Change point and trend analyses of annual expectile curves of tropical storms

P.Burdejova *
W.K.Härdle*
P.Kokoszka**
Q.Xiong*

BERLIN

ECONOMIC RISK

SFB 649

* Humboldt-Universität zu Berlin, Germany
** Colorado State University, United States of America

Change point and trend analyses of annual expectile curves of tropical storms[*]

P. Burdejova   W. K. Härdle [†]   P. Kokoszka[‡]   Q. Xiong

May 27, 2015

Abstract

Motivated by the conjectured existence of trends in the intensity of tropical storms, this paper proposes new inferential methodology to detect a trend in the annual pattern of environmental data. The new methodology can be applied to data which can be represented as annual curves which evolve from year to year. Other examples include annual temperature or log–precipitation curves at specific locations. Within a framework of a functional regression model, we derive two tests of significance of the slope function, which can be viewed as the slope coefficient in the regression of the annual curves on year. One of the tests relies on a Monte Carlo distribution to compute the critical values, the other is pivotal with the chi–square limit distribution. Full asymptotic justification of both tests is provided. Their finite sample properties are investigated by a simulation study. Applied to tropical storm data, these tests show that there is a significant trend in the shape of the annual pattern of upper wind speed levels of hurricanes.

JEL classification: C12, C15, C32, Q54
Keywords: change point, trend test, tropical storms, expectiles, functional data analysis

1   Introduction

A great deal of research in environmental and climate sciences has been dedicated to detecting change points and trends in various time series, including those related to temperature, precipitation and wind speed. In a typical setting, a scalar time series

---

$X_1, X_2, \ldots, X_N$ is analyzed. Sometimes several correlated series are considered. Most environmental and climate series exhibit a pronounced annual periodicity which must be removed, or otherwise accounted for, before statements on change–points or trends can be inferred. Sometimes, it is difficult to approximate the periodic component by a Fourier expansion due to the irregular domain and amplitude of observations within a year. The data that motivate this work are tropical storm wind speed data, examples are shown in Figure 1. The onset and end of typhoon and hurricane seasons, as well as their intensity, can change from year to year. We therefore propose to treat the data available for a whole year as a single high–dimensional data object and perform the change point and trend analyses on these objects rather than the scalar observations directly. Such an approach is now relatively well–established in the field of functional data analysis (FDA), the monographs of Horváth and Kokoszka (2012) or Ferraty and Vieu (2006) contain many examples. Methodological foundations of FDA are addressed in Ramsay and Silverman (2005), its mathematical foundations in Hsing and Eubank (2015). While the amount of information available in the data is invariably reduced by various smoothing and dimension reduction methods, the most important and relevant features of the data come into focus. In the problems we study in this paper, we are interested in the evolution of the annual pattern of tropical storms strength over several decades, not in specific hourly measurements.

Our functional methodology is combined with recent advances in expectile curve estimation, see Appendix A. We thus focus not only on the average pattern but on change points and trends in annual curves which describe the behavior at various levels of wind speed. This is illustrated in Figure 1. The curves in the middle summarize the pattern of average wind speed. These curves will exhibit some evolution from year to year. The curves above them summarize the annual patterns of the highest speeds; they may exhibit a different evolution than the average curves. This issue is well–known in climate research; typically trends in the averages are contrasted with trends in extremes. In our application, no modeling of extreme behavior is required, the expectile curves are within the range of the data points. They provide information of behavior which lies between the typical behavior and the unobservable extreme behavior. Following the work of Smith (1989), evaluation of trends in extremes has attracted a great deal of attention, with respect to change point analysis of extremes, we are aware only of the work of Dierckx and Teugels (2010).

The data objects that this paper studies have the form $X_n(t)$, where n refers to year, and t to time within the year. In the framework of functional data analysis, t is viewed as a continuous argument. The data are observed at a regular or irregular grid, but are converted to functional objects by means of various basis expansions which are defined for every t. We consider a sequence of curves $X_n(t, \tau)$ for several expectile levels $\tau \in (0, 1)$; these are similar to quantile levels. We are interested in detecting change points and trends in the functional time series $X_1(\cdot, \tau), X_2(\cdot, \tau), \ldots, X_N(\cdot, \tau)$. For this purpose, we use the existing change point test of Berkes et al. (2009) and develop two trend tests. No trend tests have presently been available for the data structure described above. These two tests form a methodological contribution to statistics, while the analysis of the expectile
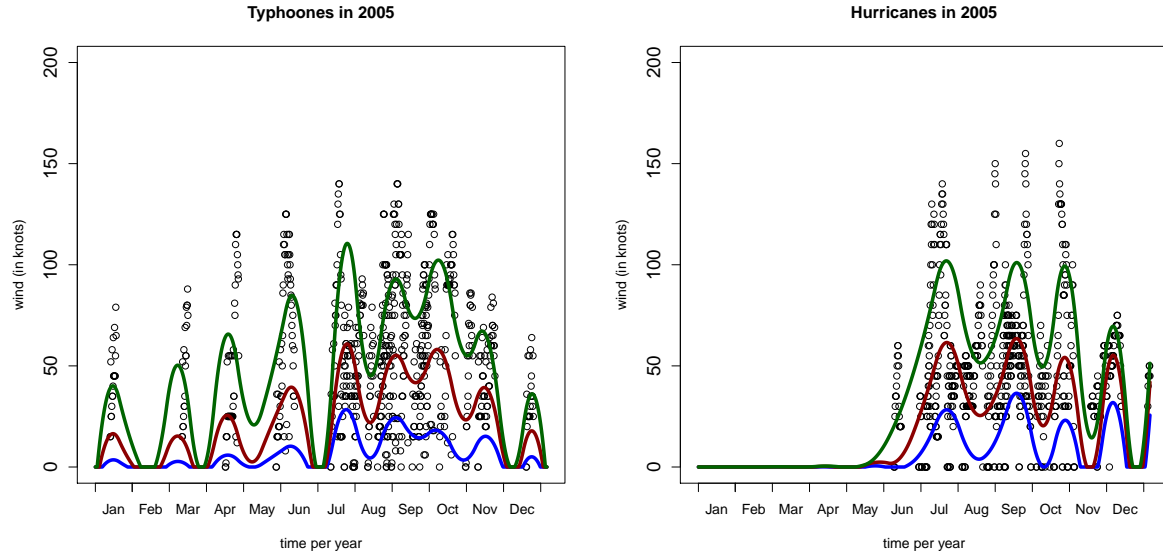
2

Typhoones in 2005   **Hurricanes in 2005**

Figure 1: Typhoons (left) and hurricanes (right) data in 2005 with expectile curves for $\tau = 0.1, 0.5$ and 0.9. ◉ data_load_hurricanes.R

curves of tropical storms provides an insight to climate science.

The paper is organized as follows. In Section 2 we review the test of Berkes et al. (2009) and present the two trend tests. These tools are applied in Section 3 to the analysis of expectile curves. Three appendices contain, respectively, background on expectile curves, a simulation study, and the details of the asymptotic theory for the trend tests. All codes are available as Quantlets ◉ on Quantnet (2015).

2   Change point and trend tests

This section presents the significance tests that will be applied to tropical storm data in Section 3. The change point test described in Section 2.1 was derived by Berkes et al. (2009), it is also described in Chapter 6 of Horváth and Kokoszka (2012). Trend tests introduced in Section 2.2 are new; their derivation and full large sample justification are presented in Section C. In both inferential settings, we consider as sequence of curves $X_n(t), t \in [0, 1], \ n = 1, 2, \ldots N$. The index n can be identified with year, the index t with time within the year normalized to unit interval. The exposition that follows uses now fairly standard concepts of functional data analysis, including functional principal components (FPC's) and their empirical counterparts (EFPC's), see e.g. Chapter 3 of Horváth and Kokoszka (2012).

2.1 Change point test

In change point tests, the null hypothesis is that the mean function does not change with year:

$$H_0: \ \mathsf{E}X_1 = \mathsf{E}X_2 = \ldots = \mathsf{E}X_N.$$

The specific value of the mean is not part of the null hypothesis. The alternative is that there are change points $k_1^*, k_2^*, \ldots, k_M^*$ such that the means X_i are not the same in all segments $(k_{m-1}^*, k_m^*]$. The theory and practice of change points tests have been described in many textbooks, e.g. Brodsky and Darkhovsky (1993), Csörgő and Horváth (1997), Chen and Gupta (2011), so we do not dwell on the background on move on to the description of the test of Berkes et al. (2009).

The test is based on the normalized differences of estimated mean functions:

$$P_k(t, \tau) = \frac{k(N-k)}{N} \left\{ \hat{\mu}_k(t, \tau) - \tilde{\mu}_k(t, \tau) \right\},$$

where

$$\hat{\mu}_k(t, \tau) = k^{-1} \sum_{i=1}^{k} X_i(t, \tau), \quad \tilde{\mu}_k(t, \tau) = (N-k)^{-1} \sum_{i=k+1}^{N} X_i(t, \tau).$$

Next, we compute the estimated functional principal components $\hat{v}_\ell$ of the curves X_n and calculate the scores

$$(2.1) \qquad \hat{\xi}_{j,n} = \int_0^1 \left\{ X_n(t) - \bar{X}_N(t) \right\} \hat{v}_j(t) dt.$$

The scores are calculated using the function `pca.fd` in the R package `fda`, see Chapter 7 of Ramsay et al. (2009). This function also produces estimated eigenvalues $\hat{\lambda}_\ell$ and the percentage of variance explained by the first d eigenvalues. We find the smallest d such that 85% of the variance is explained and calculate the test statistic

$$\widehat{S}_d = \frac{1}{N^2} \sum_{j=1}^{d} \frac{1}{\hat{\lambda}_j} \sum_{k=1}^{N} \left(\sum_{1 \le i \le k} \hat{\xi}_{j,i} - \frac{k}{N} \sum_{1 \le i \le k} \hat{\xi}_{j,i} \right).$$

For large N, the statistics $\widehat{S}_d$ has approximately the same distribution as the random variable K_d whose critical values are given Table 1, see Horváth and Kokoszka (2012) for more details.

2.2 Trend tests

Suppose the functions $X_n(t)$ follow the trend model

$$(2.2) \qquad\qquad X_n(t) = \alpha(t) + \beta(t)n + \varepsilon_n(t)$$

in which the error functions ε_n are iid and square integrable: $\mathsf{E} \, ||\varepsilon_n||^2 = \mathsf{E} \int_0^1 \varepsilon_n^2(t) dt < \infty$. The testing problem in this setting is

$$H_0: \ \beta = 0, \quad \text{vs.} \quad H_A: \ \beta \ne 0.$$

4

d	5	6	7	8	9	10	11	12
10%	1.2797	1.4852	1.6908	1.8974	2.0966	2.2886	2.4966	2.6862
5%	1.4690	1.6847	1.8956	2.1242	2.3227	2.5268	2.7444	2.9490
1%	1.8667	2.1260	2.3423	2.5893	2.8098	3.0339	3.2680	3.4911

Table 1: Critical values of the distribution of K_d, which approximates the distribution of the statistic $\widehat{S}_d$ for large N.

The parameter functions α, β are assumed to be elements of the space $L^2 = L^2([0, 1])$, so technically, $\beta = 0$ means that $\beta(t) = 0$ for almost all t.

A natural approach to testing is based on an estimator of β. If this estimator is small for all $t \in [0, 1]$, there is not enough evidence to reject H_0. The least squares estimator of β under H_0, cf. Section C, is given by

$$(2.3) \qquad \hat{\beta}(t) = \frac{6}{N(N+1)(N-1)} \sum_{k=1}^{N} (2k - N - 1)X_k(t).$$

Our first approach is based on the statistic $\int_0^1 \hat{\beta}^2(t)dt$. To describe its asymptotic distribution additional notation is needed. Introduce the covariance function of the errors $c_\varepsilon(t, s) = \mathsf{E}[\varepsilon_n(t)\varepsilon_n(s)]$. Denote by $\lambda_j, j = 1, 2, \ldots$ the eigenvalues of c_ε. Next, define the residuals

$$(2.4) \qquad \hat{\varepsilon}_n(t) = X_n(t) - \hat{\alpha}_n(t) - \hat{\beta}_n(t)n,$$

where

$$(2.5) \qquad \hat{\alpha}(t) = \frac{2}{N(N-1)} \sum_{k=1}^{N} (2N + 1 - 3k)X_k(t).$$

Denote by $\hat{\lambda}_j$ the eigenvalues of the empirical covariance function

$$(2.6) \qquad \hat{c}_\varepsilon(t, s) = \frac{1}{N} \sum_{n=1}^{N} \hat{\varepsilon}_n(t)\hat{\varepsilon}_n(s).$$

Theorem 2.1 describes large sample properties of the suitably normalized statistic $\int_0^1 \hat{\beta}^2(t)dt$.

THEOREM 2.1 *(i) Under H_0,*

$$(2.7) \qquad \widehat{\Lambda}_N = \frac{N^3}{12} \int_0^1 \left(\hat{\beta}(t) \right)^2 dt \xrightarrow{\mathcal{L}} \Lambda_\infty \stackrel{def}{=} \sum_{j=1}^{\infty} \lambda_j Z_j^2,$$

where $\{Z_j, j \geq 1\}$ are independent standard normal variables, and the λ_j are the eigenvalues of the covariance function c_ε.

5

(ii) Under H_A,

(2.8)
$$\mathrm{P}\left\{\widehat{\Lambda}_N > q_N(\alpha)\right\} \to 1, \quad \text{as } N \to \infty,$$

where $q_N(\alpha)$ is the $(1-\alpha)$th quantile of the distribution of $\Lambda_N = \sum_{j=1}^N \hat{\lambda}_j Z_j^2$.

Theorem 2.1 is proven in Section C.

The distribution of Λ_∞ can be approximated by the distribution of

(2.9)
$$\Lambda_N = \sum_{j=1}^N \hat{\lambda}_j Z_j^2.$$

This leads to the Monte Carlo test whose consistency is claimed in part (ii) of Theorem 2.1. To implement the test, we generate a large number, say $R = 10^4$, of independent replications of Λ_N (the $\hat{\lambda}_j$ are estimated only once, from the original sample). Denote these replications by $\Lambda_{N,r}, 1 \leq r \leq R$. The P–value of the test is computed as the fraction of the $\Lambda_{N,r}$ which are greater than $\widehat{\Lambda}_N$ (computed from the data).

It is also possible to develop a test similar to the test of Berkes et al. (2009) in the sense that a limit distribution is independent of the distribution of the data. In fact, in the trend model, the limit distribution is the usual chi–square distribution. This is stated in Theorem 2.2, in which we use the inner product notation $\langle f, g \rangle = \int_0^1 f(t)g(t)dt$.

THEOREM 2.2 *Suppose $\mathsf{E}\,||\varepsilon||^4 < \infty$ and*

(2.10)
$$\lambda_1 > \lambda_2 > \ldots > \lambda_q > \lambda_{q+1} > 0.$$

i) Under H_0,

(2.11)
$$\widehat{T}_N = \frac{N^3}{12} \sum_{j=1}^q \hat{\lambda}_j^{-1} \left\langle \hat{\beta}, \hat{v}_j \right\rangle^2 \xrightarrow{\mathcal{L}} \chi_q^2.$$

ii) If for some $1 \leq j \leq q$, $\langle \beta, v_j \rangle \neq 0$, then the test is consistent, i.e.

(2.12)
$$\mathrm{P}\left\{\widehat{T}_N > q(\alpha)\right\} \to 1, \quad \text{as } N \to \infty,$$

where $q(\alpha)$ is the $(1-\alpha)$th quantile of the chi–square distribution with q degrees of freedom.

Theorem 2.2 is proven in Section C.

Observe that to establish the consistency, it is not enough to assume $\beta \neq 0$ in L^2. Since the statistic $\widehat{T}_N$ is based on projections on the first q EFPC's, we must assume that the slope function β is not orthogonal to the subspace spanned by the first q FPC's.

3 Application to typhoon and hurricane data

In this section we apply the tests of Section 2 to annual expectile curves of wind speed data. The data have the form $X_n(t_i)$, where the times t_i are separated by six hours, and the index n stands for year. The value $X_n(t_i)$ is the wind speed in knots (1 kn = 0.5144 m/s). We work with two data sets: typhoons in the West Pacific area over the period 1946–2010, and hurricanes across the North Atlantic basin over the period 1947-2011. Both datasets are accessible free of charge at the website of Unisys Weather Information, UNISYS (2015).

Since there are about 1,460 time points t_i per year, we treat time $0 \leq t \leq T$ within a year as continuous, and the observed curves as functional data. For each year n, we construct expectile curves $X_n(t, \tau)$, for $\tau = 0.1, 0.2, \ldots, 0.9$. Examples of expectile curves we study are given in Figure 1. The index $\tau \in (0, 1)$ has the following interpretation. If $\tau = 0.5$, the curve $X_n(t, \tau)$ describes the mean strength of tropical storms throughout the year. If τ is close to 1, the curve $X_n(t, \tau)$ captures the annual pattern of highest wind speeds. If τ is close to zero, it does the same for the lowest wind speeds. More details are provided in Appendix A.

3.1 Change point analysis

The results of the application of the change–point test of Section 2.1 are shown in Table 2. For both data sets and at all levels τ, the test rejects the null hypothesis that the mean pattern does not change. As explained in Appendix A, the construction of the expectile curves involves the selection of a smoothing parameter λ. Table 2 shows the results for λ selected by the AIC criterion. To validate our conclusions, we performed the same analysis using λ which is either twice or half of the λ selected by AIC. In both cases, all empirical significance levels remained under 5%.

τ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
d	10	11	12	12	12	12	12	12	12
$\widehat{S}_d$	3.3522	3.2291	3.4317	3.4978	3.6564	3.8554	4.0342	4.2317	4.5084
	***	**	**	***	***	***	***	***	***
τ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
d	5	5	5	6	6	6	7	7	7
$\widehat{S}_d$	2.7440	3.3993	3.8759	4.4640	4.7141	4.8680	5.0366	4.9247	4.5740
	***	***	***	***	***	***	***	***	***

Table 2: Results of the application of the change point test of Section 2.1 to typhoon (upper panel) and hurricane (lower panel) expectile curves. Usual significance codes are used: ** – significant at 5% level, *** - at 1% level.

The change point test shows that for all expectile levels τ, there are statistically significant changes in the annual pattern. It is instructive to complement the above inferential analysis by simple exploratory analysis that reveals some dependence on the level τ.
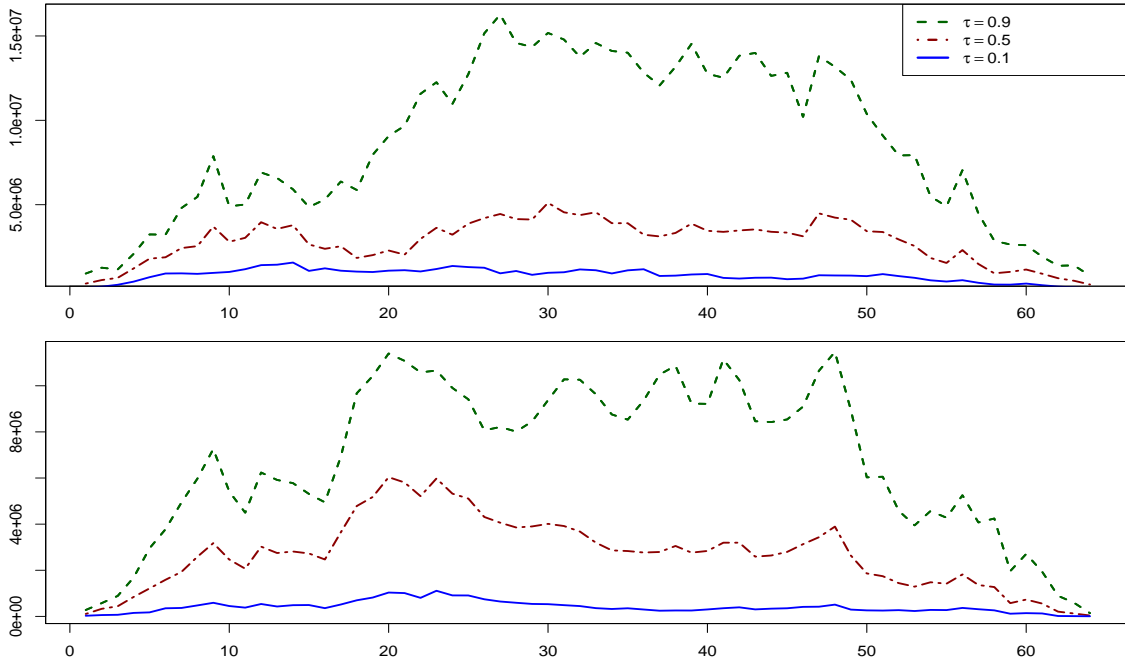
Figure 2: The squared norms $P_n(\tau)$ showing the magnitude of change in mean annual pattern for expectile curves of typhoons (upper panel) and hurricanes (lower panel). The largest changes occur in the expectile curves corresponding to $\tau = 0.9$.

🔴 P_beta_est.R

Consider squared norms

$$P_n(\tau) = \int_0^T P_n^2(t, \tau)dt$$

of the normalized differences introduced in Section 2.1. The plot of $P_n(\tau)$ against the year index n shows the magnitude of change of the mean function. We display such plots in Figure 2. They show that the largest changes occur for the expectile levels τ close to one; there is more change when focusing on the highest wind speeds than on median or lowest speeds.

The change point analysis above shows that the pattern of typhoon and hurricane wind speeds cannot be treated as stable over the sample periods we study. In the next section, we investigate if this instability can be attributed to systematic trends.

3.2   Trend analysis

We now apply the trend tests introduced in Section 2.2 to typhoon and hurricane expectile curves. In the Monte Carlo test based on Theorem 2.1, we use 10^4 replications of

the random variable Λ_N defined by (2.9). In the chi–square test based on Theorem 2.2, we determine q as the smallest number which explains at least 85% of the variance of the residual curves $\hat{\varepsilon}_n$ defined by (2.4). The results of the tests are presented in Tables 3 and 4.

τ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
typhoons P–value	0.365	0.537	0.545	0.495	0.438	0.381	0.329	0.316	0.269
hurricanes P–value	0.439	0.239	0.133	0.081	0.062	0.047	0.038	0.040	0.055

Table 3: P–values for the Monte Carlo trend test based on Theorem 2.1

τ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
q	10	11	12	12	12	12	12	12	12
typhoons P–value	0.534	0.705	0.722	0.688	0.587	0.466	0.382	0.371	0.453
q	5	5	5	6	6	6	7	7	7
hurricanes P–value	0.069	0.024	0.015	0.006	0.003	0.003	0.004	0.006	0.035

Table 4: P–values for the chi–square trend test based on Theorem 2.2

For the typhoon data, none of the two tests finds evidence of a trend. For the Hurricane data, the Monte Carlo test based on Theorem 2.1 indicates the existence of a trend for expectile levels $\tau = 0.6 - 0.9$ while the chi–square test of Theorem 2.2 for all τ except $\tau = 0.1$. Simulations reported in Appendix B show that the chi–square test tends to overreject for data generating processes (DGP's) of length and error structure similar to the tropical storm expectile curves. We therefore conclude that there is evidence for the existence of a trend for upper expectile functions of hurricane data. The estimated slope functions $\hat{\beta}$ are plotted in Figure 3.

We conclude the trend analysis by showing in Figure 5 the dependence on τ of the norm $\|\hat{\beta}\| = \sqrt{\int \hat{\beta}^2(t)dt}$ of the estimated slope function. Even though there is statistical evidence for nonzero slope function only for the upper expectiles of hurricane data, the exploratory analysis of the norms indicates that there is a very clear increasing dependence of the slope on τ. Based on the estimates $\hat{\beta}$, the slope functions β appear to be much smaller for lower expectiles, and to detect them, if they are in fact nonzero, larger sample sizes N would be needed. There is not much difference between the size of $\hat{\beta}$, for typhoon and hurricane data, but the $\hat{\beta}$ for hurricanes show a clear pattern with positive mass around July and November, and negative mass in early Fall. For the typhoon curves the pattern of mass accumulation is spread more uniformly throughout the year, with a pronounced negative mass in November. The significance tests we developed provide a statistical justification for these fairly subtle visual differences.

3.3   Main conclusions of data analysis

The change point tests have shown that the annual pattern of wind speeds for both hurricanes and typhoons cannot be treated as constant, no matter what expectile level is
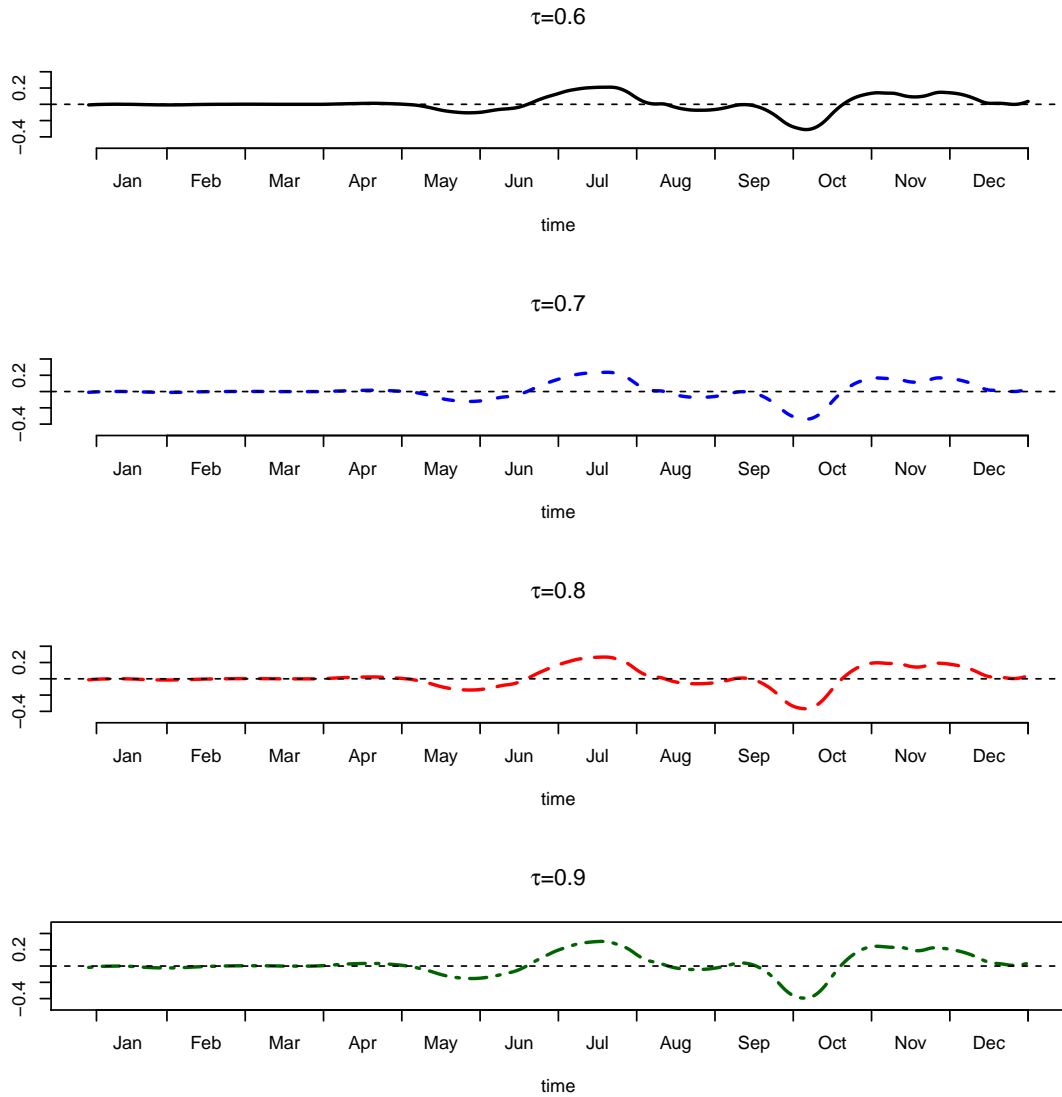
Figure 3: Estimated slope functions, $\hat{\beta}$, for upper expectile curves of **hurricane** data.
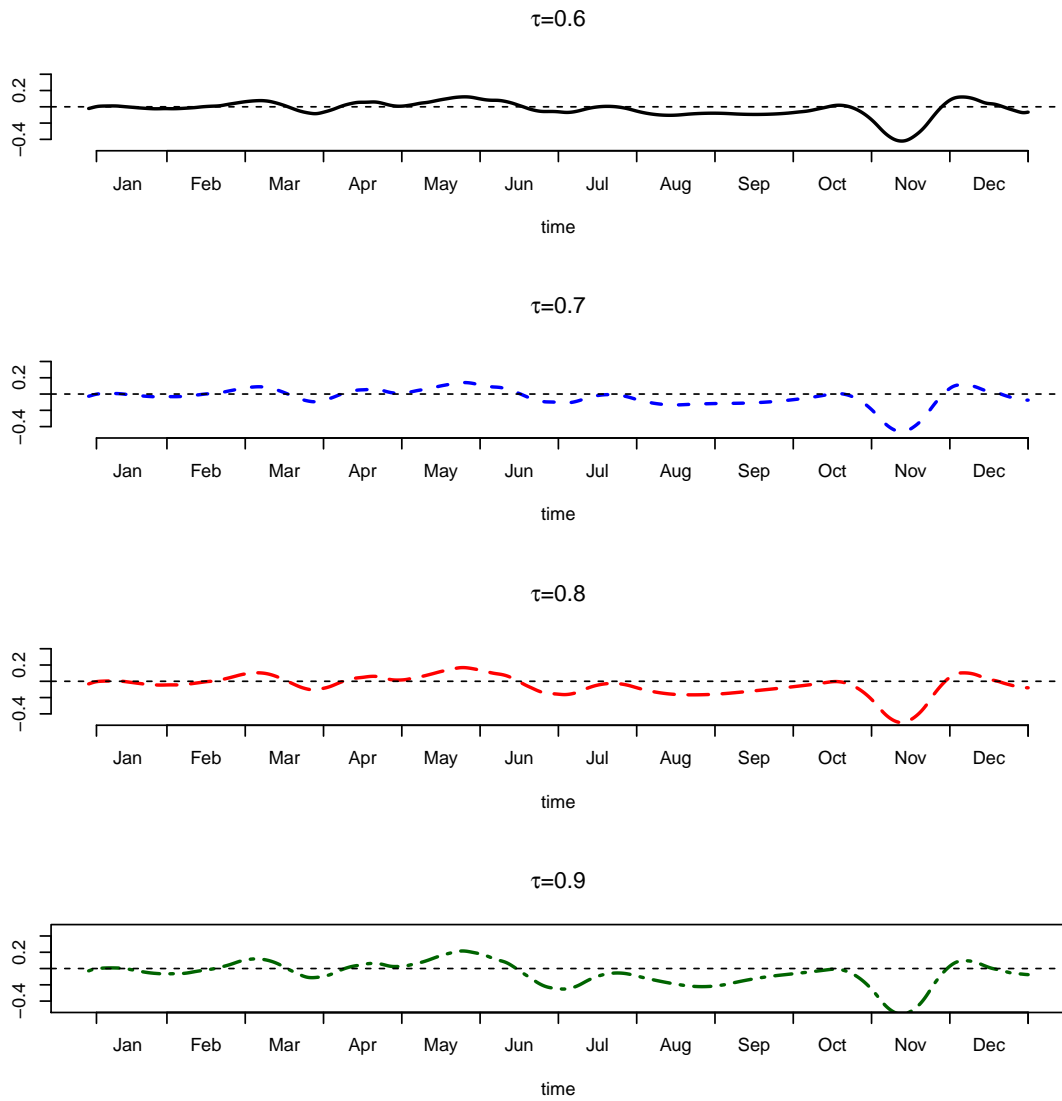
P_beta_est.R

Figure 4: Estimated slope functions, $\hat{\beta}$, for upper expectile curves of **typhoon** data.
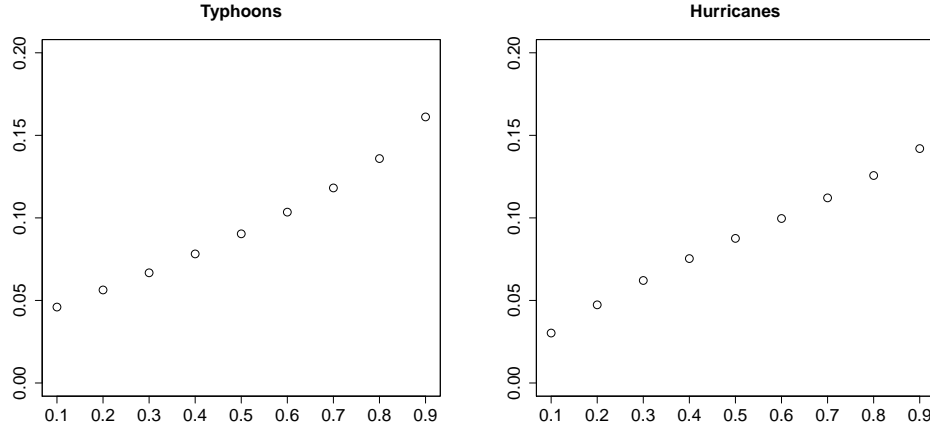
Figure 5: Norm of the slope function estimate, $\hat{\beta}$, as a function of the expectile level τ; typhoons (left), hurricanes (right).   Q P_beta_est.R

considered. The application of the new trend tests has focused on a more subtle question, which has however received a fair deal of attention: is there a trend in the intensity of tropical storms. A review of relevant research is not our aim, the paper of Kossin et al. (2013) provides background and references. There are two novel aspects to our approach: 1) focus on the annual curves, 2) separate analysis for each intensity level. Based on sixty years of data, our tests detect a trend in the upper wind speeds of Atlantic hurricanes. Exploratory analysis suggests a similar conclusion for Pacific typhoons, but it cannot be supported by low P–values with the amount of available data. These conclusions are similar to the findings of Kossin et al. (2013) who use different, custom–prepared, data sets. Their P–value for the existence of a trend in North Atlantic is less that 10^{-3}, but for the North–West pacific it is 0.03 (for South Pacific it is 0.09, 0.06 for the South Indian Ocean). Their analysis is concerned with the trend in the scalar data, not a trend in the annual pattern. They find all trends to be positive. In a sense, such trend coefficients can be viewed as averages of the annual curves like those displayed in Figures 3 and 4. The hurricane curves indeed have more positive mass, whereas for the typhoon curves the negative mass is larger (the typhoon curves are not statistically different from zero, according to our tests). The slope functions of the hurricanes indicate increasing intensity in summer and late fall, and decreasing intensity in early fall. For typhoons, these curves indicate decreasing intensity in November.

The conclusions of this paper which are supported by significance tests and do not contradict existing research are as follows:

1. The annual pattern of wind speeds of both hurricanes and typhoons has been changing at all wind speed levels over the last 60 years.

2. There is a significant trend in the shape of this pattern for upper wind speed levels of hurricanes.

12

A Construction of annual expectile curves

In this section we provide some background needed to understand how the expectile curves studied in Section 3 are constructed. The underlying concept of expectiles was first discussed by Newey and Powell (1987) and further analyzed in several directions, e.g. Efron (1991) or Rossi and Harwey (2009) focused on time-varying expectiles. Most relevant to our setting is the paper by Schnabel and Eilers (2009), which extended the work of Eilers and Marx (1996). It combined the LAWS (least average weighted squares) algorithm with P-splines in order to estimate expectile curves. Recent applications include Guo and Härdle (2012), Sobotka et al. (2013) and Guo et al. (2015) or more applicable one in finance by Taylor (2008), where Value at risk (VaR) and Expected shortfall (ES) were estimated using expectiles.

Suppose Y is a square integrable random variable with density f. Fix $\tau \in (0,1)$. The τ-th expectile $e = e_\tau$ is the number which minimizes

$$\mathsf{E}_\tau(e) = (1-\tau)\mathsf{E}_-(e) + \tau\mathsf{E}_+(e),$$

where

$$\mathsf{E}_-(e) = \int_{-\infty}^{e} (u-e)^2 f(u)du = \mathsf{E}\left[(X-e)^2\,\mathbf{I}\{X \le e\}\right];$$

$$\mathsf{E}_+(e) = \int_{e}^{\infty} (u-e)^2 f(u)du = \mathsf{E}\left[(X-e)^2\,\mathbf{I}\{X > e\}\right].$$

Let $F(y)$ be a cumulative distribution function of Y and $G(y) = \int_{-\infty}^{y} uf(u)du$ its first partial moment. For every $\tau \in (0,1)$, the expectile $e = e_\tau$ can be obtained as τ-quantile of the distribution function $T(x)$, where

$$T(x) = \frac{G(x) - xF(x)}{2\{G(x) - F(x)\} + \left\{x - \int_{-\infty}^{\infty} udF(u)\right\}}.$$

If $\tau = 1/2$, then $2\mathsf{E}_\tau(e) = \mathsf{E}\left[(X-e)^2\right]$, so $e = \mathsf{E}X$. Expectiles have a similar interpretation as quantiles, but have some desirable properties outlined in the references cited above.

Consider now a scatter plot of points (t_i, x_i), $1 \le i \le I$. In our applications, the t_i correspond to times within a year at which wind speed is measured and x_i to the wind speed. Since the form of the dependence of the x_i on the t_i is unknown, a B-spline expansion is used. We thus assume that

$$x_i \approx g_a(t_i) = \sum_{j=1}^{J} a_j B_j(t_i),$$

and find coefficients $a = (a_1, a_2, \ldots, a_J)$ which minimize

$$S_\tau(a) = (1-\tau)S_-(a) + \tau S_+(a),$$

13

where

$$S_-(a) = \sum_{x_i \le g_a(t_i)} \{x_i - g_a(t_i)\}^2, \quad S_+(a) = \sum_{x_i > g_a(t_i)} \{x_i - g_a(t_i)\}^2.$$

If τ is close to 1, then $S_+(a)$ must be made small. This means that the curve g_a will be above most of the points (t_i, x_i).

Denote a matrix of B-splines differences as D. In order to control the smoothness of curves we can add penalization and minimize

$$S_\tau(a) + \lambda a^\top D^\top D a,$$

with λ as shrinkage parameter chosen by a desired criterion.

In our calculations, we chose λ according to AIC criterion. The whole algorithm is implemented in the R package expectreg, see Sobotka et al. (2014) for further details.

In some applications, particularly to small and volatile datasets, one can observe crossing curves, even if this is not possible in theory. For that reason expectile curves can be modeled as bivariate functions of the form

$$\mu(x, \tau) = \sum_{k=1}^{K} \sum_{\ell=1}^{L} a_{k\ell} B_k(x) C_\ell(p).$$

The coefficients $a_{k\ell}$ are estimated by minimizing

$$S(a) = \sum_{i=1}^{n} \sum_{j=1}^{J} w_i(\tau_j) \{x_i - \mu(t_i, \tau_j)\}^2$$

where

$$w_i(\tau_j) = \begin{cases} 1 - \tau & \text{if } x_i \le \mu(t_i, \tau_j) \\ \tau & \text{if } x_i > \mu(t_i, \tau_j). \end{cases}$$

Expectile and quantile sheets are explained in more detail by Schnabel (2011) or Schnabel and Eilers (2013).

B Trend tests: finite sample performance

A simulation study validating the change point test of Section 2.1 is reported in Berkes et al. (2009). In this section, we examine the finite sample performance of the trend tests introduced in Section 2.2.

Data generating processes We consider two models for the error functions $\varepsilon_n(t)$. The first is a generic Gaussian model in which we take the $\varepsilon_n(t)$ to be Brownian bridges $B_n(t)$. We represent Brownian bridge as a Fourier series with stochastic coefficients (the Karhunen–Loéve expansion, e.g. Bosq (2000)):

$$B_n(t) = \sqrt{2} \sum_{j=1}^{\infty} Z_{nj} \frac{\sin(j\pi t)}{j\pi} \approx \sqrt{2} \sum_{j=1}^{J} Z_{nj} \frac{\sin(j\pi t)}{j\pi},$$

where $\{Z_j, j \geq 1\}$ are independent standard normal random variables. We set $J = 100$ so the trajectories of the B_n have similar smoothness as the typhoon and hurricane expectile curves.

The second model for the ε_n is based more directly on the tropical storm data. We proceed as follows. We consider $\tau = 0.1, 0.5, 0.9$. For each level τ, we compute the sample mean function and the sample functional principal components $\hat{v}_j(t; \tau)$ of the expectile curves $X_n(t, \tau)$. Next we compute the scores $\xi_{jn}(\tau)$ according to (2.1). Denote by $\sigma_j(\tau)$ the standard deviation of the $\xi_{jn}(\tau), 1 \leq n \leq N, (N = 65)$. The ε_n are generated as independent realizations of the random function

$$\varepsilon(t; \tau) = \sum_{j=1}^{q} \sigma_j(\tau) Z_j \hat{v}_j(t; \tau), \quad Z_j \sim \text{ iid N}(0, 1),$$

with q determined from the original expectile curves according to the 85% rule. We thus have four models for the error curves which we refer to as BB, E1, E5, E9. The errors E1, E5, E9 are different depending on whether hurricane or typhoon data are used. The empirical rejection rates are however very similar in both cases. We display the results for the errors based on the hurricane data.

We generate artificial data according to the specification

$$X_n(t) = b\beta(t)n + \varepsilon_n(t).$$

To find empirical size, we set $\beta(t) = \beta_0(t) = 0$. To find empirical power, power, we use the slope functions

$$\beta_1(t) = -\frac{\cos\left(\frac{t\pi 3}{2}\right)}{100}; \quad \beta_2(t) = \frac{\sin(t\pi 20)}{100},$$

which are graphed in Figure 6. The constant b is used to adjust the magnitude of the departure from the null hypothesis. For E1, E5 and E9 error curves we set $b = 20$, for BB errors we use $b = 1$. The different values are used to ensure similar signal to noise ration for both types of errors.

We consider sample sizes $N = 30, 60, 120$ Empirical rejection rates are shown in Tables 5 and 6. The Monte Carlo test, generally has slightly better size and power, but the pivotal chi–square test performs well too. It tends to overreject under H_0 (for $N = 60$ and $N = 120$), which may explain the smaller P–values in Table 4 as compared to Table 3.
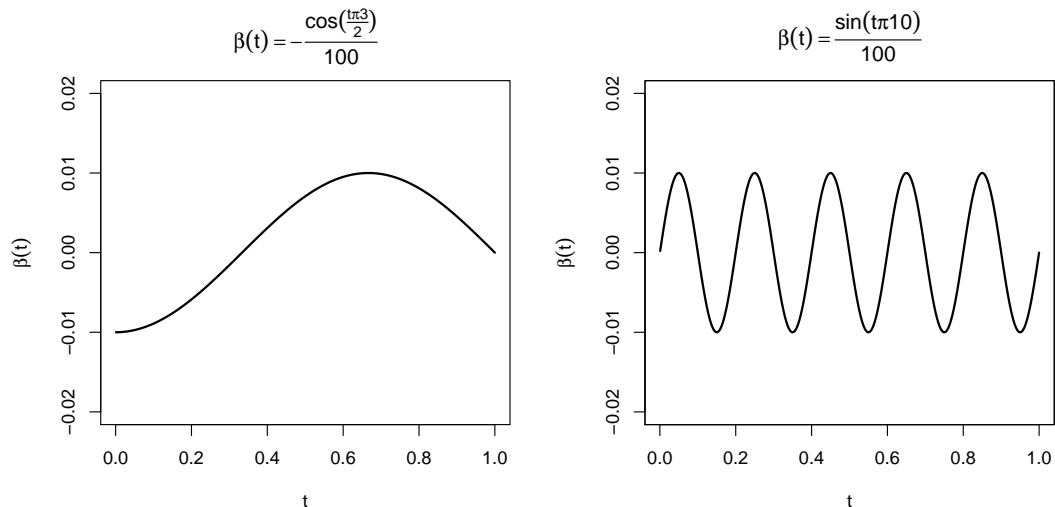
15

Figure 6: Slope functions used to assess power.

BB	β_0	β_1	β_2		E1	β_0	β_1	β_2
$N{=}30$	0.055	0.175	0.136		$N{=}30$	0.060	0.082	0.078
$N{=}60$	0.056	0.967	1.000		$N{=}60$	0.045	0.438	0.440
$N{=}120$	0.064	1.000	1.000		$N{=}120$	0.042	1.000	1.000

E5	β_0	β_1	β_2		E9	β_0	β_1	β_2
$N{=}30$	0.042	0.072	0.060		$N{=}30$	0.069	0.081	0.091
$N{=}60$	0.047	0.435	0.438		$N{=}60$	0.058	0.435	0.404
$N{=}120$	0.044	1.000	1.000		$N{=}120$	0.042	1.000	1.000

Table 5: Rejection rates of the **Monte Carlo test**. Columns corresponding to β_0 report empirical size, those to β_1 and β_2, empirical power.

BB	β_0	β_1	β_2		E1	β_0	β_1	β_2
$N{=}30$	0.064	0.344	0.053		$N{=}30$	0.053	0.071	0.089
$N{=}60$	0.058	0.995	0.085		$N{=}60$	0.058	0.215	0.220
$N{=}120$	0.069	1.000	0.238		$N{=}120$	0.056	0.975	0.971

E5	β_0	β_1	β_2		E9	β_0	β_1	β_2
$N{=}30$	0.047	0.065	0.044		$N{=}30$	0.051	0.075	0.085
$N{=}60$	0.064	0.249	0.193		$N{=}60$	0.065	0.216	0.234
$N{=}120$	0.049	0.982	0.898		$N{=}120$	0.058	0.929	0.967

Table 6: Rejection rates of the **Chi–square test**. Columns corresponding to β_0 report empirical size, those to β_1 and β_2, empirical power.

C Trend tests: large sample justification

Representing trend model (2.2) as the regression

$$
\begin{bmatrix} X_1(t) \\ \vdots \\ X_N(t) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ \vdots & \vdots \\ 1 & N \end{bmatrix} \cdot \begin{bmatrix} \alpha(t) \\ \beta(t) \end{bmatrix} + \begin{bmatrix} \varepsilon_1(t) \\ \vdots \\ \varepsilon_N(t) \end{bmatrix},
$$

we obtain the least squares estimators (2.3) and (2.5). Direct verification shows that

$$
c_\beta(t,s) \stackrel{\text{def}}{=} \mathrm{Cov}\left\{ \hat\beta(t), \hat\beta(s) \right\} = A_N c_\varepsilon(t,s),
$$

where

$$
A_N = \frac{12}{N(N+1)(N-1)}.
$$

The constant A_N is repeatedly used in the proofs of Theorems 2.1 and 2.2.

C.1 Proof of Theorem 2.1

PROOF OF PART (I): Under H_0 ($\beta = 0$),

$$
\hat\beta(t) = A_N \sum_{k=1}^{N} k\varepsilon_k(t) - \frac{1}{2} A_N(N+1) \sum_{k=1}^{N} \varepsilon_k(t).
$$

Using the identity

(C.1)
$$
\sum_{k=1}^{N} k\varepsilon_k = N \sum_{n=1}^{N} \varepsilon_n - \sum_{k=1}^{N-1}\sum_{n=1}^{k} \varepsilon_n,
$$

we have

(C.2)
$$
\begin{aligned}
\hat\beta(t) &= A_N \sum_{k=1}^{N} k\varepsilon_k(t) - \frac{1}{2} A_N(N+1) \sum_{k=1}^{N} \varepsilon_k(t) \\
&= A_N \left(N \sum_{n=1}^{N} \varepsilon_n(t) - \sum_{k=1}^{N-1}\sum_{n=1}^{k} \varepsilon_n(t) \right) - \frac{1}{2} A_N(N+1) \sum_{n=1}^{N} \varepsilon_n(t).
\end{aligned}
$$

To determine the limit behavior of $\hat\beta(t)$, we thus need an invariance principle for the partial sum process.

$$
S_N(x,t) = \frac{1}{\sqrt{N}} \sum_{1 \le n \le [Nx]} \varepsilon_n(t), \quad 0 \le x, t \le 1.
$$

A result of this type has recently been established by Berkes et al. (2013). It states that

(C.3)
$$
S_N(x,t) \xrightarrow{\mathcal{L}} \Gamma(x,t),
$$

where $\Gamma(x,t)$ is the two parameter Gaussian process which admits the representation

$$(C.4) \qquad \Gamma(x,t) = \sum_{j=1}^{\infty} \sqrt{\lambda_j} W_j(x) v_j(t),$$

where $\{W_j(x), 0 \le x \le 1\}$ are independent standard Wiener processes on $[0,1]$. The λ_j and the v_j are, respectively, the eigenvalues and the eigenfunctions of the covariance function $c_\varepsilon(t,s) = \mathsf{E}[\varepsilon_n(t)\varepsilon_n(s)]$. In (C.3), and whenever weak convergence of two parameter processes is concerned, $\xrightarrow{d}$ denotes the convergence in the Skorokhod space $D([0,1], L^2)$.

Since $A_N \sim 12N^{-3}$, (C.2) implies

$$\hat{\beta}(t) = A_N N^{\frac{3}{2}} S_N(1,t) - A_N N^{\frac{1}{2}} \sum_{k=1}^{N-1} S_N\left(\frac{k}{N},t\right) - \frac{1}{2} A_N (N+1) N^{\frac{1}{2}} S_N(1,t)$$

$$\sim 12 N^{-\frac{3}{2}} S_N(1,t) - 12 N^{-\frac{3}{2}} \left\{ \frac{1}{N} \sum_{k=1}^{N-1} S_N\left(\frac{k}{N},t\right) \right\} - 6 N^{-\frac{3}{2}} S_N(1,t)$$

$$= 6 N^{-\frac{3}{2}} S_N(1,t) - 12 N^{-\frac{3}{2}} \left\{ \frac{1}{N} \sum_{k=1}^{N-1} S_N\left(\frac{k}{N},t\right) \right\}.$$

By the continuous mapping theorem and (C.3)

$$\frac{1}{N} \sum_{k=1}^{N-1} S_N\left(\frac{k}{N},t\right) \xrightarrow{\mathcal{L}} \int_0^1 \Gamma(x,t)dx,$$

Thus

$$(C.5) \qquad \frac{N^{\frac{3}{2}}}{6} \hat{\beta}(t) \xrightarrow{\mathcal{L}} \Gamma(1,t) - 2 \int_0^1 \Gamma(x,t)dx.$$

Using the continuous mapping theorem again, we obtain

$$\frac{N^3}{36} \int_0^1 \left\{ \hat{\beta}(t) \right\}^2 dt \xrightarrow{\mathcal{L}} \int_0^1 \left\{ \Gamma(1,t) - 2 \int_0^1 \Gamma(x,t)dx \right\}^2 dt.$$

Set

$$(C.6) \qquad D_j = W_j(1) - 2 \int_0^1 W_j(x)dx,$$

so that, by (C.4), we have

$$\Gamma(1,t) - 2 \int_0^1 \Gamma(x,t)dx = \sum_{j=1}^{\infty} \sqrt{\lambda_j} D_j v_j(t).$$

Then, by Parseval's identity,

$$(C.7) \qquad \int_0^1 \left\{ \Gamma(1,t) - 2 \int_0^1 \Gamma(x,t)dx \right\}^2 dt = \left\| \sum_{j=1}^\infty \sqrt{\lambda_j} D_j v_j \right\|^2 = \sum_{j=1}^\infty \lambda_j D_j^2.$$

The random variables D_j are independent normal with mean zero and variance

$$\mathrm{Var}[D_j] = \mathsf{E}\left[W(1) - 2 \int_0^1 W(x)dx \right]^2$$

$$= \mathsf{E}W^2(1) - 4\mathsf{E}\left[W(1) \int_0^1 W(x)dx \right] + 4\mathsf{E}\left[\int_0^1 W(x)dx \right]^2$$

$$= \frac{1}{3}.$$

We can write $D_j = \frac{1}{\sqrt{3}} Z_j$, where Z_j are standard normal variables. By (C.7)

$$\int_0^1 \left\{ \Gamma(1,t) - 2 \int_0^1 \Gamma(x,t)dx \right\}^2 dt = \frac{1}{3} \sum_{j=1}^\infty \lambda_j Z_j^2.$$

Thus (2.7) is proven.

PROOF OF PART (II): The proof follows from several lemmas. It is assumed throughout that H_A holds, i.e $||\beta|| > 0$. The argument relies on Lemma C.1 whose proof follows from the relevant definitions, and so is omitted.

LEMMA C.1 *Suppose $\{X_n\}$ and $\{q_n\}$ are sequences of random variables. Suppose further that $\{X_n\}$ diverges to infinity in probability and $\{q_n\}$ is bounded in probability , i.e. for each $M, \lim_{n\to\infty} \mathrm{P}(X_n > M) = 1$ and for each $\varepsilon > 0$, there is M such that $\mathrm{P}(q_n > M) < \varepsilon$, if $n > n_0$. Then*

$$\lim_{n\to\infty} \mathrm{P}(X_n > q_n) = 1.$$

Relation (2.8) now follows from Lemmas C.2 and C.3.

LEMMA C.2 *The statistic $\widehat{\Lambda}_N$ defined by (2.7) satisfies $\widehat{\Lambda}_N \xrightarrow{P} \infty$.*

PROOF: Decompose $\hat{\beta}(t)$ as

$$(C.8) \qquad\qquad \hat{\beta}(t) = \beta(t) + G_N(t),$$

where

$$G_N(t) = \frac{1}{2} A_N \sum_{k=1}^N (2k - N - 1)\varepsilon_k(t).$$

19

Observe that $G_N(t)$ is equal to the estimator $\hat{\beta}(t)$ under H_0. Therefore, by (C.5),

$$N^{3/2}G_N(t) \xrightarrow{\mathcal{L}} 6\left\{\Gamma(1,t) - 2\int_0^1 \Gamma(x,t)dx\right\} \stackrel{def}{=} U(t).$$

Consequently, as $N \to \infty$

$$N^3 \int \hat{\beta}^2(t)dt = \int \left\{N^{3/2}\beta(t) + N^{3/2}G_N(t)\right\}^2 dt \sim \int \left\{N^{3/2}\beta(t) + U(t)\right\}^2 dt \xrightarrow{P} \infty.$$

More precisely,

$$N^{-3}\widehat{\Lambda}_N \sim \frac{1}{12}\int \left\{\beta(t) + N^{-3/2}U(t)\right\}^2 dt \xrightarrow{P} \frac{1}{12}\int \beta^2(t)dt.$$

LEMMA C.3 *The sequence $\{\Lambda_N\}$ defined by (2.9) is bounded in probability, i.e. $\mathcal{O}_p(1)$.*

PROOF: Since the $\hat{\lambda}_j$ are fixed in the generation of the replications in the Monte Carlo test, the variables Z_j are independent of the $\hat{\lambda}_j$. Therefore, since $\mathsf{E}Z_j^2 = 1$,

$$\mathsf{E}\Lambda_N = \sum_{j=1}^N \mathsf{E}\hat{\lambda}_j.$$

The definition of the $\hat{\lambda}_j$ as the eigenvalues of the covariance operator with $\hat{c}_\varepsilon(\cdot,\cdot)$ defined by (2.4) and (2.6) implies that

$$\sum_{j=1}^N \hat{\lambda}_j = \frac{1}{N}\sum_{n=1}^N ||\hat{\varepsilon}_n||^2.$$

This is the decomposition of functional sample variance, see Horváth and Kokoszka (2012), p. 40. Therefore, if we can show that

(C.9) $$\limsup_{N\to\infty} \frac{1}{N}\sum_{n=1}^N \mathsf{E}||\hat{\varepsilon}_n||^2 < \infty,$$

then we can conclude that $\limsup_{N\to\infty} \mathsf{E}\Lambda_N < \infty$, which in turn implies that the sequence $\{\Lambda_N\}$ bounded in probability.

The decomposition

(C.10) $$\hat{\varepsilon}_n(t) = \varepsilon_n(t) + \left\{\alpha(t) - \hat{\alpha}(t)\right\} + n\left\{\beta(t) - \hat{\beta}(t)\right\},$$

implies that for some constant C,

(C.11) $$||\hat{\varepsilon}_n||^2 \leq C\left(||\varepsilon_n||^2 + ||\hat{\alpha} - \alpha||^2 + ||n(\hat{\beta} - \beta)||^2\right).$$

First note that

$$\frac{1}{N}\sum_{n=1}^{N}\mathsf{E}||\varepsilon_n||^2 = \mathsf{E}\left[\int\left\{\frac{1}{N}\sum_{n=1}^{N}\varepsilon_n^2(t)\right\}dt\right]$$

$$= \int\left\{\mathsf{E}\left[\frac{1}{N}\sum_{n=1}^{N}\varepsilon_n^2(t)\right]\right\}dt$$

$$= \int \mathsf{E}\varepsilon_1^2(t)dt = \mathsf{E}||\varepsilon_1||^2 < \infty.$$

Next, observe that

$$\frac{1}{N}\sum_{n=1}^{N}\mathsf{E}||\hat{\alpha} - \alpha||^2 = \mathsf{E}||\hat{\alpha} - \alpha||^2$$

$$= \int\left\{\mathsf{E}\left[\hat{\alpha}(t) - \alpha(t)\right]^2\right\}dt$$

$$= \int \mathsf{E}\left[\frac{2}{N(N-1)}\sum_{k=1}^{N}(2N+1-3k)\varepsilon_k(t)\right]^2 dt$$

$$= \frac{2(2N+1)}{N(N-1)}\mathsf{E}||\varepsilon_1||^2 \to 0.$$

Similarly,

$$\frac{1}{N}\sum_{n=1}^{N}\mathsf{E}||n(\hat{\beta} - \beta)||^2 = \frac{(N+1)(2N+1)}{6}\mathsf{E}||\hat{\beta} - \beta||^2$$

$$= \frac{(N+1)(2N+1)}{6}\int\left\{\mathsf{E}\left[\hat{\beta}(t) - \beta(t)\right]^2\right\}dt$$

$$= \frac{(N+1)(2N+1)}{6}\int \mathsf{E}\left[\frac{6}{N(N-1)(N+1)}\sum_{k=1}^{N}(2k-N-1)\varepsilon_k(t)\right]^2 dt$$

$$= \frac{2(2N+1)}{N(N-1)}\mathsf{E}||\varepsilon_1||^2 \to 0.$$

Thus (C.9) holds. Therefore $\sup_N \mathsf{E}\Lambda_N =: C_\Lambda < \infty$, and so $\mathsf{P}(\Lambda_N > M) \le M^{-1}C_\Lambda$ can be made arbitrarily small by choosing M sufficiently large. The conclusion follows.

C.2 Proof of Theorem 2.2

PROOF OF PART (I): Under H_0, by (C.2), (C.4) and consistency of estimated eigenfunctions $\hat{v}_j$, $(\hat{v}_j \xrightarrow{P} v_j)$,

$$\left\langle \frac{N^{\frac{3}{2}}}{6}\hat{\beta}, \hat{v}_j \right\rangle^2 \xrightarrow{\mathcal{L}} \left\langle \Gamma(1, \cdot) - 2\int_0^1 \Gamma(x, \cdot)dx, v_j \right\rangle^2$$

$$= \left\langle \sum_{k=1}^{\infty} \sqrt{\lambda_k}W_k(1)v_k - 2\int_0^1 \sum_{k=1}^{\infty} \sqrt{\lambda_k}W_k(x)v_k, v_j \right\rangle^2$$

$$= \left[\sum_{k=1}^{\infty} \sqrt{\lambda_k}\left\{ W_k(1) - 2\int_0^1 W_k(x)dx \right\} \langle v_k, v_j \rangle \right]^2$$

$$= \lambda_j \left\{ W_j(1) - 2\int_0^1 W_j(x)dx \right\}^2$$

$$= \lambda_j D_j^2 = \frac{1}{3}\lambda_j Z_j^2,$$

with the random variables D_j defined in (C.6), and Z_j standard normal variables. It follows that

$$\widehat{T}_N = \frac{N^3}{12}\sum_{j=1}^{q} \hat{\lambda}_j^{-1} \left\langle \hat{\beta}, \hat{v}_j \right\rangle^2 = 3\sum_{j=1}^{q} \hat{\lambda}_j^{-1} \left\langle \frac{N^{\frac{3}{2}}}{6}\hat{\beta}, \hat{v}_j \right\rangle^2 \xrightarrow{\mathcal{L}} \sum_{j=1}^{q} Z_j^2 \overset{\mathcal{L}}{=} \chi_q^2.$$

PROOF OF PART (II): We must show that $\widehat{T}_N \xrightarrow{P} \infty$, if $\langle \beta, v_j \rangle \neq 0$ for some $1 \leq j \leq q$. It is enough to show that

$$\sum_{j=1}^{q} \hat{\lambda}_j^{-1} \left\langle \hat{\beta}, \hat{v}_j \right\rangle^2 \xrightarrow{P} \sum_{j=1}^{q} \lambda_j^{-1} \langle \beta, v_j \rangle^2,$$

because the right–hand side is positive. The verification of the above convergence reduces to

(C.12)
$$\left\| \hat{\beta} - \beta \right\| \xrightarrow{P} 0$$

and, for $1 \leq j \leq q$,

(C.13)
$$\|\hat{v}_j - v_j\| \xrightarrow{P} 0, \quad \hat{\lambda}_j \xrightarrow{P} \lambda_j.$$

To prove relation (C.12), observe first that by decomposition (C.8),

$$\mathsf{E}\left\| \hat{\beta} - \beta \right\| = \mathsf{E}\left\| G_N \right\| \leq \left\{ \mathsf{E}\left\| G_N \right\|^2 \right\}^{\frac{1}{2}} = \left\{ \mathsf{E}\int G_N^2(t)dt \right\}^{\frac{1}{2}}.$$

To calculate the last expected value, we will use the identity

$$\frac{1}{4}A_N \sum_{k=1}^{N}(2k-N-1)^2 = 1,$$

which follows from algebraic manipulations. The independence of the ε_k thus implies that

$$\mathsf{E}\int G_N^2(t)dt = \frac{1}{4}A_N^2 \sum_{k=1}^{N}(2k-N-1)^2 \mathsf{E}\int \varepsilon_k^2(t)dt = A_N\mathsf{E}\,||\varepsilon||^2 = \mathcal{O}(N^{-3}).$$

By Lemmas 2.2. and 2.3 of Horváth and Kokoszka (2012), relations (C.13) will follow from $||\hat{c}_\varepsilon - c_\varepsilon||_\mathcal{S} \xrightarrow{P} 0$, where the subscript $\mathcal{S}$ denotes the Hilbert–Schmidt norm. Proposition C.1 states that, in fact, $\mathsf{E}\,||\hat{c}_\varepsilon - c_\varepsilon||_\mathcal{S}^2 = \mathcal{O}(N^{-1})$. It thus extends a well–known result, e.g. Theorem 2.5. of Horváth and Kokoszka (2012), which states that

(C.14)
$$\mathsf{E}\int \left(\frac{1}{N}\sum_{i=1}^{N} \varepsilon_i(t)\varepsilon_i(s) - \mathsf{E}\left[\varepsilon(t)\varepsilon(s)\right] \right)^2 dtds = \mathcal{O}(N^{-1}).$$

The covariance function $\hat{c}_\varepsilon$ is defined in terms of the residuals $\hat{\varepsilon}_n$, cf. (2.4), (2.6). Estimation of the intercept and slope functions introduces many additional terms which are, however, all asymptotically negligible.

PROPOSITION C.1 *Suppose model (2.2) holds and $\mathsf{E}\,||\varepsilon||^4 < \infty$. Then the sample covariance function $\hat{c}_\varepsilon$, defined by (2.4) and (2.6), satisfies $\mathsf{E}\,||\hat{c}_\varepsilon - c_\varepsilon||_\mathcal{S}^2 = \mathcal{O}(N^{-1})$.*

PROOF: Observe that

$$\mathsf{E}\,||\hat{c}_\varepsilon - c_\varepsilon||_\mathcal{S}^2 = \mathsf{E}\iint \left(\hat{c}_\varepsilon(t,s) - c_\varepsilon(t,s) \right)^2 dtds$$

$$= \mathsf{E}\iint \left\{ \frac{1}{N}\sum_{k=1}^{N} \left(\hat{\varepsilon}_k(t)\hat{\varepsilon}_k(s) - \mathsf{E}\left[\varepsilon(t)\varepsilon(s)\right] \right) \right\}^2 dtds$$

$$= \frac{1}{N^2}\iint \sum_{n,m=1}^{N} \mathsf{E}\left[\left(\hat{\varepsilon}_n(t)\hat{\varepsilon}_n(s) - \mathsf{E}[\varepsilon(t)\varepsilon(s)] \right)\left(\hat{\varepsilon}_m(t)\hat{\varepsilon}_m(s) - \mathsf{E}[\varepsilon(t)\varepsilon(s)] \right) \right] dtds.$$

We will work with the following decomposition of the residuals:

$$\hat{\varepsilon}_n(t) = \varepsilon_n(t) + \left\{ \alpha(t) - \hat{\alpha}(t) \right\} + n\left\{ \beta(t) - \hat{\beta}(t) \right\}$$

$$= \varepsilon_n(t) + \frac{2}{N(N-1)}\sum_{k=1}^{N}(3k-2N-1)\varepsilon_k(t) + \frac{6n}{N(N+1)(N-1)}\sum_{k=1}^{N}(N+1-2k)\varepsilon_k(t)$$

$$= \varepsilon_n(t) + \frac{2}{N(N-1)(N+1)}\sum_{k=1}^{N}\left\{ 3(N+1-2n)k + (N+1)(3n-2N-1) \right\}\varepsilon_k(t)$$

$$= \varepsilon_n(t) + B_N\sum_{k=1}^{N}C_{N,n,k}\varepsilon_k(t),$$

where $B_N = \frac{2}{N(N-1)(N+1)}$ and $C_{N,n,k} = 3(N+1-2n)k + (N+1)(3n-2N-1)$.
Thus

$$\hat{\varepsilon}_n(t)\hat{\varepsilon}_n(s) - \mathsf{E}[\varepsilon(t)\varepsilon(s)] = \left(\varepsilon_n(t)\varepsilon_n(s) - \mathsf{E}[\varepsilon(t)\varepsilon(s)]\right) + B_N\varepsilon_n(t)\sum_{k=1}^{N} C_{N,n,k}\varepsilon_k(s)$$

$$+ B_N\varepsilon_n(s)\sum_{k=1}^{N} C_{N,n,k}\varepsilon_k(t) + B_N^2 \sum_{i,j=1}^{N} C_{N,n,i}C_{N,n,j}\varepsilon_i(t)\varepsilon_j(s),$$

and so the expression for $\mathsf{E}\,||\hat{c}_\varepsilon - c_\varepsilon||_\mathcal{S}^2$, i.e.

$$(\text{C.15}) \quad \frac{1}{N^2}\iint \sum_{n,m=1}^{N} \mathsf{E}\left[\left(\hat{\varepsilon}_n(t)\hat{\varepsilon}_n(s) - \mathsf{E}[\varepsilon(t)\varepsilon(s)]\right)\left(\hat{\varepsilon}_m(t)\hat{\varepsilon}_m(s) - \mathsf{E}[\varepsilon(t)\varepsilon(s)]\right)\right]\,dt\,ds$$

contains 16 terms. The leading term coincides with the left–hand side of (C.14):

$$\frac{1}{N^2}\iint \sum_{n,m=1}^{N} \mathsf{E}\left[\left(\varepsilon_n(t)\varepsilon_n(s) - \mathsf{E}[\varepsilon(t)\varepsilon(s)]\right)\left(\varepsilon_m(t)\varepsilon_m(s) - \mathsf{E}[\varepsilon(t)\varepsilon(s)]\right)\right]\,dt\,ds$$

$$=\mathsf{E}\iint \left(\frac{1}{N}\sum_{n=1}^{N} \varepsilon_n(t)\varepsilon_n(s) - \mathsf{E}[\varepsilon(t)\varepsilon(s)]\right)^2 dt\,ds = \mathcal{O}(N^{-1}).$$

The remaining 15 cross–terms are of the order $\mathcal{O}(N^{-2})$. We will display the verification for several of them, to illustrate the technique. It uses the independence of the ε_k and

the following bounds on sums involving the coefficients $C_{N,m,k}$:

$$\sum_{n=1}^{N} C_{N,n,n} = -N(N+1)(N-1),$$

$$\sum_{n,m=1}^{N} C_{N,n,n} C_{N,m,n}^2 = \mathcal{O}(N^8),$$

$$\sum_{n,m=1}^{N} \sum_{k \neq n} C_{N,n,n} C_{N,m,k}^2 = \mathcal{O}(N^9),$$

$$\sum_{n,m=1}^{N} \sum_{k \neq n} C_{N,n,k} C_{N,m,n} C_{N,m,k} = \mathcal{O}(N^9),$$

$$\sum_{n,m,k=1}^{N} C_{N,n,k}^2 C_{N,m,k}^2 = \mathcal{O}(N^{11}),$$

$$\sum_{n,m=1}^{N} \sum_{i \neq j}^{N} C_{N,n,i} C_{N,n,j} C_{N,m,i} C_{N,m,j} = \mathcal{O}(N^{12}),$$

$$\sum_{n,m=1}^{N} \sum_{i \neq j}^{N} C_{N,n,i}^2 C_{N,m,j}^2 = \mathcal{O}(N^{13}),$$

$$\sum_{n,m=1}^{N} \sum_{i \neq j}^{N} C_{N,n,i}^2 C_{N,m,j}^2 = \mathcal{O}(N^{13}).$$

For the cross–term involving the first and second terms, we have

$$\left| \frac{1}{N^2} \iint \sum_{n,m=1}^{N} \mathsf{E}\left[\left(\varepsilon_n(t)\varepsilon_n(s) - \mathsf{E}[\varepsilon(t)\varepsilon(s)] \right) B_N \varepsilon_m(t) \sum_{k=1}^{N} C_{N,m,k}\varepsilon_k(s) \right] dtds \right|$$

$$= \left| \frac{1}{N^2} \iint \sum_{n,m=1}^{N} \sum_{k=1}^{N} B_N C_{N,m,k} \mathsf{E}\left[\left(\varepsilon_n(t)\varepsilon_n(s) - \mathsf{E}[\varepsilon(t)\varepsilon(s)] \right) \varepsilon_m(t)\varepsilon_k(s) \right] dtds \right|$$

$$= \left| \frac{1}{N^2} \iint \sum_{n=1}^{N} B_N C_{N,n,n} \mathsf{E}\left[\left(\varepsilon_n(t)\varepsilon_n(s) - \mathsf{E}[\varepsilon(t)\varepsilon(s)] \right) \varepsilon_n(t)\varepsilon_n(s) \right] dtds \right|$$

$$= \frac{1}{N^2} \frac{2}{N(N-1)(N+1)} N(N+1)(N-1) \iint \left\{ \mathsf{E}\varepsilon^2(t)\varepsilon^2(s) - \left(\mathsf{E}[\varepsilon(t)\varepsilon(s)] \right)^2 \right\} dtds$$

$$\leq \frac{1}{N^2} \iint \mathsf{E}[\varepsilon^2(t)\varepsilon^2(s)] dtds = \frac{1}{N^2} \mathsf{E} \|\varepsilon\|^4.$$

Next, we turn to the cross–term involving second and fourth terms:

$$\left| \frac{1}{N^2} \iint \sum_{n,m=1}^{N} \mathsf{E}\left[B_N \varepsilon_n(t) \sum_{k=1}^{N} C_{N,n,k} \varepsilon_k(s) B_N^2 \sum_{i,j=1}^{N} C_{N,m,i} C_{N,m,j} \varepsilon_i(t) \varepsilon_j(s) \right] dtds \right|$$

$$\leq \frac{B_N^3}{N^2} \left| \iint \sum_{n,m=1}^{N} C_{N,n,n} C_{N,m,n}^2 \mathsf{E}[\varepsilon^2(t) \varepsilon^2(s)] dtds \right| \qquad (k = i = j = n)$$

$$+ \frac{B_N^3}{N^2} \left| \iint \sum_{n,m=1}^{N} \sum_{k \neq n} C_{N,n,n} C_{N,m,k}^2 \Big(\mathsf{E}[\varepsilon(t)\varepsilon(s)] \Big)^2 dtds \right| \qquad (k = n \neq i = j)$$

$$+ \frac{B_N^3}{N^2} \left| \iint \sum_{n,m=1}^{N} \sum_{k \neq n} C_{N,n,k} C_{N,m,n} C_{N,m,k} \mathsf{E}[\varepsilon^2(t)] \mathsf{E}[\varepsilon^2(s)] dtds \right| \qquad (i = n \neq j = k)$$

$$+ \frac{B_N^3}{N^2} \left| \iint \sum_{n,m=1}^{N} \sum_{k \neq n} C_{N,n,k} C_{N,m,k} C_{N,m,n} \Big(\mathsf{E}[\varepsilon(t)\varepsilon(s)] \Big)^2 dtds \right| \qquad (j = n \neq i = k)$$

$$= \mathcal{O}(N^{-11}) \mathcal{O}(N^8) \iint \mathsf{E}[\varepsilon^2(t)\varepsilon^2(s)] dtds + \mathcal{O}(N^{-11}) \mathcal{O}(N^9) \iint \Big(\mathsf{E}[\varepsilon(t)\varepsilon(s)] \Big)^2 dtds$$

$$+ \mathcal{O}(N^{-11}) \mathcal{O}(N^9) \iint \mathsf{E}[\varepsilon^2(t)] \mathsf{E}[\varepsilon^2(s)] dtds + \mathcal{O}(N^{-11}) \mathcal{O}(N^9) \iint \Big(\mathsf{E}[\varepsilon(t)\varepsilon(s)] \Big)^2 dtds$$

$$= \mathcal{O}(N^{-2}) \iint \left\{ \mathsf{E}[\varepsilon^2(t)] \mathsf{E}[\varepsilon^2(s)] + 2 \Big(\mathsf{E}[\varepsilon(t)\varepsilon(s)] \Big)^2 \right\} dtds$$

$$= \mathcal{O}(N^{-2}) \iint \mathsf{E}[\varepsilon^2(t)] \mathsf{E}[\varepsilon^2(s)] dtds = \mathcal{O}(N^{-2}) \Big(\mathsf{E} \, ||\varepsilon||^2 \Big)^2.$$

The bound for the most complex cross-term involving the fourth terms is established

as follows:

$$\left| \frac{1}{N^2} \iint \sum_{n,m=1}^{N} \mathsf{E}\left[B_N^2 \sum_{i_1,j_1=1}^{N} C_{N,n,i_1} C_{N,n,j_1} \varepsilon_{i_1}(t)\varepsilon_{j_1}(t) B_N^2 \sum_{i_2,j_2=1}^{N} C_{N,m,i_2} C_{N,m,j_2} \varepsilon_{i_2}(t)\varepsilon_{j_2}(t) \right] dtds \right|$$

$$\leq \frac{B_N^4}{N^2} \left| \iint \sum_{n,m,k=1}^{N} C_{N,n,k}^2 C_{N,m,k}^2 \mathsf{E}[\varepsilon^2(t)\varepsilon^2(s)] dtds \right| \qquad (k=i=j=n)$$

$$+ \frac{B_N^4}{N^2} \left| \iint \sum_{n,m=1}^{N} \sum_{i\neq j}^{N} C_{N,n,i} C_{N,n,j} C_{N,m,i} C_{N,m,j} \mathsf{E}[\varepsilon^2(t)]\mathsf{E}[\varepsilon^2(s)] dtds \right| \qquad (k=n\neq i=j)$$

$$+ \frac{B_N^4}{N^2} \left| \iint \sum_{n,m=1}^{N} \sum_{i\neq j}^{N} C_{N,n,i}^2 C_{N,m,j}^2 \Big(\mathsf{E}[\varepsilon(t)\varepsilon(s)]\Big)^2 dtds \right| \qquad (i=n\neq j=k)$$

$$+ \frac{B_N^4}{N^2} \left| \iint \sum_{n,m=1}^{N} \sum_{i\neq j}^{N} C_{N,n,i} C_{N,n,j} C_{N,m,i} C_{N,m,j} \Big(\mathsf{E}[\varepsilon(t)\varepsilon(s)]\Big)^2 dtds \right| \qquad (j=n\neq i=k)$$

$$= \mathcal{O}(N^{-14})\mathcal{O}(N^{11}) \iint \mathsf{E}[\varepsilon^2(t)\varepsilon^2(s)] dtds + \mathcal{O}(N^{-14})\mathcal{O}(N^{12}) \iint \mathsf{E}[\varepsilon^2(t)]\mathsf{E}[\varepsilon^2(s)] dtds$$

$$+ \mathcal{O}(N^{-14})\mathcal{O}(N^{12}) \iint \Big(\mathsf{E}[\varepsilon(t)\varepsilon(s)]\Big)^2 dtds + \mathcal{O}(N^{-14})\mathcal{O}(N^{12}) \iint \Big(\mathsf{E}[\varepsilon(t)\varepsilon(s)]\Big)^2 dtds$$

$$= \mathcal{O}(N^{-2}) \iint \left\{ \mathsf{E}[\varepsilon^2(t)]\mathsf{E}[\varepsilon^2(s)] + 2\Big(\mathsf{E}[\varepsilon(t)\varepsilon(s)]\Big)^2 \right\} dtds$$

$$= \mathcal{O}(N^{-2}) \iint \mathsf{E}[\varepsilon^2(t)]\mathsf{E}[\varepsilon^2(s)] dtds = \mathcal{O}(N^{-2})\Big(\mathsf{E}\,||\varepsilon||^2\Big)^2.$$

References

Berkes, I., Gabrys, R., Horváth, L. and Kokoszka, P. (2009). Detecting changes in the mean of functional observations. *Journal of the Royal Statistical Society (B)*, **71,** 927–946.

Berkes, I., Horváth, L. and Rice, G. (2013). Weak invariance principles for sums of dependent random functions. *Stochastic Processes and their Applications*, **123,** 385–403.

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

Brodsky, B. E. and Darkhovsky, B. S. (1993). *Nonparametric Methods in Change–Point Problems*. Kluwer.

Chen, J. and Gupta, A. K. (2011). *Parametric Statistical Change Point Analysis: With Applications to Genetics, Medicine, and Finance*. Birkhäuser.

Csörgő, M. and Horváth, L. (1997). *Limit Theorems in Change-Point Analysis*. Wiley.

Dierckx, G. and Teugels, J. L. (2010). Change point analysis of extreme values. *Environmetrics*, **21,** 661–686.

Efron, B. (1991). Regression percentiles using asymmetric squared loss. *Statistica Sinica*, **1,** 93–125.

Eilers, P. H. C. and Marx, B. D. (1996). Flexible smoothing with B–splines and penalties. *Statistical Science*, **11,** 89–121.

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

Guo, M. and Härdle, W. K. (2012). Simultaneous confidence bands for expectile functions. *AStA Advances in Statistical Analysis*, **96,** 517–541.

Guo, M., Zhou, L., Huang, J. Z. and Härdle, W. K. (2015). Functional data analysis of generalized regression quantiles. *Statistics and Computing*, **25,** 189–202.

Horváth, L. and Kokoszka, P. (2012). *Inference for Functional Data with Applications.* Springer.

Hsing, T. and Eubank, R. (2015). *Theoretical Foundations of Functional Data Analysis, with an Introduction to Linear Operators.* Wiley.

Kossin, J. P., Olander, T. L. and Knapp, K. R. (2013). Trend analysis with a new global record of tropical cyclone intensity. *Journal of Climate*, **26,** 9960–9976.

Newey, W. K. and Powell, J. L. (1987). Asymmetric least squares estimation and testing. *Econometrica*, **55,** 819–847.

Quantnet, Humboldt Universität (2015). Project: Dytec. `http://quantnet.hu-berlin.de`. Accessed: May 26, 2015.

Ramsay, J., Hooker, G. and Graves, S. (2009). *Functional Data Analysis with R and MATLAB.* Springer.

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

Rossi, G. De and Harwey, A. (2009). Quantiles, expectiles and splines. *Journal of Econometrics*, **152,** 179–185.

Schnabel, S. K. (2011). Expectile smoothing: new perspectives on asymmetric least squares. an application to life expectancy. Ph.D. Thesis. Utrecht University.

Schnabel, S. K. and Eilers, P. H. C. (2009). Optimal expectile smoothing. *Computational Statistics & Data Analysis*, **53,** 4168–4177.

Schnabel, S. K. and Eilers, P. H. C. (2013). Simultaneous estimation of quantile curves using quantile sheets. *AStA Advances in Statistical Analysis*, **97,** 77–87.

Smith, R. L. (1989). Extreme value analysis of environmental time series: An application to trend detection in ground–level ozone. *Statistical Science*, **4,** 367–377.

Sobotka, F., Kauermann, G., Schulze-Waltrup, L. and Kneib, T. (2013). On confidence intervals for semiparametric expectile regression. *Statistics and Computing*, **23,** 135148.

Sobotka, F., Schnabel, S., Schulz-Waltrup, L., Eilers, P., Kneib, T. and Kauermann, G. (2014). *R package: expectreg.* Version: 0.39, published: March 05, 2014.

Taylor, J. (2008). Estimating value at risk and expected shortfall using expectiles. *Journal of Financial Econometrics*, **6,** 231–252.

UNISYS, Unisys Weather Information Systems (2015). Data in atlantic and west pacific. `http://weather.unisys.com/hurricane/index.php`. Accessed: February 20, 2015.

SFB 649 Discussion Paper Series 2015

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

SFB 649 Discussion Paper Series 2015

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.