 locations throughout the globe, but sufficiently complete data are available at only 30-40 locations which are very unevenly spread; for example, there is a dense network of observatories over Europe and practically no data over the oceans, see Figure S.1.2. The study of this data set has been motivated by the hypothesis of Roble and Dickinson (1989) who suggested that the increasing amounts of greenhouse gases should lead to global cooling in mesosphere and thermosphere, as opposed to the global warming in lower troposphere. Rishbeth (1990) pointed out that such cooling would result in a thermal contraction and the global lowering of the ionospheric peak densities, which can be computed from the critical frequency $foF2$. The last twenty years have seen very extensive research in this area, see Lastovicka *et al.* (2008) for a partial overview. One of the difficulties is in finding a global trend for curves which appear to exhibit trends in opposing directions over various regions. Ulich *et al.* (2003) stressed that to make any trends believable, a suitable statistical modeling, and a proper treatment of “errors and uncertainties” is called for. Space physics data measured at terrestrial observatories always come in the form of temporal curves at fixed spatial locations. Maslova *et al.* (2009, 2010a, 2010b) used the tools of functional data analysis to study such data, but

FIGURE S.1.2 Locations of 218 ionosonde stations. Circles represent the 32 stations with the longest complete records.



the spatial dependence of the curves was not fully exploited.

S.2 Additional remarks to Assumption 2

At first sight, Assumption 2 may appear a bit daunting. It is used to obtain a rate of convergence for the empirical covariance operator. In this section we will provide some details and discussion.

Suppose $X(\mathbf{s}; t)$ is an arbitrary functional random field observed at locations $\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_N$. Then a direct calculation shows that quite generally

$$(S.1) \quad NE \|\widehat{C}_N - C\|_{\mathcal{S}}^2 = N^{-1} \sum_{k, \ell=1}^N \iint \text{Cov}(X(\mathbf{s}_k; t)X(\mathbf{s}_k; u), X(\mathbf{s}_\ell; t)X(\mathbf{s}_\ell; u)) dt du.$$

Without any further assumptions, a sufficient condition for the EFPC's to be consistent with the rate $N^{-1/2}$ is that the right-hand side of (S.1) is bounded from above by a constant. Under additional assumptions, more precise sufficient conditions are possible.

EXAMPLE S.2.1 Suppose $X_k = X(\mathbf{s}_k)$, where $\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_N$ are points in an arbitrary metric space, and the random field $X(\cdot)$ is such that $X(\mathbf{s})$ is independent of $X(\mathbf{s}')$ if the distance between \mathbf{s} and \mathbf{s}' , $d(\mathbf{s}, \mathbf{s}')$, is greater than m (m is a fixed distance and we continue to assume that the X_k have the same distribution). In this case Assumption 2 is clearly satisfied. Set

$$B_N(m) = \{(k, \ell) : 1 \leq k, \ell \leq N \text{ and } d(\mathbf{s}_k, \mathbf{s}_\ell) \leq m\},$$

and denote by $|B_N(m)|$ the count of pairs in $B_N(m)$. A brief calculation, which uses the Cauchy inequality twice, leads to the bound

$$NE \|\widehat{C}_N - C\|_{\mathcal{S}}^2 \leq N^{-1} |B_N(m)| E \|X(\mathbf{s})\|^4.$$

If the \mathbf{s}_k are the points in \mathbb{R}^d with integer coordinates, then $|B_N(m)|$ is asymptotically proportional to mN , implying $\limsup_{N \rightarrow \infty} N^{-1} |B_N(m)| < \infty$, and the standard rate of consistency.

Suppose $\{e_j, j \geq 1\}$ is an arbitrary *fixed* orthonormal basis in L^2 . Under very mild assumptions, every constant mean functional random field admits the representation

$$(S.2) \quad X(\mathbf{s}) = \mu + \sum_{j \geq 1} \xi_j(\mathbf{s}) e_j,$$

where the $\xi_j(\mathbf{s})$ are zero mean random variables. In principle, all properties of X , including the spatial dependence structure, can be equivalently stated as properties of the family of the scalar fields ξ_j . In general Assumption 2 cannot be verified using only conditions on the covariances of the scalar fields ξ_j in (S.2) because these covariances do not specify the 4th order structure of the model. A very convenient assumption is that the ξ_j are *independent* scalar fields. Before we discuss how such an assumption could be used to simplify Assumption 2, let us motivate why assuming independence is reasonable in the spatial context.

EXAMPLE S.2.2 (Kriging). Assume that $\mu = 0$ and suppose we want to predict $X(\mathbf{s}_0)$ using a linear combination of the curves $X(\mathbf{s}_1), X(\mathbf{s}_2), \dots, X(\mathbf{s}_N)$, i.e. we want to minimize

$$(S.3) \quad E \left\| X(\mathbf{s}_0) - \sum_{n=1}^N a_n X(\mathbf{s}_n) \right\|^2 \\ = E \langle X(\mathbf{s}_0), X(\mathbf{s}_0) \rangle - 2 \sum_{n=1}^N a_n E \langle X(\mathbf{s}_n), X(\mathbf{s}_0) \rangle + \sum_{k,\ell=1}^N a_k a_\ell E \langle X(\mathbf{s}_k), X(\mathbf{s}_\ell) \rangle.$$

Thus for the problem of the least squares linear prediction of a mean zero spatial process we need to know only

$$(S.4) \quad K(\mathbf{s}, \mathbf{s}') = E [\langle X(\mathbf{s}), X(\mathbf{s}') \rangle].$$

By the orthonormality of the e_j in (S.2) (use Parseval's identity),

$$E [\langle X(\mathbf{s}), X(\mathbf{s}') \rangle] = \sum_{j=1}^{\infty} E [\xi_j(\mathbf{s}) \xi_j(\mathbf{s}')].$$

Thus, the functional covariances (S.4) are fully determined by the covariances

$$(S.5) \quad K_j(\mathbf{s}, \mathbf{s}') = E [\xi_j(\mathbf{s}) \xi_j(\mathbf{s}')].$$

As we do not need to know the cross covariances $E[\xi_j(\mathbf{s}) \xi_i(\mathbf{s}')] for $i \neq j$, for the kriging we can assume that the spatial processes $\xi_j(\cdot)$ in (S.2) are independent. Such an assumption simplifies the verification of some fourth order properties which we will discuss now.$

LEMMA S.2.1 Let $X(\mathbf{s})$ have representation (S.2) with zero mean and $E\|X(\mathbf{s})\|^4 < \infty$. Assume further that $\xi_i(\cdot)$ and $\xi_j(\cdot)$ are independent if $i \neq j$. Then

$$\begin{aligned} & |E\langle X(\mathbf{s}_1) \otimes X(\mathbf{s}_1) - C, X(\mathbf{s}_2) \otimes X(\mathbf{s}_2) - C \rangle_S| \\ & \leq \left| \sum_{j \geq 1} \text{Cov}(\xi_j^2(\mathbf{s}_1), \xi_j^2(\mathbf{s}_2)) \right| + \left| \sum_{j \geq 1} E[\xi_j(\mathbf{s}_1)\xi_j(\mathbf{s}_2)] \right|^2. \end{aligned}$$

PROOF: If $\xi_i(\cdot)$ and $\xi_j(\cdot)$ are independent for $i \neq j$, then the e_j are the eigenvalues of C , and the $\xi_j(\mathbf{s})$ are the principal component scores with $E\xi_j^2(\mathbf{s}) = \lambda_j$. Using continuity of the inner product and dominated convergence we obtain

$$\begin{aligned} & |E\langle X(\mathbf{s}_1) \otimes X(\mathbf{s}_1) - C, X(\mathbf{s}_2) \otimes X(\mathbf{s}_2) - C \rangle_S| \\ & = \left| E \sum_{j \geq 1} \left\langle \langle X(\mathbf{s}_1), e_j \rangle X(\mathbf{s}_1) - C(e_j), \langle X(\mathbf{s}_2), e_j \rangle X(\mathbf{s}_2) - C(e_j) \right\rangle \right| \\ & = \left| E \sum_{j \geq 1} \left\langle \xi_j(\mathbf{s}_1) \sum_{\ell \geq 1} \xi_\ell(\mathbf{s}_1) e_\ell - \lambda_j e_j, \xi_j(\mathbf{s}_2) \sum_{k \geq 1} \xi_k(\mathbf{s}_2) e_k - \lambda_j e_j \right\rangle \right| \\ & = \left| E \sum_{j \geq 1} \left\{ \xi_j(\mathbf{s}_1)\xi_j(\mathbf{s}_2) \sum_{\ell \geq 1} \xi_\ell(\mathbf{s}_1)\xi_\ell(\mathbf{s}_2) + \lambda_j^2 - \lambda_j\xi_j^2(\mathbf{s}_1) - \lambda_j\xi_j^2(\mathbf{s}_2) \right\} \right| \\ & \leq \left| \sum_{j \geq 1} \text{Cov}(\xi_j^2(\mathbf{s}_1), \xi_j^2(\mathbf{s}_2)) \right| + \left| \sum_{j \geq 1} \sum_{\ell \neq j} E[\xi_j(\mathbf{s}_1)\xi_j(\mathbf{s}_2)] \times E[\xi_\ell(\mathbf{s}_1)\xi_\ell(\mathbf{s}_2)] \right| \\ & \leq \left| \sum_{j \geq 1} \text{Cov}(\xi_j^2(\mathbf{s}_1), \xi_j^2(\mathbf{s}_2)) \right| + \left| \sum_{j \geq 1} E[\xi_j(\mathbf{s}_1)\xi_j(\mathbf{s}_2)] \right|^2. \end{aligned}$$

■

EXAMPLE S.2.3 Suppose first that representation (S.2) holds with independent strictly stationary scalar fields ξ_j . Define the covariances

$$E[\xi_j(\mathbf{s}_k)\xi_j(\mathbf{s}_\ell)] = \gamma_j(\mathbf{s}_k - \mathbf{s}_\ell), \quad \text{Cov}(\xi_j^2(\mathbf{s}_k), \xi_j^2(\mathbf{s}_\ell)) = \tau_j(\mathbf{s}_k - \mathbf{s}_\ell).$$

Using (S.1), we see that under these assumptions,

$$NE\|\widehat{C} - C\|_S^2 = N^{-1} \sum_{k, \ell=1}^N \left\{ \sum_{i \neq j} \gamma_i(\mathbf{s}_k - \mathbf{s}_\ell)\gamma_j(\mathbf{s}_k - \mathbf{s}_\ell) + \sum_{j=1}^{\infty} \tau_j(\mathbf{s}_k - \mathbf{s}_\ell) \right\}.$$

Thus the standard rate holds, if

$$(S.6) \quad \limsup_{N \rightarrow \infty} N^{-1} \sum_{k, \ell=1}^N \left\{ \sum_{j=1}^{\infty} \gamma_j(\mathbf{s}_k - \mathbf{s}_\ell) \right\}^2 < \infty$$

and

$$(S.7) \quad \limsup_{N \rightarrow \infty} N^{-1} \sum_{k,\ell=1}^N \sum_{j=1}^{\infty} |\tau_j(\mathbf{s}_k - \mathbf{s}_\ell)| < \infty.$$

For the next example we need the following Lemma.

LEMMA S.2.2 *Suppose X and Y are jointly normal mean zero random variables such that $EX^2 = \sigma^2$, $EY^2 = \nu^2$, $E[XY] = \rho\sigma\nu$. Then*

$$\text{Cov}(X^2, Y^2) = 2\rho^2\sigma^2\nu^2.$$

EXAMPLE S.2.4 We impose now as further assumption that X is Gaussian. If

$$(S.8) \quad E[\xi_j(\mathbf{s}_k)\xi_j(\mathbf{s}_\ell)] = \sigma_j^2 \exp\{-\rho_j^{-1}d(\mathbf{s}_k, \mathbf{s}_\ell)\}$$

then by Lemma S.2.2

$$(S.9) \quad \text{Cov}(\xi_j^2(\mathbf{s}_k), \xi_j^2(\mathbf{s}_\ell)) = 2\sigma_j^4 \exp\{-2\rho_j^{-1}d(\mathbf{s}_k, \mathbf{s}_\ell)\}.$$

If the scalar fields ξ_j are independent and if (2.11) in our paper holds, it follows that for some large enough constant A ,

$$\begin{aligned} & \left| \sum_{j \geq 1} \text{Cov}(\xi_j^2(\mathbf{s}_1), \xi_j^2(\mathbf{s}_2)) \right| + \left| \sum_{j \geq 1} E[\xi_j(\mathbf{s}_1)\xi_j(\mathbf{s}_2)] \right|^2 \\ & \leq A \exp(-2\rho^{-1}\|\mathbf{s}_1 - \mathbf{s}_2\|_2). \end{aligned}$$

Hence by Lemma S.2.1, Assumption 2 holds with $H(x) = A \exp(-2\rho^{-1}\|\mathbf{s}_1 - \mathbf{s}_2\|_2)$.

S.3 Additional remarks to the regular sampling designs

In this section we provide the details to the regular sampling designs described in Section 2.4 and 2.5 of our paper. For ease of reference, we reformulate the results of these two sections.

S.3.1 Non-random regular design

Let $\mathcal{Z}(\boldsymbol{\delta})$ be a lattice in \mathbb{R}^d with increments δ_i in the i -th direction. Let $\delta_0 = \min\{\delta_1, \dots, \delta_d\}$, $\Delta^d = \prod_{i=1}^d \delta_i$ and let $R_N = \alpha_N R_0$, where R_0 is some bounded Riemann measurable Borel-set in \mathbb{R}^d containing the origin. A set is Riemann measurable if its indicator function is

Riemann integrable. This condition excludes highly irregular sets R_0 . The scaling parameters $\alpha_N > 0$ are assumed to be non-decreasing and will be specified below in Lemma S.3.2. We assume without loss of generality that $\text{Vol}(R_0) = 1$, hence $\text{Vol}(R_N) = \alpha_N^d$. Typical examples are $R_0 = \{x \in \mathbb{R}^d : \|x\| \leq z_{1,d}\}$, with $z_{1,d}$ equal to the radius of the d -dimensional sphere with volume 1, or $R_0 = [-1/2, 1/2]^d$. The sampling points \mathfrak{S}_N are defined as $\{\mathbf{s}_{k,N}, 1 \leq k \leq S_N\} = \mathcal{Z}(\eta_N \boldsymbol{\delta}) \cap R_N$, where η_N is chosen such that the sample size $S_N \sim N$. It is intuitively clear that $\text{Vol}(R_N) \approx \eta_N^d \Delta^d S_N$, suggesting

$$(S.1) \quad \eta_N = \frac{\alpha_N}{\Delta N^{1/d}}.$$

A formal proof that η_N in (S.1) assures $S_N \sim N$ is immediate from the following

LEMMA S.3.1 *Let K be a bounded set in \mathbb{R}^d , and assume that K is Riemann measurable with $\text{Vol}(K) = 1$. If $\beta_N \rightarrow 0$, then*

$$|K \cap \mathcal{Z}(\beta_N \boldsymbol{\delta})| \sim \frac{1}{\Delta^d \beta_N^d}.$$

PROOF: Let $K \subset M_1 \subset M_2$ where M_1 and M_2 are rectangles in \mathbb{R}^d having no intersecting margin (M_1 is an inner subset of M_2). The points $\{x_{i,N}\} = \mathcal{Z}(\beta_N \boldsymbol{\delta}) \cap M_2$ can be seen as the vertices of rectangles $J_{i,N} = x_{i,N} + \{\mathbf{t} \circ \beta_N \boldsymbol{\delta}, \mathbf{t} \in [0, 1]^d\}$, where \circ denotes the Hadamard (entrywise) product. For large enough N , the sets $L_{i,N} = J_{i,N} \cap M_1$ define a partition of M_1 . Then, by the assumed Riemann measurability,

$$\begin{aligned} & \int_{M_1} I_K(x) dx \\ &= \liminf_{N \rightarrow \infty} \beta_N^d \Delta^d \sum_i \inf\{I_K(x) : x \in L_{i,N}\} \\ &\leq \liminf_{N \rightarrow \infty} \beta_N^d \Delta^d \sum_i I_K(x_{i,N}) \\ &\leq \limsup_{N \rightarrow \infty} \beta_N^d \Delta^d \sum_i I_K(x_{i,N}) \\ &\leq \limsup_{N \rightarrow \infty} \beta_N^d \Delta^d \sum_i \sup\{I_K(x) : x \in L_{i,N}\} \\ &= \int_{M_1} I_K(x) dx. \end{aligned}$$

■

LEMMA S.3.2 *In the above described design the following pairs of statements are equivalent:*

- (i) α_N remains bounded \Leftrightarrow Type A sampling;
- (ii) $\alpha_N \rightarrow \infty$ and $\alpha_N = o(N^{1/d}) \Leftrightarrow$ Type B sampling;
- (iii) $\alpha_N \gg N^{1/d} \Leftrightarrow$ Type C sampling.

PROOF: Let $U_\varepsilon(x)$ be the sphere in \mathbb{R}^d with center x and radius ε . Assume first that $\alpha_N = o(N^{1/d})$, which covers (i) and (ii). In this case the volume of the rectangles $L_{i,n}$ as described in the proof of Lemma S.3.1 satisfies

$$(S.2) \quad \text{Vol}(L_{i,n}) = \Delta^d \eta_N^d = \frac{\alpha_N^d}{N} \rightarrow 0.$$

Hence $|U_\rho(x) \cap \mathcal{Z}(\eta_N \boldsymbol{\delta})|$ is asymptotically proportional to

$$\text{Vol}(U_\rho(x))/\text{Vol}(L_{i,n}) = V_d \left(\frac{\rho}{\alpha_N} \right)^d N,$$

where V_d is the volume of the d -dimensional unit sphere. Now if we fix an arbitrary $\rho_0 > 0$ then there are constants $0 < C_L < C_U < \infty$, such that for any $\rho \geq \rho_0$ and $N \geq N_0$ and $x \in \mathbb{R}^d$

$$C_L \left(\frac{\rho}{\alpha_N} \right)^d \leq \frac{|U_\rho(x) \cap \mathcal{Z}(\eta_N \boldsymbol{\delta})|}{N} \leq C_U \left(\frac{\rho}{\alpha_N} \right)^d.$$

By the required Riemann measurability we can find an $x \in R_0$ such that for some small enough ε we have $U_{2\varepsilon}(x) \subset R_0$. Then $U_{2\varepsilon\alpha_N}(\alpha_N x) \subset R_N$. Hence for any $2\rho_0 \leq \rho \leq \varepsilon\alpha_N$,

$$C_L \left(\frac{\rho}{\alpha_N} \right)^d \leq \frac{|U_{\rho/2}(\alpha_N x) \cap \mathfrak{S}_N|}{N} \leq I_\rho(\mathfrak{S}_N) \leq \frac{|U_{2\rho}(\alpha_N x) \cap \mathfrak{S}_N|}{N} \leq C_U \left(\frac{\rho}{\alpha_N} \right)^d.$$

With the help of the above inequalities (i) and (ii) are easily checked.

Now we prove (iii). We notice that by (S.2) $\alpha_N \gg N^{1/d}$ is equivalent to $\text{Vol}(L_{i,n})$ does not converge to 0. Assume first that we have Type C sampling. Then by the arguments above we find an x and a $\rho > 0$ such that $U_\rho(\alpha_N x) \subset R_N$. Thus

$$|U_\rho(\alpha_N x) \cap \mathcal{Z}(\eta_N \boldsymbol{\delta})| \leq S_N I_\rho(\mathfrak{S}_N).$$

As this quantity remains bounded, $\text{Vol}(L_{i,n})$ does not converge to 0.

On the other hand, if $\text{Vol}(L_{i,n})$ does not converge to 0 then for any $\rho > 0$ and any $x \in \mathbb{R}^d$ we have $\limsup_{N \rightarrow \infty} |U_\rho(x) \cap \mathcal{Z}(\eta_N \boldsymbol{\delta})| < \infty$ and thus for arbitrary large ρ

$$I_\rho(\mathfrak{S}_N) \leq \sup_x \frac{|U_\rho(x) \cap \mathcal{Z}(\eta_N \boldsymbol{\delta})|}{S_N} \rightarrow 0.$$

The claim follows immediately. ■

S.3.2 Randomized design

Let $\{\mathbf{s}_k, 1 \leq k \leq N\}$ be iid random vectors with a density $f(\mathbf{s})$ which has support on a Borel set $R_0 \subset \mathbb{R}^d$ containing the origin and satisfying $\text{Vol}(R_0) = 1$. Again we assume

Riemann measurability for R_0 to exclude highly irregular sets. For the sake of simplicity we shall assume that on R_0 the density is bounded away from zero, so that we have $0 < f_L \leq \inf_{x \in R_0} f(x)$. The point set $\{\mathbf{s}_{k,N}, 1 \leq k \leq N\}$ is defined by $\mathbf{s}_{k,N} = \alpha_N \mathbf{s}_k$ for $k = 1, \dots, N$. For fixed N , this is equivalent to: $\{\mathbf{s}_{k,N}, 1 \leq k \leq N\}$ is an iid sequence on $R_N = \alpha_N R_0$ with density $\alpha_N^{-d} f(\alpha_N^{-1} \mathbf{s})$.

We cannot expect to obtain a full analogue of Lemma S.3.2 in the randomized setup. For Type C sampling, the problem is much more delicate, and a closer study shows that it is related to the oscillation behavior of multivariate empirical processes. While Stute (1984) gives almost sure upper bounds, we would need here sharp results on the moments of the modulus of continuity of multivariate empirical process. Such results exist, see Einmahl and Ruymgaart (1987), but are connected to technical assumptions on the bandwidth for the modulus (here determined by α_N) which are not satisfied in our setup. Hence a detailed treatment would go beyond the scope of this paper. We thus state here the following lemma.

LEMMA S.3.3 *In the above described sampling scheme the following statements hold:*

- (i) α_N remains bounded \Rightarrow Type A sampling;
- (ii) $\alpha_N \rightarrow \infty$ and $\alpha_N = o(N^{1/d}) \Rightarrow$ Type B sampling;

PROOF. By Jensen's inequality we infer that

$$\begin{aligned} EI_\rho(\mathfrak{S}_N) &= E \sup_{x \in R_N} \frac{1}{N} \sum_{k=1}^N I\{\mathbf{s}_{k,N} \in U_\rho(x) \cap R_N\} \\ &\geq \sup_{x \in R_N} P(\mathbf{s}_{1,N} \in U_\rho(x) \cap R_N) \\ &= \sup_{x \in R_0} P(\mathbf{s}_1 \in U_{\rho/\alpha_N}(x) \cap R_0) \\ &= \sup_{x \in R_0} \int_{U_{\rho/\alpha_N}(x) \cap R_0} f(s) ds. \end{aligned}$$

We have two scenarios. First, α_N remains bounded. Then we can choose ρ big enough such that $U_{\rho/\alpha_N}(0)$ covers R_0 for all N . It follows that $\limsup_{N \rightarrow \infty} EI_\rho(\mathfrak{S}_N) = 1$ and (i) follows.

Second, $\alpha_N \rightarrow \infty$. Then for large enough N , R_0 contains a ball with radius ρ/α_N . It follows that

$$(S.3) \quad EI_\rho(\mathfrak{S}_N) \geq f_L V_d \left(\frac{\rho}{\alpha_N} \right)^d.$$

Now statement (ii) follows easily. ■

S.4 A remark to the sharpness of the convergence rates

We see from Examples 3.1 and 3.2 of our paper, that imposing certain regularity assumptions in the sampling design will provide faster rates of consistency. More specifically we saw that the bounds provided by Proposition 3.2 (4.2), are sharper than those given by Proposition 3.1 and 4.1, respectively. The question is whether we can improve the rates obtained in Proposition 3.1 (4.1) under the general formulation.

The answer is *no*. The rate (3.2) is optimal under the general assumptions of Proposition 3.1. We construct an examples which attains the bound.

EXAMPLE S.4.1 Let $X(s; t) = \psi(s)e(t)$, with $s \in \mathbb{R}$, $t \in [0, 1]$, $\int_0^1 e^2(t)dt = 1$ and

$$\psi(s) = \sum_{k \in \mathbb{Z}} I\{s \in (U + k, U + k + 1]\} \delta_k,$$

where $\{\delta_k\}$ is an iid sequence with $\delta_1 = \pm 1$, each with probability 1/2 and U is uniformly distributed on $[0, 1]$ and independent of $\{\delta_k\}$. A simple calculation shows that $EX(s; t) = 0$ for all s, t and that $E\langle X(u), X(v) \rangle = (1 - |u - v|)I\{|u - v| \leq 1\}$. Let

$$\mathfrak{S}_N = \left\{ \frac{k\alpha_N}{N}, 1 \leq k \leq N \right\}.$$

This sampling scheme is of Type A, B or C, depending on whether α_N remains bounded, $\alpha_N \rightarrow \infty$, $\alpha_N = o(N)$ and $N = O(\alpha_N)$, respectively. In the latter case let us assume for the sake of simplicity that $\alpha_N = N$. Using the explicit formula for $E\langle X(u), X(v) \rangle$ we obtain

$$\begin{aligned} E \left\| \frac{1}{N} \sum_{k=1}^N X(s_k) - \mu \right\|^2 &= \frac{1}{N^2} \sum_{k=1}^N \sum_{\ell=1}^N \left(1 - |k - \ell| \frac{\alpha_N}{N} \right) I \left\{ |k - \ell| \leq \frac{N}{\alpha_N} \right\} \\ &= \frac{1}{N} + \frac{2}{N^2} \sum_{h=1}^{N/\alpha_N} \left(1 - \frac{h\alpha_N}{N} \right) (N - h) \\ &\sim \begin{cases} N^{-1} & \text{if } \alpha_N = N; \\ \alpha_N^{-1} & \text{if } \alpha_N \rightarrow \infty, \alpha_N = o(N); \\ 2 \int_0^{\min\{\alpha^{-1}, 1\}} (1 - \alpha|x|)(1 - |x|)dx & \text{if } \alpha_N \rightarrow \alpha. \end{cases} \end{aligned}$$

For Type B and Type C sampling, the optimal bound using Proposition 3.1 is obtained setting $\rho_N = 1$, in which case we have that the r.h.s. in (3.2) is $I_1(\mathfrak{S}_N) = N^{-1}$ if $\alpha_N = N$ and $I_1(\mathfrak{S}_N) = \alpha_N^{-1}$ if $\alpha_N = o(N)$. Under Type A sampling the r.h.s. in (3.2) remains bounded away from zero and the same holds true for the exact quadratic loss.

References

- Carroll, S. S. and Cressie, N. (1996). A comparison of geostatistical methodologies used to estimate snow water equivalent. *Water Resources Bulletin*, **32**, 267–278.
- Carroll, S. S., Day, G. N., Cressie, N. and Carroll, T. R. (1995). Spatial modeling of snow water equivalent using airborne and ground-based snow data. *Environmetrics*, **6**, 127–139.
- Delaigle, A. and Hall, P. (2010). Defining probability density function for a distribution of random functions. *The Annals of Statistics*, **38**, 1171–1193.
- Einmahl, J. H. J and Ruymgaart, F. G. (1987). The order of magnitude of the moments of the modulus of continuity of the multiparameter Poisson and empirical processes. *Journal of Multivariate Analysis*, **21**, 263–273.
- Kaiser, M. S., Daniels, M. J., Furakawa, K. and Dixon, P. (2002). Analysis of particulate matter air pollution using Markov random field models of spatial dependence. *Environmetrics*, **13**, 615–628.
- Lastovicka, J., Akmaev, R. A., Beig, G., Bremer, J., Emmert, J. T., Jacobi, C., Jarvis, J. M., Nedoluha, G., Portnyagin, Yu. I. and Ulich, T. (2008). Emerging pattern of global change in the upper atmosphere and ionosphere. *Annales Geophysicae*, **26**, 1255–1268.
- Maslova, I., Kokoszka, P., Sojka, J. and Zhu, L. (2009). Removal of nonconstant daily variation by means of wavelet and functional data analysis. *Journal of Geophysical Research*, **114**, A03202.
- Maslova, I., Kokoszka, P., Sojka, J. and Zhu, L. (2010a). Estimation of Sq variation by means of multiresolution and principal component analyses. *Journal of Atmospheric and Solar–Terrestrial Physics*, **72**, 625–632.
- Maslova, I., Kokoszka, P., Sojka, J. and Zhu, L. (2010b). Statistical significance testing for the association of magnetometer records at high-, mid- and low latitudes during substorm days. *Planetary and Space Science*, **58**, 437–445.
- Ramsay, J. O. and Silverman, B. W. (2005). *Functional Data Analysis*. Springer.
- Rishbeth, H. (1990). A greenhouse effect in the ionosphere? *Planet. Space Sci.*, **38**, 945–948.
- Roble, R. G. and Dickinson, R. E. (1989). How will changes in carbon dioxide and methane modify the mean structure of the mesosphere and thermosphere? *Geophys. Res. Lett.*, **16**, 1441–1444.
- Stute, W. (1984). The oscillation behavior of empirical processes: the multivariate case. *The Annals of Probability*, **12**, 361–379.
- Ulich, T., Clilverd, M. A. and Rishbeth, H. (2003). Determining long-term change in the ionosphere. *Eos, Transactions American Geophysical Union*, **84**, 581–585.