The Canadian Journal of Statistics
Vol. 36, No. 2, 2008, Pages 1–16
La revue canadienne de statistique

Testing for lack of dependence in the functional linear model

Piotr KOKOSZKA, Inga MASLOVA, Jan SOJKA and Lie ZHU

Key words and phrases: Functional linear model; geomagnetism; hypothesis testing; independence test.

MSC 2000: Primary 62G08; secondary 62F03, 86A25.

Abstract: The authors consider the linear model $Y_n = \Psi X_n + \varepsilon_n$ relating a functional response with explanatory variables. They propose a simple test of the nullity of Ψ based on the principal component decomposition. The limiting distribution of their test statistic is chi-squared, but this distribution is also an excellent approximation in finite samples. The authors illustrate their method using data from terrestrial magnetic observatories.

Un test d'absence de dépendance dans un modèle fonctionnel linéaire

Résumé : Les auteurs s'intéressent au modèle linéaire $Y_n = \Psi X_n + \varepsilon_n$ liant une variable réponse fonctionnelle à des variables explicatives. Ils proposent un test simple de nullité de Ψ fondé sur la décomposition en composantes principales. La loi limite de leur statistique est une khi-deux, mais cette loi fournit aussi une excellente approximation à taille finie. Les auteurs illustrent leur méthode au moyen de données provenant d'observatoires du champ magnétique terrestre.

1. INTRODUCTION

The last two decades have seen the emergence of new technology allowing the collection and storage of data consisting of finely sampled records over some natural repeated time or space interval. Examples include minute by minute values of a speculative asset, meteorogical and pollution monitoring data, seismic data and a plethora of examples in all fields of science and engineering. The common feature of such data is that a single observation is a curve, rather than a point or a vector. Functional data analysis (FDA) is a rapidly growing body of statistical tools designed to analyze such data.

One of the most popular models of functional data analysis is the functional linear model; see Ramsay & Silverman (2005, chs 12–17). A brief review is presented in Chiou, Müller & Wang (2004). This model is defined by the equation

$$Y_n = \Psi X_n + \varepsilon_n, \quad n = 1, \ldots, N. \tag{1}$$

The curves Y_n and X_n, as well as the unobservable functional errors, ε_n, are assumed to lie in the Hilbert space $L^2[0,1)$. The operator $\Psi\colon L^2 \to L^2$ is a bounded linear operator. Detailed assumptions are stated in Section 2. Typically Ψ is assumed to be a Hilbert–Schmidt integral operator, i.e., it can be represented by a kernel function $\psi(t,s)$ which is square integrable over $[0,1) \times [0,1)$. In that case equation (1) becomes

$$Y_n(t) = \int_0^1 \psi(t,s) X_n(s)\, ds + \varepsilon_n(t), \quad n = 1, \ldots, N.$$

Testing the null hypothesis of no effect, i.e., testing whether Ψ is zero is often a question of practical interest, which exhibits new features in the functional setting due to the fact that the data are infinitely dimensional and every dimension reduction technique restricts the domain of Ψ, and so leads to a loss of information about Ψ and complicates invertibility arguments. These issues are addressed in different contexts in Cuevas, Febrero & Fraiman (2002) and Cardot, Ferraty,

Mas & Sarda (2003). The data that motivated our research requires that model (1) be fully functional with random explanatory variables, i.e., the Y_n, X_n, ε_n are all random curves. The testing procedure we propose is similar to that developed in Cardot, Ferraty, Mas & Sarda (2003) who consider scalar responses Y_n. It turns out that the more symmetric fully functional formulation actually leads to a somewhat simpler, and more symmetric in Y_n and X_n, test statistic which does not require additional estimation of the noise variance and can be readily computed using the principal components decompositions of the the Y_n and the X_n. Our test statistic has limiting chi-square distribution which is a good approximation for sample sizes around 50. Our asymptotic argument carefully distinguishes between population and estimated functional principal components, a point recently emphasized by Hall & Hosseini-Nasab (2006, 2007). Other recent contributions dealing with the functional linear model are Cai & Hall (2006), Müller & Stadtmüller (2005), and Yao, Müller & Wang (2005), among others.

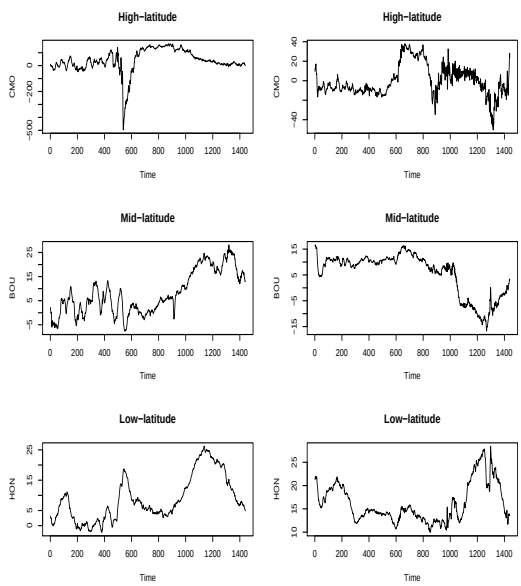

FIGURE 1: Horizontal intensities of the magnetic field measured at high-, mid- and low-latitude stations during a substorm (left column) and a quiet day (right column). Note the different vertical scales for high-latitude records.

The research presented in this paper is to a large extent motivated by our work with magnetometer data. Figure 1 shows examples of magnetometer records. Technical details are explained in Section 5. Here we merely note that each panel shows one curve which we treat as a single functional observation.

The paper is organized as follows. After introducing the notation and the assumptions in Section 2, we present the test procedure and establish its asymptotic validity in Section 3. The finite-sample performance is examined in Section 4, whereas the application to magnetometer data is presented in Section 5. The proofs of the asymptotic results of Section 3 are developed in the Appendix.

2. NOTATION AND ASSUMPTIONS

We assume that the response variables Y_n, the explanatory variables X_n and the errors ε_n are random elements of the Hilbert space $L^2[0,1)$. Recall that the expectation of a random element Z, say, of $L^2[0,1)$ is a function in $L^2[0,1)$ defined by $(\mathrm{E}\,Z)(t) = \mathrm{E}\,[Z(t)]$, $t \in [0,1)$. The inner product of $x, y \in L^2[0,1)$ is defined by $\langle x, y\rangle = \int_0^1 x(t)y(t)\,dt$, and the norm by $\|x\|^2 = \int_0^1 x^2(t)\,dt$. If Z is a random element of $L^2[0,1)$, then $\|Z\|$ is a random variable.

The theory developed below is valid under the following assumption.

ASSUMPTION 1. The triples $(Y_n, X_n, \varepsilon_n)$ form a sequence of independent identically distributed random elements such that ε_n is independent of (Y_n, X_n) and

$$\mathrm{E}\, X_n = 0 \quad \text{and} \quad \mathrm{E}\, \varepsilon_n = 0; \tag{2}$$

$$\mathrm{E}\, \|X_n\|^4 < \infty \quad \text{and} \quad \mathrm{E}\, \|\varepsilon_n\|^4 < \infty. \tag{3}$$

Our next assumption requires that the empirical eigenelements be close to the population eigenelements of the covariance operators of the X_n and the Y_n. This point is often overlooked in empirical work, but assumptions of this type are needed to develop a rigorous asymptotic theory, see Bosq (2000, ch. 4) and Hall & Hosseini-Nasab (2006).

Introduce the operators:

$$\Gamma x = \mathrm{E}\, [\langle X_1, x\rangle X_1], \quad \Lambda x = \mathrm{E}\, [\langle Y_1, x\rangle Y_1], \quad \Delta x = \mathrm{E}\, [\langle X_1, x\rangle Y_1].$$

Denote their empirical counterparts by Γ_N, Λ_N, Δ_N, e.g.,

$$\Gamma_N x = \frac{1}{N} \sum_{n=1}^{N} \langle X_n, x\rangle X_n.$$

Define the eigenelements by

$$\Gamma v_k = \gamma_k v_k, \quad \Lambda u_j = \lambda_j u_j.$$

Empirical eigenelements are defined correspondingly and denoted by $(\hat{\gamma}_k, \hat{v}_k)$, $(\hat{\lambda}_j, \hat{u}_j)$.

ASSUMPTION 2. The eigenvalues of the operators Γ and Λ satisfy, for some $p > 0$ and $q > 0$,

$$\gamma_1 > \cdots > \gamma_{p+1}, \quad \lambda_1 > \cdots > \lambda_{q+1}. \tag{4}$$

Assumption 2 implies that the eigenspaces corresponding to the first largest p (respectively q) eigenvalues are one dimensional. Therefore, the corresponding normalized principal components are well defined (up to the sign) and orthogonal. No formal test is currently available to verify Assumption 2, but it is very natural as the estimated eigenvalues are always positive and distinct in applications.

In the proofs, we will often use the relations

$$\limsup_{N\to\infty} N\, \mathrm{E}\, \|v_k - \hat{v}_k\|^2 < \infty, \quad \limsup_{N\to\infty} N\, \mathrm{E}\, \|u_j - \hat{u}_j\|^2 < \infty; \tag{5}$$

$$\limsup_{N\to\infty} N\, \mathrm{E}\, \left[|\gamma_k - \hat{\gamma}_k|^2\right] < \infty, \quad \limsup_{N\to\infty} N\, \mathrm{E}\, \left[|\lambda_j - \hat{\lambda}_j|^2\right] < \infty, \tag{6}$$

which hold for each $k \le p$ and $j \le q$ under Assumptions 1 and 2, see Bosq (2000, ch. 4).

3. TEST PROCEDURE AND ASYMPTOTIC RESULTS

Assuming model (1), we wish to test

$$\mathcal{H}_0 : \ \Psi = 0 \quad \text{versus} \quad \mathcal{H}_A : \ \Psi \neq 0.$$

We thus test the null hypothesis that the curves X_n have no effect on the curves Y_n. This is analogous to testing in the scalar linear model $y_i = \beta_0 + \beta_1 x_{i,1} + \cdots + \beta_{p-1} x_{i,p-1} + \varepsilon_i$ whether the regression on the regressors $X_1, \ldots, X_{p-1}$ is significant. In this standard setting, the F-test is used. In the particular case of straight line regression, $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, the F-test is equivalent to the usual t-test for non-zero slope, see, e.g., Seber & Lee (2003, ch. 4). In

our functional setting the slope corresponds to a linear operator which transforms functions into functions. Just as in the case of straight line regression, the nullity of Ψ does not mean that there is no dependence between the curves X_n and Y_n, but that if there is a dependence, it cannot be described by a functional linear model.

The testing procedure involves restrictions of the operators defined in Section 2 to certain finite dimensional subspaces. This is a dimension reduction procedure which necessarily involves some loss of information about the action of Ψ. The subspace $\mathcal{V}_p = \mathrm{sp}\{v_1, \ldots, v_p\}$, which is isomorphic to R^p, contains the best approximations to the X_n which are linear combinations of the first p principal components, see Ramsay & Silverman (2005,§ 8.2.3). Similarly, $\mathcal{U}_q = \mathrm{sp}\{u_1, \ldots, u_q\}$ is a good approximation to sp$\{Y_1, \ldots, Y_n\}$.

Since, by (1), $\Delta = \Psi\Gamma$, we have, for $k \leq p$,

$$\Psi v_k = \gamma_k^{-1} \Delta v_k. \tag{7}$$

Thus, by Assumption 2, Ψ vanishes on $\mathrm{sp}\{v_1, \ldots, v_p\}$ if and only if $\Delta v_k = 0$ for each $k = 1, \ldots, p$. Observe that

$$\Delta v_k \approx \Delta_N v_k = \frac{1}{N} \sum_{n=1}^{N} \langle X_n, v_k \rangle Y_n.$$

As $\mathrm{sp}\{Y_1, \ldots, Y_N\}$ is well approximated by $\mathcal{U}_q$, we can develop a test by checking whether

$$\langle \Delta_N v_k, u_j \rangle = 0, \quad k = 1, \ldots, p, \; j = 1, \ldots, q. \tag{8}$$

If such a test accepts $\mathcal{H}_0$, this means that for every $x \in \mathcal{V}_p$, Ψx is not in $\mathcal{U}_q$. Intuitively, we see that up to a small error arising from the approximations by the principal components and a random error, no function Y_n, $n = 1, \ldots, N$, can be expressed as a linear combination of functions X_n, $n = 1, \ldots, N$.

Theorem 1 shows that the test statistic

$$\widehat{T}_N(p, q) = N \sum_{k=1}^{p} \sum_{j=1}^{q} \hat{\gamma}_k^{-1} \hat{\lambda}_j^{-1} \langle \Delta_N \hat{v}_k, \hat{u}_j \rangle^2 \tag{9}$$

has a parameter-free asymptotic distribution.

THEOREM 1. *Under $\mathcal{H}_0$ and the assumptions of* Section 2,

$$\widehat{T}_N(p, q) \xrightarrow{d} \chi^2_{pq}.$$

If $\mathcal{H}_0$ fails, then $\Psi v_k \neq 0$ for some $k \geq 1$. If we impose conditions only on the first p largest eigenvalues, the test will be consistent only if Ψ does not vanish on one of the v_k, $k = 1, \ldots, p$. The test has no power if Ψ does not vanish on the orthogonal complement of sp$\{v_1, \ldots, v_p\}$. Further, to ensure consistency, one of the v_k, $k = 1, \ldots, p$ must be mapped into sp$\{u_1, \ldots, u_q\}$. These restrictions are intuitively appealing because we want to test whether the main sources of the variability of the responses Y can be explained by the main sources of the variability of the explanatory variables X.

The following theorem formalizes these ideas and establishes the consistency of the test.

THEOREM 2. *If the assumptions of* Section 2 *hold, and $\langle \Psi v_k, u_j \rangle \neq 0$ for some $k \leq p$ and $j \leq q$, then $\widehat{T}_N(p, q) \xrightarrow{P} \infty$, as $N \to \infty$.*

In a linear regression setting, it is often of interest to test whether specific covariates have no effect on the responses. In our setting, we could ask whether specific principal components v_k have no effect. It is easy to see from the proof of Theorem 1 (see Lemma 1 in particular) that if we want to test whether principal components $v_{i(1)}, \ldots, v_{i(p')}$ have no effect, we must modify the statistic (9) by including only these components. The limit χ^2 distribution will then have $p'q$ degrees of freedom. A further obvious modification can be made if we want to check whether there is an effect in the subspace spanned by some principal components of the responses Y_k. Modifications of this type are useful if some principal components have obvious physical interpretations. This is sometimes the case in space physics applications, see W-Y. Xu & Y. Kamide (2004), but when the X_n are high-latitude records the v_k cannot, at this point, be readily interpreted. See Section 5.

Summary of the testing procedure.

1. Check the linearity assumption using FPC score predictor-response plots, see Section 5.

2. Select the number of important PC's?, p and q using both the scree test and CPV, see Section 5.

3. Compute the test statistics $\widehat{T}_N(p,q)$ (9). Note that

$$\langle \triangle_N \hat{v}_k, \hat{u}_j \rangle = \left\langle \frac{1}{N} \sum_{n=1}^{N} \langle X_n, \hat{v}_k \rangle Y_n, \hat{u}_j \right\rangle = \frac{1}{N} \sum_{n=1}^{N} \langle X_n, \hat{v}_k \rangle \langle Y_n, \hat{u}_j \rangle,$$

   where $\langle X_n, \hat{v}_k \rangle$ is the kth score of the X_n, and $\langle Y_n, \hat{u}_j \rangle$ is jth score of the Y_n. These scores and the eigenvalues $\hat{\gamma}_k$ and $\hat{\lambda}_j$ are output of functions available in the R package fda.

4. If $\widehat{T}_N(p,q) > \chi^2_{pq}(\alpha)$, reject the null hypothesis of no linear effect. The critical value $\chi^2_{pq}(\alpha)$ is the $(1-\alpha)$th quantile of the chi-squared distribution with pq degrees of freedom.

4. A SMALL SIMULATION STUDY

In this section, we present the results of a small simulation study intended to evaluate the empirical size and power of the test in standard Gaussian settings.

We used $R = 1000$ replications of samples of processes ε_n, X_n and Y_n, $n = 1, \ldots, N$. In order to evaluate the empirical size, we generated samples of pairs (ε_n, Y_n) with independent components. To find the empirical power, we generated samples of pairs (ε_n, X_n) with independent components, and calculated Y_n according to (1). As ε_n, X_n and Y_n, we used Brownian bridge and motion processes in various combinations. The computations were performed using the R package fda. We used both Fourier and splines bases.

Since the Brownian bridge and motion have very regular Karhunen–Loève decompositions, see, e.g., Bosq (2000, p. 26), it is not surprising that the size and power of the test do not depend appreciably on p and q. Figures 2 and 3 illustrate this point. The horizontal axes represent various combinations of p and q; 1 stands for $p = 1$ and $q = 1$; 2 for $p = 1$, $q = 2$; 3 for $p = 1$, $q = 3$, etc. All combinations of $p \leq 4$, $q \leq 4$ were considered in the size study and $p \leq 6$, $q \leq 6$ in the power study. The results for Brownian bridges and motions and Fourier and spline bases are practically the same. For this reason, we present the results only in cases when all processes are Brownian bridges, and the analysis was performed with the Fourier basis.

Naturally, the bigger the sample size the closer the empirical size of the test is to the nominal size. Nevertheless, there is little or no improvement in the size of the test starting from $N = 40-80$; these values can therefore be considered sufficient to obtain reasonable size; with $N = 40$ the test being slightly conservative.

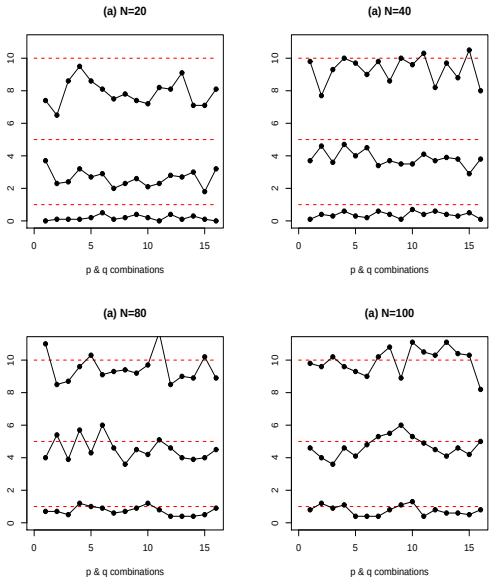

FIGURE 2: Empirical size of the test for $\alpha = 1\%, 5\%, 10\%$ (indicated by dotted lines) for different combinations of p and q. Here ε_n and Y_n, $n = 1, \ldots, N$ are two independent Brownian Bridges.

To evaluate the empirical power, we used the Gaussian kernel

$$\psi(s,t) = C \exp\,(t^2 + s^2)/2, \quad t \in [0,1],\ s \in [0,1] \tag{10}$$

with constants C such that $\|\Psi\| < 1$, i.e., $|C| < 1$. Panels (a) and (b) of Figure 3 present power when the dependence between X_n and Y_n is quite strong, $\|\Psi\| = 0.75$. For $N = 80$, the power is practically 100% if $\|\Psi\| = 0.75$. The right column of Figure 3 shows the power of the test when $\|\Psi\| = 0.5$. In this case power increases more slowly with N.

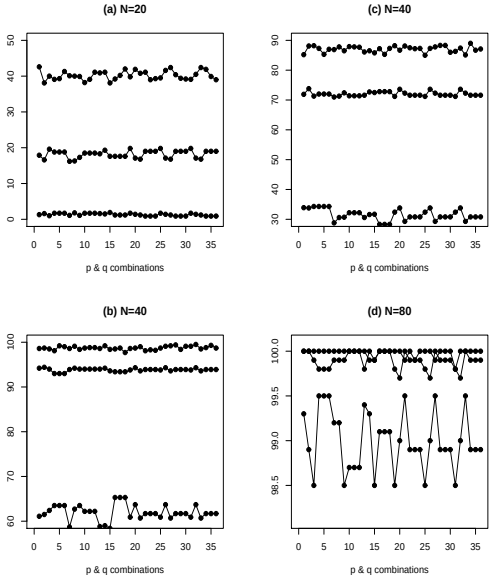

FIGURE 3: Empirical power of the test for different combinations of principal components and different sample sizes N. Here X_n and ε_n are Brownian Bridges. In panels (a), (b) $\|\Psi\| = 0.75$; in panels (c), (d) $\|\Psi\| = 0.5$.

Even though this paper is concerned with testing for no effect in the linear fully functional model, it might be interesting to see what happens if the responses depend or the regressors in a nonlinear manner. Let X_n be independent Brownian motions, and ε_n independent Brownian bridges (independent of the X_n). We computed the empirical power of the test for the following models:

$$Y_n(t) = H_2(X_n(t)) + \varepsilon_n(t), \tag{11}$$

where $H_2(x) = x^2 - 1$, and

$$Y_n(t) = X_n(t)\varepsilon_n(t). \tag{12}$$

The function H_2 in (11) is the Hermite polynomial of rank 2; in model (12) the errors are multiplicative.

For $N = 40$, in case of model (11), the empirical power for various principal component combinations is around 53% for the significance level $\alpha = 10\%$, 30% for $\alpha = 5\%$, and 9% for $\alpha = 1\%$. For the multiplicative model (12) the power is about 38% for $\alpha = 10\%$, 24% for $\alpha = 5\%$, and 6% for $\alpha = 1\%$. Just as in the case of the usual linear models, the test can detect some nonlinear dependence, but not reliably.

5. APPLICATION TO MAGNETOMETER DATA

About a hundred terrestrial geomagnetic observatories form a network, INTERMAGNET, designed to monitor and understand the behavior of electrical currents flowing in the magnetosphere and ionosphere (M-I). Interestingly, Gauss was one of the leaders of the early nineteenth century effort to establish such a network, and he pioneered the statistical analysis of the resulting measurements, see Kivelson & Russell (1997, ch. 1). Modern digital magnetometers record three components of the magnetic field in five second resolution, but the data made available by INTERMAGNET (http://www.intermagnet.org) consist of one minute averages (1440 data points per day per component per observatory). Figure 1 shows examples of magnetometer records. We work with the Horizontal (H) component of the magnetic field. This is the component lying in the Earth's tangent plane and pointing toward the magnetic North. It most directly reflects the variation of the M-I currents we wish to study. The M-I currents form a complex interactive system which at present is only partially understood, see Kamide et al. (2003; 1998 in references???). The magnetometer records contain intertwined signatures of many currents, and an effort has been under way to deconvolute the signatures of various currents. So far this has been done by preprocessing records from every individual station, and then combining the filtered signals from stations at the same magnetic latitude (e.g. equatorial stations, or auroral stations). For a recent example of such an approach, see Jach, Kokoszka, Sojka & Zhu (2007). It is however believed, see, e.g., Rostoker (2000), that the auroral currents may have an impact, perhaps indirect, on the equatorial and mid-latitude currents. The present paper has been motivated by this problem and shows that this is indeed the case using the proposed test of significance. Our goal in this section is to illustrate the methodology using an interesting data set rather than to present a comprehensive case study. A detailed analysis with a deeper discussion of physical insights is presented in Kokoszka, Maslova, Sojka & Zhu (2007).

The data consist of minute-by-minute records of the horizontal intensity of the magnetic field measured in 2001 at observatories listed in Table 1. The observatories in each of the four groups are roughly aligned along the same magnetic longitude. The functional observations consist of daily (in UT) curves (1440 records per curve). Examples of such curves are shown in Figure 1.

The question of interest is whether the auroral geomagnetic activity reflected in the high-latitude curves has an effect on the processes in the equatorial belt reflected by the mid- and high-latitude curves. This question is of particular interest for days during which a high-latitude activity known as a substorm occurs. Its most spectacular manifestation are the Northern Lights caused by high-energy electrojets flowing for a few hours in the auroral belt. The top left panel of Figure 1 shows a signature of a substorm. It is believed that there is energy transfer between the

auroral electrojets and lower latitude currents, but the direct physical mechanisms which might be responsible for this interaction are a matter of debate.

TABLE 1: Geomagnetic observatories used in this study.

Latitude	I	II	III	IV
High	College (CMO)	–	–	–
Mid	Boulder (BOU)	Fredericksburg (FRD)	Tihany (THY)	Memambetsu (MMB)
Low	Honolulu (HON)	San Juan (SJG)	Hermanus (HER)	Kakioka (KAK)

The question can be cast into the setting of the functional linear model (1) in which the X_n are centered high-latitude records and Y_n are centered mid- or low-latitude records. This postulates an approximate statistical model for the data and allows us to the the null hypothesis $\Psi = 0$. If the null hypothesis is true, we conclude that the high-latitude curves X_n have no linear effect on the lower latitude curves. If the null hypothesis is rejected, this indicates the existence of an effect, which can be approximately linear (functionally. Other modeling settings are conceivable; for instance, an adaptation of a nonparametric approach advocated by Ferraty & Vieu (2006) might be appealing and could provide additional insights.

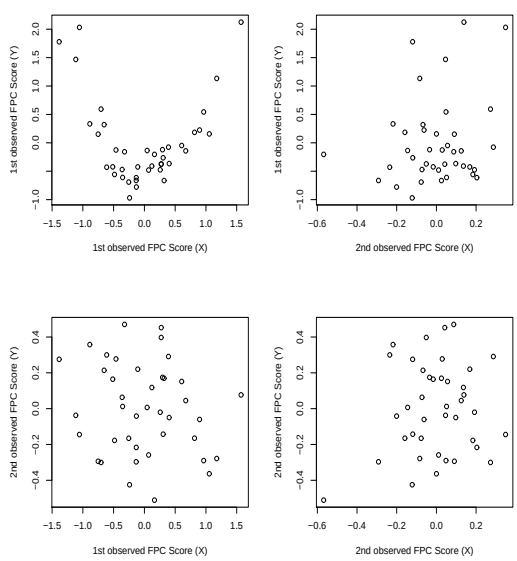

FIGURE 4: Functional predictor-response plots of functional principal component scores of response functions versus functional principal component scores of predictor functions for $Y_n(t) = H_2(X_n(t)) + \varepsilon_n(t)$, where $H_2(x) = x^2 - 1$, $n = 1, \ldots, 40$.

In the analysis below, we use $N = 41$ days in January–August 2001, a period which contained a medium strength substorm. Thus X_n is the curve on the nth day with a medium strength substorm and Y_n is the curve on the same (UT) day measured at mid- or low-latitude station. In addition, we consider mid- and low-latitude curves 1, 2 and 3 days after the day with a substorm. This is intended to check how long the effects of a substorm persist. The independence of the cases (X_n, Y_n) can be assumed to hold approximately because the substorm days are typically separated by quiet days during which the M-I system resets itself. The independence of the X_n is also confirmed by the application of the test developed by Gabrys & Kokoszka (2007). The same holds true for the Y_n.

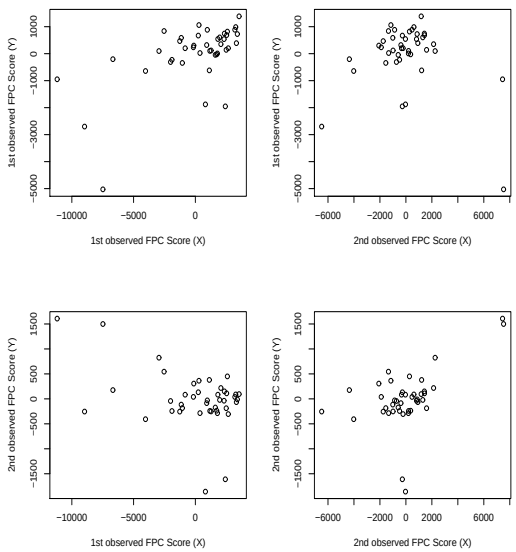

FIGURE 5: Functional predictor-response plots of functional principal component scores of response functions versus those of predictor functions for magnetometer data (CMO vs THY0).

As mentioned in Section 4, to ensure that the test gives reliable results, the linearity assumption must be checked. For this purpose, visual techniques introduced by Chiou & Müller (2007) can be used. Functional principal component (FPC) scores are used to check the linearity assumption. In case of linear dependence, the FPC score plots are roughly football-shaped. When the dependence is not linear, these plots exhibit different patterns. For example, for model (11) introduced in Section 4, the scatterplot of the first FPC clearly shows a quadratic trend, see Figure 4.

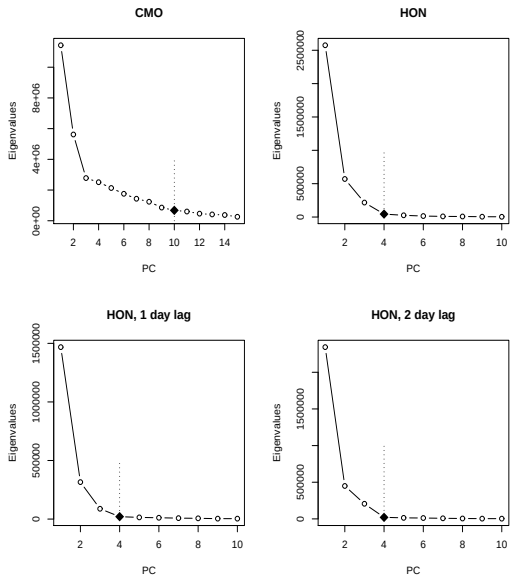

FIGURE 6: Eigenvalues for different principal components of the substorm days that occurred from March until May, 2001, from College (CMO), Honolulu (HON) stations.

Figure 5 is an example of the relationship between the response and the predictor FPC scores for magnetometer data. We used CMO records as X, and THY with no lag as Y. These scatterplots indicate linear relationship with some outliers. Since we do not require Gaussianity, only finite fourth moment, these outliers need not invalidate our conclusions. In case of other pairs of functional data, the FPC score plots look similar. We conclude that a linear model is approximately appropriate for our application.

To apply the test, we need to decide which values of p and q should be used. We propose to use all values up to some meaningful upper bounds, and to look at the pattern of rejections and acceptances as a function of p and q. One of the ways to pick the most important principal components is to use the scree test, which is a graphical method first proposed by Cattell (1966). To apply the scree method one plots the successive eigenvalues against the corresponding principal components (see Figure 6). The method suggests finding the place where the smooth decrease of eigenvalues appears to level off. To the right of this point one finds only factorial scree ("scree" is a geological term used to refer to the debris which collects on the lower part of a rocky slope).

Another way to pick the unknown number of principal components from the data is to compute the cumulative percentage of total variance (CPV), as in multivariate principal component analysis. So, the CPV explained by the first p functional principal components is

$$\mathrm{CPV}(p) = \frac{\sum_{k=1}^{p} \lambda_k}{\sum_{k=1}^{\infty} \lambda_k}.$$

Table 2 gives the upper limits on p and q together with the CPV explained by these components. Visual examination of the principal components beyond the upper bound confirms that they resemble random noise.

In most cases, there is a clear rejection or acceptance for almost all combinations of p and q, for all small values which correspond to the most important principal components. For such cases, we can with reasonable confidence reject ("1") or fail to reject ("0"). However, there are some cases where it is not clear what conclusion to draw. We denote them by "1?" (inclined toward rejecting $\mathcal{H}_0$), "0?" (inclined toward accepting $\mathcal{H}_0$), "1?0?" (inconclusive). Figure 7 gives examples of such cases.

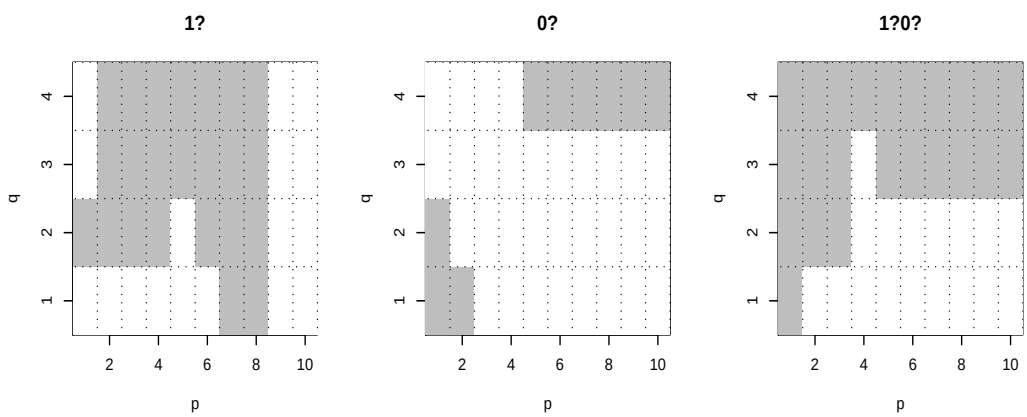

FIGURE 7: Examples of rejection/acceptance plots at 5% level which are difficult to interpret. Grey area: reject $\mathcal{H}_0$, white: fail to reject $\mathcal{H}_0$.

TABLE 2: Number of principal components retained by the scree test, and percentage of total variability explained, during medium strength substorm days that occurred from January until August, 2001.

Stations	PC	%
College (CMO)	10	81.97
Boulder (BOU)	3	86.99
Boulder (BOU) one-day lag	4	81.68
Boulder (BOU) two-day lag	2	91.18
Boulder (BOU) three-day lag	3	95.15
Honolulu (HON)	2	93.85
Honolulu (HON) one-day lag	4	93.89
Honolulu (HON) two-day lag	3	98.16
Honolulu (HON) three-day lag	2	98.80
Fredericksburg (FRD)	4	92.83
Fredericksburg (FRD) one-day lag	4	89.52
Fredericksburg (FRD) two-day lag	3	94.35
Fredericksburg (FRD) three-day lag	4	96.77
San Juan (SJG)	2	90.86
San Juan (SJG) one-day lag	3	86.40
San Juan (SJG) two-day lag	2	94.32
San Juan (SJG) three-day lag	3	96.63
Tihany (THY)	3	89.57
Tihany (THY) one-day lag	4	83.75
Tihany (THY) two-day lag	2	89.64
Tihany (THY) three-day lag	3	94.33
Hermanus (HER)	2	90.91
Hermanus (HER) one-day lag	3	89.64
Hermanus (HER) two-day lag	2	93.84
Hermanus (HER) three-day lag	3	96.53
Memambetsu (MMB)	2	89.99
Memambetsu (MMB) one-day lag	3	85.36
Memambetsu (MMB) two-day lag	3	95.64
Memambetsu (MMB) three-day lag	3	97.04
Kakioka (KAK)	2	92.89
Kakioka (KAK) one-day lag	2	92.89
Kakioka (KAK) two-day lag	3	97.08
Kakioka (KAK) three-day lag	3	98.04

Table 3 presents the results of our analysis. It shows that the effect of a substorm persists for about one day. Beyond that time, the magnetometer data at mid and low latitudes are not linearly dependent (functionally) on the high latitude records. We note however that a slightly more complex picture emerges for different seasons in 2001 and for special subcategories of substorms. These issues are discussed in Kokoszka, Maslova, Sojka & Zhu (2007).

TABLE 3: Results of the test for medium strength substorm days from January to August 2001.

CMO							
BOU0	BOU1	BOU2	BOU3	HON0	HON1	HON2	HON3
1?	0	0	0?	1?	1?0?	0	0?
FRD0	FRD1	FRD2	FRD3	SJG0	SJG1	SJG2	SJG3
1?	0	0	0?	1?	0	0	0?
THY0	THY1	THY2	THY3	HER0	HER1	HER2	HER3
1?	0?	0	0?	0?	0?	0	0?
MMB0	MMB1	MMB2	MMB3	KAK0	KAK1	KAK2	KAK3
0?	0	0	0?	1?	1?	0	0?

APPENDIX

Proof of Theorem 1. Theorem 1 follows from Corollary 1, which is arrived at through a series of lemmas. Lemma 1 shows that the χ^2 limit holds for the population eigenelements. The remaining lemmas show that the differences between the empirical and population eigenelements have asymptotically negligible effect.

LEMMA 1. *Under $\mathcal{H}_0$ and the assumptions of* Section 2, *for each $j \le q$, $k \le p$,*

$$\sqrt{N}\,\langle \Delta_N v_k, u_j\rangle \xrightarrow{d} \eta_{kj}\sqrt{\gamma_k\lambda_j}\,, \tag{13}$$

with $\eta_{kj} \sim \mathsf{N}(0,1)$. Moreover, η_{kj} and $\eta_{k'j'}$ are independent if $(k,j) \neq (k',j')$.

Proof of Lemma 1. Under $\mathcal{H}_0$,

$$\sqrt{N}\,\langle \Delta_N v_k, u_j\rangle = N^{-1/2}\sum_{n=1}^{N}\langle X_n, v_k\rangle\langle \varepsilon_n, u_j\rangle.$$

The summands have mean zero and variance $\gamma_k\lambda_j$, so (13) follows.

To verify that η_{kj} and $\eta_{k'j'}$ are independent if $(k,j) \neq (k',j')$, it suffices to show that $\sqrt{N}\,\langle \Delta_N v_k, u_j\rangle$ and $\sqrt{N}\,\langle \Delta_N v_{k'}, u_{j'}\rangle$ are uncorrelated. Observe that

$$\begin{aligned}
&\mathrm{E}\left[\sqrt{N}\,\langle \Delta_N v_k, u_j\rangle, \sqrt{N}\,\langle \Delta_N v_{k'}, u_{j'}\rangle\right] \\
&= \frac{1}{N}\,\mathrm{E}\left[\sum_{n=1}^{N}\langle X_n, v_k\rangle\langle \varepsilon_n, u_j\rangle \sum_{n'=1}^{N}\langle X_{n'}, v_{k'}\rangle\langle \varepsilon_{n'}, u_{j'}\rangle\right] \\
&= \frac{1}{N}\sum_{n,n'=1}^{N}\mathrm{E}\left[\langle X_n, v_k\rangle\langle X_{n'}, v_{k'}\rangle\right]\mathrm{E}\left[\langle \varepsilon_n, u_j\rangle\langle \varepsilon_{n'}, u_{j'}\rangle\right] \\
&= \frac{1}{N}\sum_{n=1}^{N}\mathrm{E}\left[\langle X_n, v_k\rangle\langle X_n, v_{k'}\rangle\right]\mathrm{E}\left[\langle \varepsilon_n, u_j\rangle\langle \varepsilon_n, u_{j'}\rangle\right] \\
&= \langle \Gamma v_k, v_{k'}\rangle\langle \Lambda u_j, u_{j'}\rangle \;=\; \lambda_k\delta_{kk'}\lambda_j\delta_{jj'}.
\end{aligned}$$

Recall that the Hilbert–Schmidt norm of a Hilbert–Schmidt operator S is defined by

$$\|S\|_{\mathcal{S}}^2 = \sum_{j=1}^{\infty}\|Se_j\|^2,$$

where $\{e_1, e_2, \ldots\}$ is any orthonormal basis. Recall also that the Hilbert–Schmidt norm dominates the operator norm: $\|S\| \leq \|S\|_{\mathcal{S}}$.

LEMMA 2. *Under $\mathcal{H}_0$ and the assumptions of* Section 2,

$$\mathrm{E}\ \|\Delta_N\|_{\mathcal{S}}^2 = N^{-1}\mathrm{E}\ \|X_1\|^2\mathrm{E}\,\|\varepsilon_1\|^2.$$

Proof of Lemma 2. Observe that

$$\|\Delta_N e_j\|^2 = N^{-2}\sum_{n,n'=1}^{N}\langle X_n, e_j\rangle\langle X_{n'}, e_j\rangle\langle Y_n, Y_{n'}\rangle.$$

Therefore, under $\mathcal{H}_0$,

$$\begin{aligned}
\mathrm{E}\ \|\Delta_N\|_{\mathcal{S}}^2 &= N^{-2}\sum_{j=1}^{\infty}\sum_{n,n'=1}^{N}\mathrm{E}\left[\langle X_n, e_j\rangle\langle X_{n'}, e_j\rangle\langle \varepsilon_n, \varepsilon_{n'}\rangle\right] \\
&= N^{-2}\sum_{j=1}^{\infty}\sum_{n=1}^{N}\mathrm{E}\ \langle X_n, e_j\rangle^2\mathrm{E}\ \|\varepsilon_n\|^2 \\
&= N^{-1}\mathrm{E}\ \|\varepsilon_1\|^2\sum_{j=1}^{\infty}\langle X_1, e_j\rangle^2 \ = \ N^{-1}\mathrm{E}\ \|\varepsilon_1\|^2\mathrm{E}\ \|X_1\|^2.
\end{aligned}$$

The following elementary lemma is stated for ease of reference.

LEMMA 3. *Suppose $\{U_N\}$ and $\{V_N\}$ are random sequences in a Hilbert space such that $\|U_N\| \xrightarrow{P} 0$ and $\|V_N\| = O_P(1)$, i.e., $\lim_{C\to\infty}\limsup_{N\to\infty} P(\|V_N\| > C) = 0$. Then*

$$\langle U_N, V_N\rangle \xrightarrow{P} 0.$$

Proof of Lemma 3. The lemma follows from the corresponding property of real random sequences and the inequality $|\langle U_N, V_N\rangle| \leq \|U_N\|\|V_N\|$.

LEMMA 4. *Under $\mathcal{H}_0$ and the assumptions of* Section 2, *for each $j \leq q, k \leq p$,*

$$\sqrt{N}\,\langle \Delta_N\hat{v}_k, \hat{u}_j\rangle \xrightarrow{d} \eta_{kj}\sqrt{\gamma_k\lambda_j}\,, \tag{14}$$

with η_{kj} equal to those in Lemma 1.

Proof of Lemma 4. It suffices to verify that

$$\sqrt{N}\,\langle \Delta_N\hat{v}_k, \hat{u}_j\rangle - \sqrt{N}\,\langle \Delta_N v_k, u_j\rangle \xrightarrow{P} 0. \tag{15}$$

Relation (15) will follow from

$$\sqrt{N}\,\langle \Delta_N v_k, \hat{u}_j - u_j\rangle \xrightarrow{P} 0 \tag{16}$$

and

$$\sqrt{N}\,\langle \Delta_N(\hat{v}_k - v_k), \hat{u}_j\rangle \xrightarrow{P} 0. \tag{17}$$

To verify (16), note that by (5), $\sqrt{N}\,(\hat{u}_j - u_j) = O_P(1)$, and by Lemma 2, $\mathrm{E}\,\|\Delta_N v_k\| \leq \mathrm{E}\,\|\Delta_N\|_{\mathcal{S}} = O(N^{-1/2})$. Thus (16) follows from Lemma 3.

To use the same argument for (17) (with (5)), we note that

$$\sqrt{N}\,\langle \Delta_N(\hat{v}_k - v_k), \hat{u}_j\rangle = \sqrt{N}\,\langle \hat{v}_k - v_k, \tilde{\Delta}_N \hat{u}_j\rangle,$$

where $\tilde{\Delta}_N x = N^{-1}\sum_{n=1}^N \langle Y_n, x\rangle X_n$. Lemma 2 shows that under $\mathcal{H}_0$, $\mathrm{E}\,\|\tilde{\Delta}_N\|_{\mathcal{S}} = \mathrm{E}\,\|\Delta_N\|_{\mathcal{S}}$.

By (6), $\hat{\gamma}_k \xrightarrow{P} \gamma_k$ and $\hat{\lambda}_j \xrightarrow{P} \lambda_j$, so we obtain

COROLLARY 1. *Under $\mathcal{H}_0$ and the assumptions of* Section 2*, for each $j \leq q, k \leq p$,*

$$\sqrt{N}\,\hat{\gamma}_k^{-1/2}\hat{\lambda}_j^{-1/2}\langle \Delta_N \hat{v}_k, \hat{u}_j\rangle \xrightarrow{d} \eta_{kj}, \tag{18}$$

with η_{kj} equal to those in Lemma 1.

Proof of Theorem 2. Denote

$$\widehat{S}_N(p,q) = \sum_{k=1}^p \sum_{j=1}^q \hat{\gamma}_k^{-1}\hat{\lambda}_j^{-1}\langle \Delta_N \hat{v}_k, \hat{u}_j\rangle^2.$$

By Lemma 7 and (6), $\widehat{S}_N(p,q) \xrightarrow{P} S(p,q) > 0$. Hence $\widehat{T}_N(p,q) = N\widehat{S}_N(p,q) \xrightarrow{P} \infty$.

To establish Lemma 7, it is convenient to split the argument into two simple lemmas: Lemma 5 and Lemma 6.

LEMMA 5. *If Y_n, $n \geq 1$, are identically distributed, then* $\mathrm{E}\,\|\Delta_N\| \leq \mathrm{E}\,\|Y_1\|^2$.

Proof of Lemma 5. For arbitrary $u \in L^2$ with $\|u\| \leq 1$,

$$\|\Delta_N u\| \leq N^{-1}\sum_{n=1}^N |\langle Y_n, u\rangle|\,\|Y_n\| \leq N^{-1}\sum_{n=1}^N \|Y_n\|^2.$$

Since the Y_n are identically distributed, the claim follows.

LEMMA 6. *Under the assumptions of* Section 2*, for any functions $v, u \in L^2$,*

$$\langle \Delta_N v, u\rangle \xrightarrow{P} \langle \Delta v, u\rangle.$$

Proof of Lemma 6. The result follows from the law of large numbers after noting that

$$\langle \Delta_N v, u\rangle = \frac{1}{N}\sum_{n=1}^N \langle X_n, v\rangle\langle Y_n, u\rangle$$

and

$$\mathrm{E}\,[\langle X_n, v\rangle\langle Y_n, u\rangle] = \mathrm{E}\,\left[\langle\langle X_n, v\rangle Y_n, u\rangle\right] = \langle \Delta v, u\rangle.$$

LEMMA 7. *Under the assumptions of* Section 2, $\langle \Delta_N \hat{v}_k, \hat{u}_j\rangle \xrightarrow{P} \langle \Delta v_k, u_j\rangle,\ j \leq q, k \leq p$.

Proof of Lemma 7. By Lemma 6, it suffices to show

$$\langle \Delta_N v_k, \hat{u}_j - u_j \rangle \quad \xrightarrow{P} \quad 0; \tag{19}$$

$$\langle \Delta_N \hat{v}_k - \Delta_N v_k, \hat{u}_j \rangle \quad \xrightarrow{P} \quad 0. \tag{20}$$

These relations follow from Lemma 3, relations (6) and Lemma 5.

ACKNOWLEDGEMENTS

We thank the Associate Editor and two referees for useful and accurate advice which helped us improve the presentation and substance of the paper.

REFERENCES

D. Bosq (2000). *Linear Processes in Function Spaces*. Springer, New York.

T. Cai & P. Hall (2006). Prediction in functional linear regression. *The Annals of Statistics*, 34, 2159–2179.

H. Cardot, F. Ferraty, A. Mas & P. Sarda (2003). Testing hypothesis in the functional linear model. *Scandinavian Journal of Statistics*, 30, 241–255.

R. B. Cattell (1966). The scree test for the number of factors. *Journal of Multivariate Behavioral Research*, 1, 245–276.

J-M. Chiou & H.-G. Müller (2007). Diagnostics for functional regression via residual processes. *Computational Statistics and Data Analysis*, 15, 4849–4863.

J-M. Chiou, H.-G. Müller & J-L. Wang (2004). Functional response models. *Statistica Sinica*, 14, 675–693.

A. Cuevas, M. Febrero & R. Fraiman (2002). Linear functional regression: the case of fixed design and functional response. *The Canadian Journal of Statistics*, 30, 285–300.

I. A. Danglis, J. U. Kozyra, Y. Kamide, D. Vassiliadis, A. S. Sharma, M. W. Liemohn, W. D. Gonzalez, B. T. Tsurutani & G. Lu (2003). Intense space storms: Critical issues and open disputes. *Journal of Geophysical Research*, 108, doi:10.1029/2002JA009722.

F. Ferraty & P. Vieu (2006). *Nonparametric Functional Data Analysis: Theory and Practice*. Springer, New York.

R. Gabrys & P. Kokoszka (2007). Portmanteau test of independence for functional observations. *Journal of the American Statistical Association*, 00, 000–000; Forthcoming.

P. Hall & M. Hosseini-Nasab (2006). On properties of functional principal components. *Journal of Royal Statistical Society Series B*, 68, 109–126.

P. Hall & M. Hosseini-Nasab (2007). Theory for high-order bounds in functional principal components analysis. Technical Report. The University of Melbourne, Australia.

A. Jach, P. Kokoszka, J. Sojka & L. Zhu (2006). Wavelet-based index of magnetic storm activity. *Journal of Geophysical Research*, 111, A09215.

Y. Kamide, W. Baumjohann, I. A. Danglis, W. D. Gonzalez, M. Grande, J. A. Joselyn, R. L. McPherron, J. L. Phillips, E. G. D. Reeves, G. Rostoker, A. S. Sharma, H. J. Singer, B. T. Tsurutani & V. M. Vasyliunas (1998). Current understanding of magnetic storms: Storm-substorm relationships. *Journal of Geophysical Research*, 103, 17705–17728.

M. G. Kivelson & C. T. Russell (1997). *Introduction to Space Physics*. Cambridge University Press.

P. Kokoszka, I. Maslova, J. Sojka & L. Zhu (2007). Effect of substroms on mid- and low-latitude horizontal intensity. Technical Report. Utah State University, city.

H.-G. Müller & U. Stadtmüller (2005). Generalized functional linear models. *The Annals of Statistics*, 33, 774–805.

J. O. Ramsay & B. W. Silverman (2005). *Functional Data Analysis*. Springer, New York.

G. Rostoker (2000). Effects of substorms on the stormtime ring current index Dst. *Annales Geophysicae*, 18, 1390–1398.

G. A. F. Seber & A. J. Lee (2003). *Linear Regression Analysis*. Wiley, New York.

W-Y. Xu & Y. Kamide (2004). Decomposition of daily geomagnetic variations by using method of natural orthogonal component. *Journal of Geophysical Research*, 109, A05218; DOI:10.1029/203JA010216.

F. Yao, H.-G. Müller & J-L. Wang (2005). Functional linear regression analysis for longitudinal data. *The Annals of Statistics*, 33, 2873–2903.

Received 13 April 2007
Accepted 1 November 2007

Piotr KOKOSZKA: piotr.kokoszka@usu.edu
Inga MASLOVA: Inga.Maslova@gmail.com
Department of Mathematics and Statistics
Utah State University, Logan
Utah 84321, USA

Jan SOJKA: fasojka@sojka.cass.usu.edu
Lie ZHU: zhu@cc.usu.edu
Center for Atmospheric and Space Sciences
Logan, Utah 84321, USA