The Canadian Journal of Statistics
Vol. xx, No. yy, 2014, Pages 1–16
La revue canadienne de statistique

Detecting trends in time series of functional data: a study of Antarctic climate change

Ricardo Fraiman [1], Ana Justel [2], Regina Liu [3] and Pamela Llop [4]*

[1]*Centro de Matemática, Facultad de Ciencias, Universidad de la República, Uruguay*
[2]*Departamento de Matemáticas, Universidad Autónoma de Madrid, Spain*
[3]*Department of Statistics, Rutgers University, New Jersey, USA*
[4]*Facultad de Ingeniería Química (UNL) and Instituto de Matemática Aplicada del Litoral (UNL - CONICET), Argentina*

Key words and phrases: Antarctica; functional data; global warming; nonparametric test for functional data; trends in functional data; trends in time series.

MSC 2010: 62P12; secondary 62G10; 62M10

Abstract: The Spanish Antarctic Station *Juan Carlos I* has been registering surface air temperatures with the frequency of one reading per ten minutes since the austral summer 1987-88. Although this data set contains valuable information about the climate patterns in and around Antarctica, it has not been utilized in any existing climate studies thus far because of the concern of its substantial missing data caused by the difficulty in collecting data in the extreme winter weather conditions there. Such data sets do not fit the standard setting covered by the existing times series techniques. However, by treating the temperature readings for each summer as a function, the temperature data can be viewed as a time series of functional data. We introduce new notions of increasing trends for general time series of functional data based on the so-called *record functions*, and also develop useful nonparametric tests for such trends. Following our analysis, the data collected from *Juan Carlos I Station* exhibit an increasing trend in the Antarctic temperature.
The Canadian Journal of Statistics xx: 1–16; 2014 © 2014 Statistical Society of Canada
Résumé: Insérer votre résumé ici. We will supply a French abstract for those authors who can't prepare it themselves. *La revue canadienne de statistique* xx: 1–16; 2014 © 2014 Société statistique du Canada

1. INTRODUCTION

In recent years, there has been an intense focus on studies of climate change, for scientific and other purposes. Many climate studies of Antarctica have been hindered by poor quality data and yielded incomplete or inconclusive findings. But, overall, the Antarctic Peninsula is believed to be an area of recent rapid regional warming, with far more significant changes than those associated with global warming. For example, Vaughan et al. (2003) reports that surface air temperatures show warning trends reaching as high as 0.57^{o}C per decade for the region

* *Author to whom correspondence may be addressed.*
E-mail: pllop@santafe-conicet.gov.ar

around Faraday/Vernadsky Research Base (see the map provided in Section 2). Similarly, Steig et al. (2009) reports a significant warming that exceeds 0.1^{o}C per decade over the past 50 years in most of West Antarctica.

Many of the meteorological data sets used in the existing climate studies of Antarctica are from different sources and with varying durations. Some data sets span more than 50 years and, obviously, contain rich information. But all data sets appear to suffer varying degrees of data quality issues. Most common is the issue of substantial missing data due to the unavoidable instrument malfunctions. More specifically, meteorological data in Antarctica are generally collected only when the weather stations are inhabited by researchers, usually in summer. Even though many automatic weather stations (AWSs) have been installed in recent decades to register measurements continuously, winter registers are often missing due to malfunctioned sensors caused by the adverse Antarctic climate. Since malfunctioned sensors are repaired only when the station is re-inhibited in summer, large amount of missing data often incur in the data sets collected from AWSs. Such temporal discontinuities due to instrument malfunctions can often cause observations to be sparse or with uneven lengths within the duration of the study. Different studies have considered different data quality control criteria which sometimes exclude substantial data from the studies. For instance, out of 42 stations studied by Steig et al. (2009), Comiso (2000) used only the data from 21 stations, and Turner et al. (2005) further discarded the data from 2 stations. More often than not, data collected with much human and economic effort would not be fully utilized.

So far, almost all climate studies on Antarctica are based on long time series data of surface air temperatures from stations which operate all year-round. In fact, some of these studies adopt specific data quality control criteria so that the selected data can be used to form long time series data and apply the existing times series approaches for annual or monthly means. However, this approach, as pointed out by Trenberth (1984) or Santer et al. (2000), often yields significant bias in annual and monthly means because of the sparse observations or missing data in patches, and thus makes the means unsuitable for direct comparisons among different years. Moreover, such sparse observations or missing data in patches no longer allow the data set be considered as series of observations equally spaced in time. These type of data sets are not suitable for analysis using the usual time series methods such as exponential smoothing, moving averages or ARIMA modeling.

In this paper, by adapting a framework of functional data (or high resolution data) analysis, we can overcome the issue of unequally spaced data and hence utilize almost all of the observed data. Specifically, we assume that there is a continuous underlying temperature curve for each summer, although this curve is observed only at finite discrete time points due to the constraint of instrument capability. To be precise, the underlying temperature curves can be expressed as a sequence of real functions $\{x_t(s) : s \in [a,b]\}_{t=1}^{T}$, where $x_t(s)$ represents the temperature at time s in the time interval $[a,b]$ in the summer of the t-th year and t runs from 1 to T. In this set-up, the surface air temperatures from stations that operate only in summer can be considered as a time series of functional data, with one sample function observed each summer. Our goal is to define as well as detect possible trends, either up- or down-ward, in such a time series of functional data. These new ideas will be illustrated in Section 4 using the data set obtained from the Spanish Antarctic Station *Juan Carlos I (JCI)*.

While the research in functional data analysis has grown extensively in recent years, (see Ramsay and Silverman, 2002; Ferraty and Vieu, 2006), its development in the setting of time series remains relatively scarce. For example, the functional autoregressive model of order 1 was introduced in Bosq (1991) and then studied by several papers (see Bosq, 2000; Kargin and Onatski, 2008). However, these studies are mostly for the purpose of forecasting future values from past ones in the context of stationary processes. To the best of our knowledge, no definitions of trends for times series of functional data have been developed in the literature. In this paper, we introduce the notion of *record function (curve)* and we use it to develop new concepts of trend for series of functional data. Moreover, we also propose new nonparametric tests for such trends. The proposed new notions of trends and the nonparametric tests are shown to be useful for carrying out the trend analysis for the data set from the Spanish Antarctic Station *Juan Carlos I (JCI)* in this paper, but they are also important theoretical advances in functional time series in their own rights.

The paper is organized as follows. In Section 2, we describe in detail the data set obtained from the Spanish Antarctic Station *Juan Carlos I (JCI)*. In Section 3, we present our approach for the analysis of series of functional data, including the definitions of the new concepts of record curves and trends, and the proposal of the new "record based" tests for trends. These tests are then applied in Section 4 to analyzing the Antarctic temperature data collected from *JCI*. The test results conclude with an increasing trend in Antarctic temperature over the years. Finally, we provide some concluding remarks in Section 5.

2. DATA DESCRIPTION

The Spanish Antarctic Station *JCI* is located on the South West coast of Hurd Peninsula, Livingston Island (62ºS, 60ºW). It has been collecting surface air temperature since February 1988.

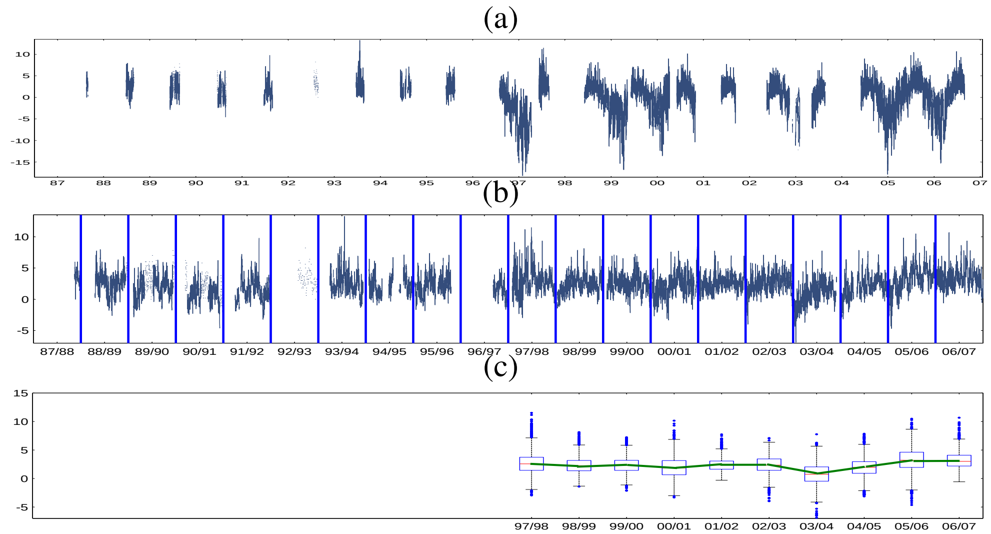

FIGURE 1: (a) Available hourly temperatures registered at *JCI* Station since 1988 to 2007. (b) Hourly temperatures at *JCI* Station for summer seasons only. (c): Boxplots and median trend for the last ten years.

The surface air temperatures are instantaneous values measured at 1.70 m above the ground with an automatic weather station at 10 minutes time intervals (Bañon et al., 2013). Since it is infrequent to have abrupt changes within the near-by 10-minute interval readings, we consider in this paper the hourly data which are the averages of all of the available 10 minute registers within each given hour from 1988 to 2007. These hourly data are plotted in Figure 1 (a), which clearly shows uneven lengths of available meteorological registers in different years, with sizable missing data in many years. Figure 1 (a) further shows that the meteorological registers succeeded in running all year long in only a few years, and that the data which are available across all 20 years are mostly during the summer seasons (i.e., during the time period December 1 to February 28). However, the plot of only the summer registers in Figure 1 (b) still shows considerable missing data, particularly at the beginning and the end of some summers which are generally the coldest days of the summer. Therefore, we further remove the first 26 and the last 18 days from consideration, which still leaves sufficient number of days in a continuous stretch in each summer. A quick visual examination by connecting the medians from the boxplots of such data sets for the last ten years (Figure 1 (c)) shows no indication of a positive or negative trend. Clearly, a more refined or powerful approach for detecting trends would be useful.

Focusing on the summer registers only, the hourly data of the t-th year can be represented by a vector of N points through the summer, such as $\{X_t(s_1), \ldots, X_t(s_N)\}$, or more generally, as $\{X_t(s_i) : i = 1, ..., N\}_{t=1}^{T}$ where s_i is the i-th time point (hour), $i = 1, ..., N$, and t is for the year t. For *JCI* data, the final data set for our analysis consists of $T = 17$ discretized curves in chronological order (from 1987/88 to 2006/07 except the years 1987/88, 1992/93 and 1996/97 which have more than two thirds of missing data, namely 93.47%, 94.87% and 69.70% respectively), and each is observed on a gird of $N = 1,103$ time points which are the hourly registers from December 27 0:00 to February 10 23:00.

3. TREND ANALYSIS OF TIME SERIES OF FUNCTIONAL DATA

In this section, we describe trends in a series of functional data based on "record functions" defined below in Section 3.1. Since our primary interest is one of determining whether there is an increasing trend in the series of Antarctic temperatures, we illustrate our approach for increasing trends throughout the paper. In this case, the "record functions" correspond to "maximum functions". The theoretical extensions for testing a decreasing trend or the two sided alternative should follow easily. For instance, the approach for decreasing trends can be modified in a straightforward manner by replacing the maximum functions with "minimum functions".

3.1. Definitions of increasing trends for functional data based on record functions

Let $\{x_t(s) : s \in [a,b]\}_{t=1}^T$ be a sequence of T real functions (curves) defined on the same compact interval $[a,b]$. In the *JCI* data setting, $x_t(s)$ represents the continuous temperature reading of the entire summer time interval $[a,b]$ of year t. For simplicity, $[a,b]$ is assumed to be $[0,1]$. One obvious increasing trend is to require that $\{x_t(s)\}_{t=1}^T$ satisfy $x_1(s) \le x_2(s) \le \ldots \le x_T(s)$ at every point $s \in [0,1]$. But the requirement of such a monotonicity at every point s on the entire time interval is clearly overly restrictive for defining trends, we resolve to pursuing increasing trends by examining *the proportion of time that each curve matches with the previous maxima or attains new maxima*. More specifically, we formulate notions of positive trend for functional data $\{x_t(s)\}_{t=1}^T$ by tracking sequentially in t the proportion of time s, during the observed time interval $[0,1]$, that each curve $x_t(s)$ has attained the maxima accumulated up to year t, namely among $\{x_1(s), \ldots, x_t(s)\}$. The rationale is that if each of the later curves in the sequence matches often the records set by all previous curves or continues to create new records, there is evidence to indicate a gradual increment or an increasing trend through the sequence. We state those definitions precisely using the following expressions. For each $t \in \{1, \ldots, T\}$, we denote by $r_t(\cdot)$ the *maximum function up to t-th curve*, where $r_t : C[0,1]^t \to C[0,1]$ and $r_t(s) \equiv \max\{x_1(s), \ldots, x_t(s)\}$ for $s \in [0,1]$. This maximum function up to t-th curve will be referred to as the *record function* up to year t. Then the proportion of the time that the curve $x_t(\cdot)$ attains (or overlaps with, or spends at) the maximum function up to the t-th curve (namely $r_t(\cdot)$) is $w_t \equiv \int_0^1 \mathcal{I}_{\{x_t(s)=r_t(s)\}} ds$, where $\mathcal{I}_A$ stands for the indicator function of set A. Similarly, we define the *minimum function up to t-th curve* as $\underline{r}_t(s) \equiv \min\{x_1(s), \ldots, x_t(s)\}$ for $s \in [0,1]$, and the proportion of time that the curve $x_t(\cdot)$ overlaps with $\underline{r}_t(\cdot)$ as $\underline{w}_t \equiv \int_0^1 \mathcal{I}_{\{x_t(s)=\underline{r}_t(s)\}} ds$.

3.1.1. Strong increasing trend in a sequence of functional data

Consider the setting of hypothesis testing where the null hypothesis H_0 is "*There does not exist a trend*". Under the null hypothesis H_0, the sequence of curves up to any t-th curve are all equally likely to achieve the maximum $r_t(s)$ at any time point s. This fact allows us to assume that the curves in the sequence are exchangeable, which thus implies that the expected value of $w_t = 1/t$ (and also $\underline{w}_t = 1/t$) for all t. If the sequence of curves progresses with the pattern of spending more than the expected share of time at the maximum curve, and less than the expected share of time at the minimum curve, it would suggest an increasing trend. This notion of increasing trend is formulated more precisely below.

Definition I. *(Strong Increasing Trend)* The sequence $\{x_t(s) : s \in [0,1]\}_{t=1}^T$ is said to have a Strong Increasing Trend if it satisfies the following conditions:

$$0 < w_2 - \frac{1}{2} < \cdots < w_T - \frac{1}{T}, \tag{1}$$

and

$$0 > \underline{w}_2 - \frac{1}{2} > \ldots > \underline{w}_T - \frac{1}{T}. \tag{2}$$

Because of the symmetry of the equations on (1) and (2), in what follows we will only deal with (1). The corresponding test for (2) can be obtained similarly. Finally, the Bonferroni method can be applied to perform both tests simultaneously.

For testing whether there is an increasing trend in a given sequence of functions, we first consider the alternative hypothesis H_a that there is an increasing trend according to Definition I. In other words, we examine whether the normalized fraction of time spent at the maximum function up to year t, $\{w_t - 1/t\}$ for $t = 2, \ldots, T$, is a strictly positive increasing sequence of t. This strictly positive increasing requirement is quite restrictive, and thus the naming of "Strong Increasing Trend" in the definition. It amounts to requiring each function in the sequence be a record with respect to all the previous years. We proceed to introduce positive trends with lesser degrees of restrictiveness below.

3.1.2. Weak increasing trends in a sequence of functional data

Definition II. *(Weak Increasing Trend)* The sequence $\{x_t(s) : s \in [0,1]\}_{t=1}^T$ is said to have a Weak Increasing Trend if it satisfies the following condition:

$$\min_{t \in \mathbb{T}} \left(w_t - \frac{1}{t} \right) > 0, \tag{3}$$

where $\mathbb{T}$ stands for the set $\{2, \ldots, T\}$. Recall that, if there is no trend in the sequence, we would expect to observe $w_t = 1/t$ for all t. Hence, Equation (3) implies that every t-th curve spends more time than expected at the maximum

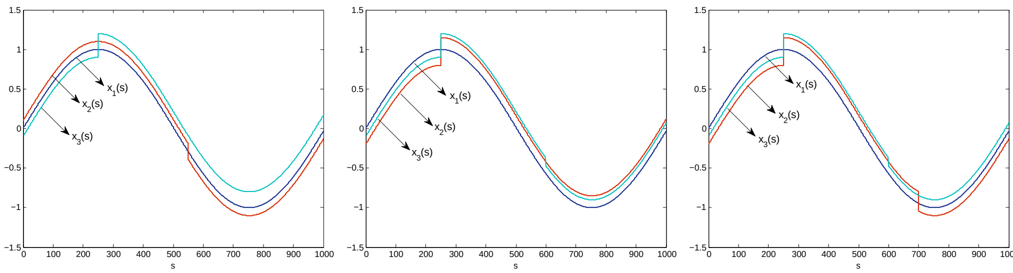

FIGURE 2: Left: Sequence with a Strong Increasing Trend. Middle: Sequence with a Weak Increasing Trend. Right: Sequence with a k_0-Weak Increasing Trend.

function up to year t. Clearly, this is a weaker notion of increasing trend than that of **Definition I**, since (1) implies (3).

Even though the condition $\min_{t\in\mathbb{T}}(w_t - 1/t) > 0$ in (3) is less stringent than that of (1), it can still be too restrictive for many practical situations, since requiring $(w_t - 1/t) > 0$ for all t would amount to requiring the curve of every single year in the sequence be a record function (in a weak sense) with respect to all its previous years. Instead, we may consider the following definition which is less restrictive and yet still captures meaningful patterns of increasing trends.

Definition III. *(k_0-Weak Increasing Trend)* The sequence $\{x_t(s) : s \in [0,1]\}_{t=1}^T$ is said to have a k_0-Weak Increasing Trend if there exists $k_0 \in \mathbb{T}$ such that

$$\bar{w}_{[k]} > 0, \qquad \forall\, k \geq k_0, \tag{4}$$

where $\bar{w}_t \equiv w_t - 1/t$ for $t = 2, \ldots, T$ whose corresponding order statistics are denoted by $\bar{w}_{[2]} \leq \bar{w}_{[3]} \leq \ldots \leq \bar{w}_{[T]}$.

Figure 2 provides some examples to illustrate the difference between the three notions of increasing trends. The left panel shows that $x_2(s)$ spends $w_2 = 55\%$ of the time at the maximum function up to year 2, whereas $x_3(s)$ spends $w_3 = 75\%$ up to year 3. This gives $0 < w_2 - 1/2 = 0.05 < w_3 - 1/3 = 0.417$, and thus satisfies condition (1). On the other hand, $x_2(s)$ spends $\underline{w}_2 = 45\%$ of the time at the minimum function up to year 2, whereas $x_3(s)$ spends $\underline{w}_3 = 25\%$ up to year 3. This gives $0 > \underline{w}_2 - 1/2 = -0.05 > \underline{w}_3 - 1/3 = -0.083$, and thus also satisfies condition (2). The middle panel shows that $x_2(s)$ spends $w_2 = 75\%$ of the time at the maximum function up to year 2, whereas $x_3(s)$ spends $w_3 = 35\%$ up to year 3. This gives $w_2 - 1/2 = 0.25$ and $w_3 - 1/3 = 0.016$, and thus satisfies (3) but not (1). Finally, the right panel shows a sequence of functions in which $x_2(s)$ spends $w_2 = 45\%$ of the time at the maximum function up to year 2, whereas $x_3(s)$ spends $w_3 = 65\%$ up to year 3. This gives $\bar{w}_{[2]} \equiv w_2 - 1/2 = -0.05$ and $\bar{w}_{[3]} \equiv w_3 - 1/3 = 0.3167$, and thus the sequence satisfies (4) with $k_0 = 3$ but not (3) nor (1).

3.2. Nonparametric tests for an increasing trends in a series of functional data

In this section, we develop nonparametric tests for the proposed increasing trends for a sequence of functional data. Let $\{X_t(s) : s \in [0,1]\}_{t=1}^{T}$ be a sample sequence of T functional data defined on a rich enough probability space $(\Omega, \mathcal{A}, P)$. For each $t \in \mathbb{T} \equiv \{2, \ldots, T\}$, let $R_t(s) \equiv \max\{X_1(s), \ldots, X_t(s)\}$ for $0 \leq s \leq 1$ be the sample maximum function up to year t. The time X_t spends at the maximum function up to year t is then $W_t \equiv \int_0^1 \mathcal{I}_{\{X_t(s)=R_t(s)\}} ds$. This quantity, with the standardized time scale $0 \leq s \leq 1$, is also the proportion of time that X_t matches with the maximum function from the first t years, whose expected value under P is $J_t \equiv E\left(W_t\right) = \int_0^1 P\left(X_t(s) = R_t(s)\right) ds$.

Under the null hypothesis of no trend in $\{X_t(s) : s \in [0,1]\}_{t=1}^{T}$, these curves are exchangeable and hence $J_t = 1/t$ for each t. For testing only the hypotheses whether a particular curve X_t is a record function up to year t, it would suffice to conduct an individual test for

$$H_{0,t} : J_t = 1/t \quad \text{vs.} \quad H_{1,t} : J_t > 1/t. \tag{5}$$

If the null hypothesis $H_{0,t}$ is rejected for (almost) all t with very small p-values, the Bonferroni device can be used to uphold simultaneously all the rejections. This approach of combining results from individual tests, however, is known to be inefficient. Therefore, we seek to develop a global test with an overall test statistic for assessing whether there exists an increasing trend. To this end, we may consider the following more complex hypotheses,

$$H_0 : \min_{t \in \mathbb{T}}(J_t - 1/t) = 0 \quad \text{vs.} \quad H_A : \min_{t \in \mathbb{T}}(J_t - 1/t) > 0. \tag{6}$$

Note that the alternative hypothesis H_A states that there is a *weak increasing trend* in the sense of Definition II. Let $\bar{\mathbf{J}} \equiv \left(\bar{J}_2, \bar{J}_3, \ldots, \bar{J}_T\right)$, where $\bar{J}_t \equiv (J_t - 1/t)$ for $t = 2, \ldots, T$ and whose corresponding order statistics are $\bar{J}_{[2]} \leq \bar{J}_{[3]} \leq \ldots \leq \bar{J}_{[T]}$. Obviously, under the null hypotheses H_0 in (6) $\bar{J}_{[2]}$ is expected to be zero, and a positive value of $\bar{J}_{[2]}$ would be evidence for rejecting H_0. Similarly, under the null hypotheses in (5), $\bar{\mathbf{J}}$ should be a vector of zeros, and $\bar{J}_{[k]} > 0$ for any $k > 1$ would be evidence in support of the alternative hypothesis in (5) with an increasing trend. This observation leads us to consider testing the following hypotheses

$$H_0 : \bar{J}_{[k]} = 0 \;\; \forall k \in \mathbb{T} \quad \text{vs.} \quad H_{A1} : \bar{J}_{[k_0]} > 0 \text{ for some } k_0 > 1. \tag{7}$$

To apply the tests above to real data, we need to replace the continuous curves by their discrete observations, since each curve is observed only at a finite number of time points. Specifically, if the t-th curve is observed only at N time points, $s_1, s_2, \ldots, s_N$, it can be expressed as the N-dimensional vector $(X_t(s_1), X_t(s_2), \ldots, X_t(s_N))$, for $t = 1, \ldots, T$. Note that $s_1, s_2, \ldots, s_N$ are not required to be equally spaced. We assume a simple nonparametric regression model for each data curve such that, for each $t = 1, \dots, T$,

$$X_t(s_j) = \mu_t(s_j) + e_t(s_j), \quad \text{for } \; j = 1, \dots, N, \tag{8}$$

where $\{e_t(s_j), \quad j = 1, \dots, N\}$ are i.i.d. random variables with mean zero, which represent random fluctuations around the mean value $\mu_t(s_j)$. Under the null hypothesis H_0, $\mu_t(s)$ is assumed to be independent of t, i.e., $\mu_t(s) = \mu(s)$ for all t. Therefore, the proportion of time that the t-th curve spends at the maximum curve can be simply accounted for by the random fluctuations $\{e_t(s_1), \dots, e_t(s_N)\}$. The maximum value and maximum error of the first t curves at the specific time point s_j are respectively

$$R_t(s_j) \equiv \max\{X_1(s_j), \dots, X_t(s_j)\} \;\text{ and }\; e^{(t)}(s_j) \equiv \max\{e_1(s_j), \dots, e_t(s_j)\}.$$

Note that the assumption that the error terms $\{e_t(s_j), \quad j = 1, \dots, N\}$ are i.i.d. implies that the maximum errors $\{e^{(t)}(s_j), \; j = 1, \dots, N\}$ are i.i.d. as well. The natural test statistic for testing (5) is the time that t-th curve spends at the maximum vector up to the t-th curve, namely $\tau_t \equiv \sum_{j=1}^{N} I_{\{X_t(s_j) = R_t(s_j)\}}$. For model (8), under the null hypothesis of (5), the distribution of τ_t is $B(N, p)$, a binomial distribution with N trials and each trial has the event probability p, where $p \equiv P(X_t(s_j) = R_t(s_j)) = P\left(e_t(s_j) = e^{(t)}(s_j)\right) = 1/t$, for any $t = 1, \dots, T$. This can be applied to performing hypothesis tests for the simplest setting (5) by noting that the proportion of time that $\{X_t(s_j), j = 1, \dots, N\}$ matches with the maximum vector up to the t-th year is $\hat{J}_t \equiv \tau_t / N$. The null hypothesis in (5) should be rejected if $\hat{J}_t - 1/t > c_\alpha$, where c_α is the critical value corresponding to a pre-fixed level α for the binomial test.

If we are to test the more complicated hypotheses in (6), we may consider the test statistic given by the first order statistic of the vector $\hat{\hat{\mathbf{J}}} \equiv \left(\hat{J}_2 - 1/2, \hat{J}_3 - 1/3, \dots, \hat{J}_T - 1/T\right)$, which is $\hat{\hat{J}}_{[2]} \equiv \min_{t \in \mathbb{T}} (\hat{J}_t - 1/t)$. Similarly, if we are to test the hypotheses in (7), we may consider the test statistic $\hat{\hat{J}}_{[k]}$ for some $k > 1$, where $\hat{\hat{J}}_{[k]}$ is the $(k-1)$-th order statistic of the elements in $\hat{\hat{\mathbf{J}}}$. For both cases, the null hypothesis should be rejected when the test statistic $\hat{\hat{J}}_{[2]}$ or $\hat{\hat{J}}_{[k]}$ is larger than a suitable critical value.

3.2.1. Calculation of critical values

We have described several tests for increasing trends so far, but the lack of independence of the random variables $\{(\hat{J}_t - 1/t) : t \in \mathbb{T}\}$ makes it difficult to calculate precise critical values for the tests, since the calculations involve a highly complex combinatorial problem. Nonetheless, we note that, under the null hypothesis and assuming that model (8) holds and also that there are no ties in observed data, the task of determining the critical values is equivalent in theory

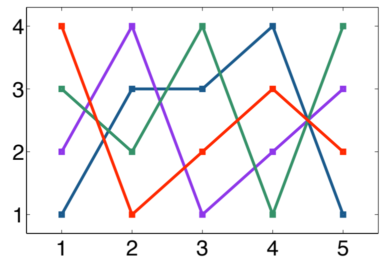

FIGURE 3: A realization of vectors of ranks on the lattice for four observed curves. The last observed curve (green) spends 40% of the time at the maximum.

to undertaking the procedure prescribed below. To find the time spent at the maximum curve (vector) by a given curve $X_t(\cdot)$, it suffices to obtain the rank of the value $X_t(s_j)$ among $\{X_1(s_j), \dots, X_T(s_j)\}$ at each s_j for all $j = 1, \dots, N$. This point is articulated in the following 3-step procedure.

Step 1: *Transform each functional to its corresponding rank vector* – At each time point s_j, we first rank all the observed values $\{X_1(s_j), \dots, X_T(s_j)\}$ and denote by λ_{tj} the rank associated with $X_t(s_j)$. For simplicity, we assume that there are no ties in the ranking. Repeat this procedure for all s_j, $j = 1, \dots, N$ and for all $t \in \{1, \dots, T\}$. Let $\Lambda_t = (\lambda_{t1}, \dots, \lambda_{tN}) \in \{1, \dots, T\}^N$ denote the vector of ranks corresponding to the t-th observed curve $(X_t(s_1), \dots, X_t(s_N))$.

Step 2: *Map functionals with permutations of ranks on a lattice* – We now detach ourselves from the observed sample functionals $\{X_t(s_1), \dots, X_t(s_N)\}_{t=1}^T$ and their probability spaces and, instead, work directly with their rank vectors as the trajectories on the lattice $\{1, \dots, T\}^N$, which is to be built up using the following algorithm:

- We build up Λ_1 by selecting a value for λ_{1j} at random from $\{1, \dots, T\}$ for each j, $j = 1, \dots, N$.
- We build up Λ_i $(i > 1)$ by selecting a value for λ_{ij} at random from $\{1, \dots, T\} \setminus \{\lambda_{1j}, \dots, \lambda_{(i-1)j}\}$ for each j, $j = 1, \dots, N$.
- We continue this way to build up all $\Lambda_1, \dots, \Lambda_T$ until the set of trajectories completely fill the lattice $\{1, \dots, T\}^N$ (Figure 3 shows a simple example of such a realization on the lattice for $T = 4$ curves on $N = 5$ time points.)

Step 3: *Calculate the values of test statistics from rank realizations* – For each rank realization $(\Lambda_1, \dots, \Lambda_T)$, we can compute the values of test statistics ($\hat{\hat{J}}_{[2]}$ or $\hat{\hat{J}}_{[k]}$).

By iterating this 3-step procedure sufficiently many times, we can obtain with reasonable accuracy an approximate distribution function of our test statistics and their desired quantiles. The key to obtain the distribution of our test statistics is to observe that, under the null hypothesis, each rank realization $(\Lambda_1, \dots, \Lambda_T)$ on the lattice $\{1, \dots, T\}^N$ obtained through the 3-step procedure above has the same probability $1/(T!)^N$.

In principle, we can obtain the exact distribution of the test statistic under the null hypothesis by considering all possible sets of trajectories $(\Lambda_1, \ldots, \Lambda_T)$ on the lattice $\{1, \ldots, T\}^N$ and calculating their corresponding values of the test statistic. Since all sets of trajectories have the same probability under the null hypothesis, we thus obtain the exact distribution of our statistic. Obviously, this task can become quite cumbersome or unrealistic when T and/or N are large, given that the set of trajectories is cardinal $(T!)^N$. In this case, we may resort to using Monte Carlo methods.

Remark 1 Note that the procedure described above can also be applied to testing the hypothesis of a strong increasing trend. In this case, for each set of trajectories, the only change needed in the procedure is to calculate the corresponding realization of the statistic $\hat{\hat{J}}_{strong} \equiv \#\{\hat{\hat{\mathbf{J}}} : 0 < \hat{J}_2 - 1/2 < \ldots < \hat{J}_T - 1/T\}$, instead of $\hat{\hat{J}}_{weak} \equiv \#\{\hat{\hat{\mathbf{J}}} : \hat{\hat{J}}_{[2]} > 0\}$. Clearly, $\hat{\hat{J}}_{strong} \leq \hat{\hat{J}}_{weak}$. Recall that $\hat{\hat{\mathbf{J}}} \equiv (\hat{J}_2 - 1/2, \ldots, \hat{J}_T - 1/T)$ and $\hat{\hat{J}}_{[2]} \equiv \min_{t=2,\ldots,T}(\hat{J}_t - 1/t)$.

Remark 2 Our proposed test procedure remains valid even if there are missing data in observed curves. In this case, some realignment of time points from individual curves is needed before implementing the 3-step procedure. Specifically, we consider only the largest subset of the original time points s_j for which all individual curve are completely observed. This modified sample of curves, though observed on a reduced number of time points, constitute a complete data set, and thus our test procedure is readily applicable. Although the number of time points N is effectively reduced to the cardinal of the subset of time points that provide a complete sample, the test results remain valid, since our approach does not require observations from each curve be equally spaced.

3.3. Multiple time series of functional data

Suppose that there are time series of functional data collected from more than one weather station, and that we are interested in synthesizing the information from different stations into a coherent combined analysis. For convenience, we discuss only the setting of two stations, as the extension to more stations is straightforward. In this case, we have two sets of observed data $X_1(s), \ldots, X_T(s)$, $Y_1(s), \ldots, Y_T(s)$ $0 \leq s \leq 1$, defined on a suitable probability space $(\Omega, \mathcal{A}, P)$.

For $t = 2, \ldots, T$, let $R_t^X(s) \equiv \max\{X_1(s), \ldots, X_t(s)\}$ and $R_t^Y(s) \equiv \max\{Y_1(s), \ldots, Y_t(s)\}$ be the two respective sample maximum function up to t-th curve. The time that X_t and Y_t spend jointly at their respective maximum function up to year t can be expressed as $W_t^{XY} \equiv \int_0^1 \mathcal{I}_{\{X_t(s)=R_t^X(s)\}}\mathcal{I}_{\{Y_t(s)=R_t^Y(s)\}}ds$, whose expected value is $J_t \equiv E\left(W_t^{XY}\right) = \int_0^1 P\left(X_t(s) = R_t^X(s), Y_t(s) = R_t^Y(s)\right) ds$. Under the null hypothesis of no trends, we again assume the following nonparametric model for the data from each station, for given $s \in [0,1]$ and $t = 1, \ldots, T$,

$$X_t(s) = \mu_t(s) + e_t(s) \quad \text{and} \quad Y_t(s) = \nu_t(s) + \eta_t(s),$$

where $e_t(s)$ and $\eta_t(s)$ are i.i.d. random variables with zero mean and represent the random fluctuations around the mean values $\mu_t(s)$ and $\nu_t(s)$ respectively. We further assume that $\mu_t(s)$ and $\nu_t(s)$ are independent of t. Therefore, under the null hypothesis, the time each curve spends at the maximum function can be explained entirely by the random fluctuations $e_t(s)$ and $\eta_t(s)$.

Assume that the curves are observed at a finite number of time points $s_1, \ldots, s_N$, not necessarily equally spaced. Similar to the one sample case, the test statistic is now the time that $X_t(\cdot)$ and $Y_t(\cdot)$ simultaneously spend at their respective maximum vectors up to the t-curve, namely, $\tau_t \equiv \sum_{j=1}^{N} I_{\{X_t(s_j)=R_t^X(s_j)\}} I_{\{Y_t(s_j)=R_t^Y(s_j)\}}$, as t progresses from 1 to T. Under the null hypothesis of no trends, τ_t follows the binomial distribution $B(N,p)$, with N trials and each with the probability of success $p = P(X_t(s_j) = R_t^X(s_j), Y_t(s_j) = R_t^Y(s_j)) = P\left(e_t(s_j) = e^{(t)}(s_j), \eta_t(s_j) = \eta^{(t)}(s_j)\right)$.

If we are to test the hypotheses whether the particular pair curve $(X_t(\cdot), Y_t(\cdot))$ matches with the pair of the maximum functions up to year t, it suffices to test for $H_0 : J_t = p \quad \text{vs.} \quad H_1 : J_t > p$. However, in this case it is not appropriate to simply assume that $p = 1/t^2$, and more involved parametric modeling of this parameter is needed. Following the modeling, we can estimate p from the observed data and then proceed to carry out the test in a similar way as in dealing with data from a single station. More precisely, we string together the observed data from both stations in a chronological ordered sequence, and treat the combined sequence as an observed function and then carry out the test procedures for one station as described in Sections 3.2. In some sense, the stringing together multiple sequences of functional data into one actually lengthens the domain of the underlying functional and thus enhances the suitability of the functional data approach proposed in this paper.

4. TREND ANALYSIS FOR ANTARCTIC TEMPERATURE DATA

We now apply our proposed approach in Section 3 to investigate possible trends in the surface air temperature data collected at *JCI* Station from 1988 to 2007. As described at length in Section 2, we have filtered out parts of the data whose data quality is in question but we have also retained as much data as possible to ensure that the selected final data set represents sufficient profile of the weather pattern through the years for a meaningful trend analysis. Specifically, the final data set consists of 17 discretized curves and each is observed on 1,103 time points which correspond the hourly registers in the middle ten weeks of the summer. This final data set corresponds to the setting of $\{X_t(s_j) : j = 1, 2, \cdots, 1103\}_{t=1}^{17}$.

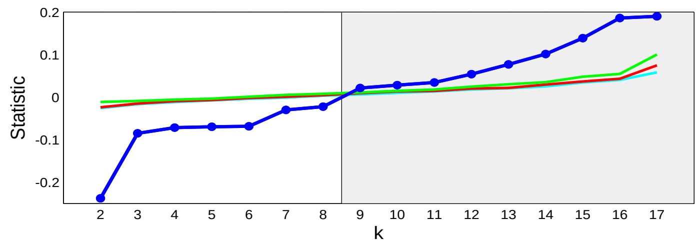

FIGURE 4: Value of $\hat{\hat{J}}_{[k]}$ (blue curve) together with respectively the 0.9 (red), 0.95 (violet) and 0.99 (green) critical values.

We perform the global tests for possibly weak increasing trends as defined in (6) and (7). The results are shown in Figure 4. At 0.10, 0.05 and 0.01 levels of significance, the global test fails to reject the null hypothesis in (6), as seen from the value of the test statistic corresponding to the plot of $\hat{\hat{J}}_{[k]}$ for $k = 2$ in Figure 4. This result also immediately implies that the global test fails to reject the null hypothesis that is associated with an alternative hypothesis of a strong increasing trend. On the other hand, Figure 4 does show that for all $\hat{\hat{J}}_{[k]}$ with $k > 8$ the test rejects the null hypothesis in (7) which is associated with a less restrictive alternative. This implies that at least half of the observed 17 summer temperature curves have spent a large amount of time at or near the record functions, and thus indicates a weak increasing trend over the years. This finding provides further support for the common belief of a rapid warming in the Antarctic Peninsula. This same finding is also claimed in IPCC (2014).

5. CONCLUDING REMARKS

Motivated by the need to explore the data set collected at Spanish Antarctic Station *JCI* between 1988 and 2007, we have introduced three notions of positive trend, *strong*, *weak* and k_0-*weak*, for a general framework of series of functional data and also developed nonparametric inference approaches for testing such trends. Under this framework, we consider the surface air temperatures registered at *JCI* for each summer as a function observed at finite grid points, and apply our approach to investigate this sequence of functions. Our findings suggest that there is indeed evidence of an increasing trend in the temperatures registered at *JCI* Station over the years. We believe that these findings can help better understand climate change in the Antarctic region.

The proposed approach is important in its own right as a theoretical advance in functional data analysis which should have broad applicability to problems in other domains. It broadens the scope of functional data analysis to include irregularly structured data sets which are observed unequally spaced and with possible missing data. The approach is applicable as long as: i) the available data have a common time span from one year to the next throughout the years in the study;

and ii) this common time span constitutes a sufficient portion of the year. This is the case for our *JCI* data set. Obviously, our approach can become unreliable if there are substantial missing data that are randomly scattered throughout and hence fail to provide a common time span for the direct comparison or ranking of observations from all years in the study.

In its original form, our *JCI* data can be viewed as high resolution data with missing values. When analyzing such a data set, there is inevitably a trade-off between the desire to keep the resolution sufficiently high in order to retain the data structure and to reduce the presence of missing data sufficiently to suit the analysis. Indeed, considering only yearly or monthly averages in our analysis can surely help “avoid’ missing data situations, but its over smoothing may risk corrupting important data structure. Also, over smoothing often results in non-negligible loss of efficiency and potential bias. A case in point is the yearly medians seen in Figure 1 (c). As for hourly averages, while they may still incur some loss of resolution, they do retain most of the data structure, since there are unlikely to be sharp phase shifts within a short time interval such as an hour. In fact, even working with hourly averages of our *JCI* data, we still have some missing data and, eventually, we can only focus on the time span for which the hourly averages are available for all years. As it turns out, this particular time span does cover more than 50 % of the central part of the summer temperature structure every year.

When dealing with high resolution temperature data, it is important to account for phase variation. Otherwise, one may obtain increasing portions of maximum temperatures without true warming condition. This seems less an issue in our case since temperatures in Antarctica tend to exhibit some regularity in their phase variations. However, in a setting of possible phase shift where the yearly curves are horizontally shifted versions of each other or a gradual shift in the onset of the seasons, the weak increasing trend may be subject to confounding with phase variation. In this case, the test results should be questioned, unless an additional test for phase shift can be devised to separate a phase shift from a real trend. Another issue on whether or not the phase variation is cyclical can also further complicate the study of trends. It would be worthwhile conducting a thorough study on this subject.

We provide two remarks on the nonparametric model (8) used in our approach. The assumption that the yearly functions are independent under the null hypothesis of no trends seems to be well accepted by most subject matter experts. This assumption is critical, as it implies in particular that, under the null hypothesis, the functional data are exchangeable for the years under consideration. On the other hand, the assumption of i.i.d. $\{e_t(s_j)\}_{j=1}^N$ in model (8), though also widely used in practice, may not necessarily hold in reality. It would be useful to consider modifications that reflect some practical correlation structure. We plan to investigate this further in the future by modeling the residuals with a functional AR(1) model. A closely related problem has been considered recently by Berkes et al. (2009), namely the problem of testing the i.i.d. functional observations $\{X_t(s), s \in [a,b]\}_{t=1}^T$ under the null assumption that the mean curve remains constant in t, i.e. $\mathbb{E}(X_t(s)) = \mu(s)$. Their test shows high power values against the alternative in which the data can be divided into several consecutive segments, with a constant mean within each segment but which differs from segment to segment. The simplest case of only two segments would correspond to a change point problem. Our approach is different, as we are interested in testing whether or not there is an increasing (or decreasing) trend among the underlying mean curves.

Finally, we note that although the methodology developed in this paper has been illustrated only with *JCI* temperature data, it is applicable to many other domains as well. For example, the methodology can be applied to the studies of consumption behaviors of restaurant goers where the data are often collected irregularly and only around holiday or low seasons.

Acknowledgements

The authors would like to thank Manuel Bañón of the Agencia Española de Meteorología for sharing the Juan Carlos I Station data and his expert knowledge of Antarctica meteorology. Ricardo Fraiman is supported by the Spanish Ministerio Ciencia e Innovación, grant MTM2007-66632. Ana Justel is supported by MINECO (Spain), grant CTM2011-28736/ANT. Regina Liu is supported by the National Science Foundation, grant DMS #1007683.

Bibliography

Bañon, M., Justel, A. Velázquez, D. and Quesada, A. (2013) Regional weather survey on Byers Peninsula, Livingston Island, South Shetland Islands, Antarctica. *Antarctic Science*, 25, 146–156.

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

Bosq, D. (1991) Modelization, nonparametric estimation and prediction for continuous time processes. *In Nonparametric functional estimation and related topics* (ed. G. Roussas), 509–530. NATO ASI Series, 335 Kluwer.

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

Comiso, J. C. (2000) Variability and trends in Antarctic surface temperatures from in situ and satellite infrared measurements. *Journal of Climate*, 13, 1674–1696.

Ferraty, F. and Vieu, P. (2006) Nonparametric modeling for functional data. *Springer-Verlag, New York.*

IPCC (2014) Climate Change 2013: The Physical Science Basis. Working Group I Contribution to the IPCC Fifth Assessment Report. *Cambridge University Press*.

Kargin, V. and Onatski, A. (2008). Curve forecasting by functional autoregression. *Journal of Multivariate Analysis*, 99, 2508–2526.

Ramsay, J. and Silverman, B. (2002) Functional Data Analysis. *Springer Series in Statistics, Springer, New York*.

Santer, B. D., Wigley, T. M. L., Boyle, J. S., Gaffen, D. J., Hnilo, J. J., Nychka, D., Parker, D. E. and Taylor, K. E. (2000) Statistical significance of trends and trend differences in layer–average temperature time series. *Journal of Geophysical Research*, 105, 7337–7356.

Steig, E. J., Schneider, D. P., Rutherford, S. D., Mann, M. E., Comiso, J. C. and Shindell, D. T. (2009) Warming of the Antarctic ice-sheet surface since the 1957 International Geophysical Year. *Nature*, 457, 459–463.

Trenberth, K. E. (1984) Signal versus noise in the southern oscillation. *Monthly Weather Review*, 112, 326–332.

Turner, J., Colwell, S., Marshall, G., Lachlan-Cope, T., Carleton, A., Jones. P., Lagun, V., Reid, P. and Iagovkina, S. (2005) Antarctic climate change during the last 50 years. *International Journal of Climatology*, 25, 279–294.

Vaughan, D. G., Marshall, G. J., Connolley, W. M., Parkinson, C., Mulvaney, R., Hodgson, D. A., King, J. C., Pudsey, C. J. and Turner, J. (2003) Recent rapid regional climate warming on the Antarctic Peninsula. *Climatic Change*, 60, 243–274.