Functional Median Polish

Ying SUN and Marc G. GENTON

This article proposes functional median polish, an extension of univariate median polish, for one-way and two-way functional analysis of variance (ANOVA). The functional median polish estimates the functional grand effect and functional main factor effects based on functional medians in an additive functional ANOVA model assuming no interaction among factors. A functional rank test is used to assess whether the functional main factor effects are significant. The robustness of the functional median polish is demonstrated by comparing its performance with the traditional functional ANOVA fitted by means under different outlier models in simulation studies. The functional median polish is illustrated on various applications in climate science, including one-way and two-way ANOVA when functional data are either curves or images. Specifically, Canadian temperature data, U.S. precipitation observations and outputs of global and regional climate models are considered, which can facilitate the research on the close link between local climate and the occurrence or severity of some diseases and other threats to human health.

Key Words: Analysis of variance; Climate models; Functional data; Health; Image data; Median polish; Robustness; Spatio-temporal data.

1. INTRODUCTION

Analysis of variance (ANOVA) is an important technique for analyzing the effect of categorical factors on a response. It decomposes the variability in the response variable among the different factors to determine which factors have significant effects on the response and how much of the variability in the response variable is attributable to each factor. A one-way ANOVA is used when the data are divided into groups according to only one factor. When more than one factor is present, a multi-way ANOVA is appropriate where both main effects and interactions among the factors may be estimated. For example, a two-way ANOVA places observations y_{ij} in a two-way table and the additive model without factor

Ying Sun is Postdoctoral Associate, Statistical and Applied Mathematical Sciences Institute, 19 T.W. Alexander Drive, Research Triangle Park, NC 27709-4006, USA (E-mail: *sunwards@samsi.info*). Marc G. Genton (✉) is Professor, Department of Statistics, Texas A&M University, College Station, TX 77843-3143, USA (E-mail: *genton@stat.tamu.edu*).

Journal of Agricultural, Biological, and Environmental Statistics, Volume 17, Number 3, Pages 354–376
DOI: 10.1007/s13253-012-0096-8

interactions is

$$y_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij}, \quad i = 1, \ldots, r, \; j = 1, \ldots, c, \tag{1.1}$$

where μ represents a grand effect, α_i and β_j denote the ith row effect and the jth column effect, respectively, ϵ_{ij} is the measurement error in the cell of the ith row and jth column, $r \geq 2$ is the number of rows and $c \geq 2$ is the number of columns.

The traditional ANOVA model is fitted by arithmetic means which is appropriate for data which have no outliers. If there are outlying observations in the data, using medians to estimate the terms is generally robust in the additive decomposition. Median polish is a technique for extracting row and column effects in a two-way table using medians rather than means. It is an iterative procedure; see Tukey (1970, 1977), Mosteller and Tukey (1977), Velleman and Hoaglin (1981), Hoaglin, Mosteller, and Tukey (1983, 1985), and Emerson and Hoaglin (1983) for details. The median polish has many applications. For instance, in spatial statistics, Cressie (1993) applied the median polish to gridded data as a robust method to remove any trend over a spatial domain. The median polish often produces a fit to the additive model that is close to being optimal in the least-absolute-residuals sense (Hoaglin, Mosteller, and Tukey 1983, p. 184). The iterative procedure lowers the L_1-norm of the residuals at each iteration. Usually it does not take very many iterations to converge but in general it does not converge to the least L_1-norm residuals. Fink (1988) provided an algorithm that converges in a finite number of steps for any real data and gives the least L_1-norm residuals.

Traditional ANOVA models fitted by means apply for univariate data. When functional data are observed, such as temporal curves or spatial surfaces, where observations are the entire functions rather than a string of numbers, functional ANOVA models are appropriate. They estimate the functional effects of categorical factors in order to determine how functions differ at different levels of these factors. Such questions arise in many fields, including meteorology, biology, medicine, and engineering. Kaufman and Sain (2010) developed a Bayesian framework for functional ANOVA modeling to estimate the effect of geographic regions on a Canadian temperature dataset and examined sources of variability in the output of climate models where the factors are the choice of regional climate models (RCMs) and the choice of their boundary conditions from global circulations models (GCMs). Based on their work, Sain, Nychka, and Mearns (2011) compared two different dynamic downscaling methods in climate model experiments and their projections of summer temperature and precipitation over North America.

Similar to the univariate case, the functional ANOVA models fitted by means are not resistant to outliers. In this article, we propose a robust functional ANOVA model fitted by medians, coined functional median polish, for functional data $y(x)$ where the index $x \in \mathscr{I}$. In classical functional data analysis, $x = t$ represents time and the index set $\mathscr{I}$ is an interval in $\mathbb{R}$. In image and surface analysis, $x = \mathbf{s}$ represents space and the index set $\mathscr{I}$ is a region in $\mathbb{R}^d$ with $d = 2$ although $d > 2$ could be considered as well. A further extension is to space-time functions where $x = (\mathbf{s}, t)$ and the index set $\mathscr{I}$ is a region in $\mathbb{R}^2 \times \mathbb{R}$. To perform functional ANOVA fitted by medians, we need to define the median for functional data. López-Pintado and Romo (2009) introduced the concept of modified band depth (MBD), a measure of the centrality of sample curves, defined through a graph-based approach. Suppose we observe n sample curves, $y_1(t), \dots, y_n(t)$. The sample band depth of a curve $y_i(t)$ is $BD_{n,J}(y_i) = \sum_{j=2}^{J} \binom{n}{j}^{-1} \sum_{1 \le i_1 < i_2 < \cdots < i_j \le n} I\{G(y_i) \subseteq B(y_{i_1}, \dots, y_{i_j})\}$, where $I\{\cdot\}$ denotes the indicator function, $B(y_{i_1}, \dots, y_{i_k}) = \{(t, z(t)) : t \in \mathscr{I}, \min_{r=1,\dots,k} y_{i_r}(t) \le z(t) \le \max_{r=1,\dots,k} y_{i_r}(t)\}$ is the band in $\mathbb{R}^2$ delimited by the curves $y_{i_1}, \dots, y_{i_k}$, $G(y_i) = \{(t, y_i(t)) : t \in \mathscr{I}\}$ and $2 \le J \le n$ is the number of curves determining a band. Since the order of curves induced by band depth is very stable in J, we use $J = 2$ to avoid computational issues. The MBD is a more flexible version which replaces the aforementioned indicator function by a function which measures the proportion of time that a curve $y_i(t)$ is in a band to prevent too many depth ties. The median curve is then defined as the one with the largest depth value. It is different from taking the median of $(y_1(t), \dots, y_n(t))$ at each $t \in \mathscr{I}$ pointwisely, which would lead to a curve that is not a real observation. If there are several curves with the largest depth value, then the median is the average of them. Consequently, the median curve is simply the average of two curves when $n = 2$ and $J = 2$. Using the modified band depth to order sample curves, Sun and Genton (2011, 2012) proposed functional boxplots and adjusted functional boxplots to visualize temporal curves. Furthermore, as an extension, they also proposed surface boxplots to visualize spatial surfaces where the band depth was extended to a volume-based depth in $\mathbb{R}^3$. Similarly, the median surface was defined as the one with the largest depth value. Having medians for both temporal curves and spatial surfaces defined by MBD, the functional median polish can be performed in an analogous fashion to the univariate median polish. The fast algorithm in Sun, Genton, and Nychka (2012) is used for MBD computation.

This article is organized as follows. Section 2 describes the functional median polish algorithm. Section 3 compares the traditional functional ANOVA fitted by means and the functional median polish by Monte Carlo simulation studies. Applications of the functional median polish on spatio-temporal datasets from climate science are reported in Section 4 and a discussion is provided in Section 5.

2. FUNCTIONAL MEDIAN POLISH ALGORITHM

For univariate data, the median polish procedure operates on the data in a two-way table by sweeping out column and row medians. The resulting estimation of the model (1.1) consists of a grand effect $\hat{\mu}$ common to all cells in the table, plus specific row effects $\hat{\alpha}_i$ and column effects $\hat{\beta}_j$. The procedure using medians to find grand, row and column effects operates iteratively on the data table, going through a number of row sweeping and column sweeping operations until no more changes occur in the row and column effects, or until changes are sufficiently small. The grand effect can depend on whether one starts with a row sweep or a column sweep. However, the order is often a matter of arbitrary choice, and the difference between the two solutions is generally unimportant for many practical purposes. For functional median polish, we describe the algorithm by starting with the row sweep.

2.1. ONE-WAY FUNCTIONAL MEDIAN POLISH

Suppose we observe functional data at each level of one categorical factor and we are interested in examining the effect of the factor, which we call the functional row effect. Then the observations can be decomposed as

$$y_{ik}(x) = \mu(x) + \alpha_i(x) + \epsilon_{ik}(x), \quad i = 1, \ldots, r,\ k = 1, \ldots, m_i, \tag{2.1}$$

where $x \in \mathscr{I}$, $r \geq 2$ is the number of rows and m_i is the number of replications in the ith row. Here $\mu(x)$ represents a functional grand effect and $\alpha_i(x)$ the ith functional row effect, with constraints, $\forall x \in \mathscr{I}$, that $\text{median}_i\{\alpha_i(x)\} = 0$ and $\text{median}_i\{\epsilon_{ik}(x)\} = 0$ for all k.

To fit this model by medians, we propose the following algorithm:

1. Compute the functional median of each row and record the functional value to the side of the row. Subtract the row functional median from each function in that particular row.
2. Compute the functional median of the row functional medians, and record the value as the functional grand effect. Subtract this functional grand effect from each of the row functional medians, and record the values as the functional row effect.
3. Repeat steps 1–2 and add the new functional grand effect and functional row effect to the current ones at each iteration until no changes occur with the row functional medians.

2.2. TWO-WAY FUNCTIONAL MEDIAN POLISH

Suppose we observe functional data at each combination of two categorical factors and we are interested in examining their effects, which we call the functional row or column effects, respectively. Then the observations can be decomposed as

$$y_{ijk}(x) = \mu(x) + \alpha_i(x) + \beta_j(x) + \epsilon_{ijk}(x), \quad i = 1, \ldots, r,\ j = 1, \ldots, c,\ k = 1, \ldots, m_{ij}, \tag{2.2}$$

where $x \in \mathscr{I}$, $r \geq 2$ is the number of rows, $c \geq 2$ is the number of columns and m_{ij} is the number of replications in the cell of the ith row and jth column. Here $\mu(x)$ represents a functional grand effect, $\alpha_i(x)$ the ith functional row effect, and $\beta_j(x)$ the jth functional column effect, with constraints, $\forall x \in \mathscr{I}$, that $\text{median}_i\{\alpha_i(x)\} = 0$, $\text{median}_j\{\beta_j(x)\} = 0$ and $\text{median}_i\{\epsilon_{ijk}(x)\} = \text{median}_j\{\epsilon_{ijk}(x)\} = 0$ for all k.

To fit this model by medians, we propose the following algorithm:

1. Compute the functional median of each row and record the functional value to the side of the row. Subtract the row functional median from each function in that particular row.
2. Compute the functional median of the row functional medians, and record the value as the functional grand effect. Subtract this functional grand effect from each of the row functional medians, and record the values as the functional row effect.

3. Compute the functional median of each column and record the functional value beneath the column. Subtract the column functional median from each function in that particular column.

4. Compute the functional median of the column functional medians, and add the value to the current functional grand effect. Subtract this functional grand effect from each of the column functional medians, and record the values as the functional column effect.

5. Repeat steps 1–4 and add the new functional grand effect, functional row and column effects to the current ones at each iteration until no changes occur with the row functional medians.

2.3. HYPOTHESIS TEST IN FUNCTIONAL MEDIAN POLISH

Based on the band depth, López-Pintado and Romo (2009) also proposed a rank test for functional data. In the functional ANOVA setting, it can be used to test if two populations Y and Y' have the same location parameter, i.e. the null hypothesis

$$H_0 : \mu_Y(x) = \mu_{Y'}(x), \quad \forall x \in \mathscr{I}. \tag{2.3}$$

If we reject H_0, the factor effect is significant in the functional ANOVA model. Suppose two functional samples $y_1(x), \dots, y_n(x)$ and $y'_1(x), \dots, y'_m(x)$ are observed and let $z_1(x), \dots, z_r(x)$ be the reference sample from one of the two populations with $r \geq \max(n, m)$. To find the position of y_i $(i = 1, \dots, n)$ or y'_j $(j = 1, \dots, m)$ with respect to the reference sample, define

$$R(y_i) = \frac{1}{n} \sum_{k=1}^{n} I\{\mathrm{MBD}(z_k) \leq \mathrm{MBD}(y_i)\},$$

$$R(y'_j) = \frac{1}{m} \sum_{k=1}^{m} I\{\mathrm{MBD}(z_k) \leq \mathrm{MBD}(y'_j)\}.$$

Then order these values from the smallest to the largest giving them the ranks from 1 to $n+m$. Let $W = \sum_{j=1}^{m} \mathrm{rank}\{R(y'_j)\}$. Then the distribution of W under H_0 is the distribution of the sum of m numbers that are randomly chosen from $\{1, 2, \dots, n+m\}$. The critical values are determined by simulations.

This rank test is nonparametric and robust which is different from the two-sample functional t-test (Ramsay, Hooker, and Graves 2009) for H_0 defined in (2.3). The functional t-test statistic is defined as

$$\max_x \frac{|\bar{Y}(x) - \bar{Y}'(x)|}{\sqrt{\frac{1}{n} s^2_{Y(x)} + \frac{1}{m} s^2_{Y'(x)}}}, \tag{2.4}$$

where $\bar{Y}(x) = \frac{1}{n}\sum_{i=1}^{n} Y_i(x)$, $\bar{Y}'(x) = \frac{1}{m}\sum_{i=1}^{m} Y'_i(x)$, $s^2_{Y(x)} = \frac{1}{n-1}\sum_{i=1}^{n}\{Y_i(x) - \bar{Y}(x)\}^2$ and $s^2_{Y'(x)} = \frac{1}{m-1}\sum_{i=1}^{m}\{Y'_i(x) - \bar{Y}'(x)\}^2$. The critical values are determined using a permutation test by randomly shuffling the labels of the two samples and calculating the maximum in (2.4) with the new labels. This is repeated by simulations and a null distribution is constructed.

3. MONTE CARLO SIMULATIONS

The functional ANOVA model in (2.2) can be fitted by means with

$$\hat{\mu}(x) = \frac{1}{\sum_{i,j} m_{ij}} \sum_{i=1}^{r} \sum_{j=1}^{c} \sum_{k=1}^{m_{ij}} y_{ijk}(x),$$

$$\hat{\alpha}_i(x) = \frac{1}{\sum_{j} m_{ij}} \sum_{j=1}^{c} \sum_{k=1}^{m_{ij}} y_{ijk}(x) - \hat{\mu}(x),$$

$$\hat{\beta}_j(x) = \frac{1}{\sum_{i} m_{ij}} \sum_{i=1}^{r} \sum_{k=1}^{m_{ij}} y_{ijk}(x) - \hat{\mu}(x).$$

To study the robustness of the fit, we compare the performance of the functional median polish with the mean version of functional ANOVA under different models.

We generate data from a true model in (2.2) with $x = t$, $r = 2$, $c = 3$, and $m = 100$ curves in each cell at $p = 50$ time points. Let $\mu(t) = 4t$, $\alpha_i(t) = i\alpha(t)$, $\beta_j(t) = j\beta(t)$, where $\alpha(t) = 5(t - 0.5)^2$ has a quadratic form and $\beta(t) = 4$, a constant form, for $i = 1, 2$, $j = 1, 2, 3$ and $t \in [0, 1]$. We introduce outliers through a stochastic Gaussian process $\epsilon_{ijk}(t)$ from different models with similar model structures as studied by Sun and Genton (2011). Model details are described as follows:

1. Model 1 is a basic one without contamination: $\epsilon_{ijk}(t) = e_{ijk}(t)$, where $e_{ijk}(t)$ is a stochastic Gaussian process with zero mean and covariance function $\gamma(t_1, t_2) = \exp\{-|t_2 - t_1|\}$.

2. Model 2 includes an asymmetric contamination: $\epsilon_{ijk}(t) = e_{ijk}(t) + c_{ijk}K$, where c_{ijk} is 1 with probability q_{ij} and 0 with probability $1 - q_{ij}$, $K = 20$ is a contamination size constant and the outlier probability q_{ij} is different for each cell according to the matrix:

$$\mathbf{Q} = \begin{pmatrix} 0.10 & 0.20 & 0.15 \\ 0.05 & 0.10 & 0.10 \end{pmatrix}.$$

3. Model 3 is partially contaminated: $\epsilon_{ijk}(t) = e_{ijk}(t) + c_{ijk}K$, if $t \geq T_{ijk}$ and $\epsilon_{ijk}(t) = e_{ijk}(t)$, if $t < T_{ijk}$, where T_{ijk} is a random number generated from a uniform distribution on $[0, 1]$.

For 1,000 replications, we use the functional boxplot introduced by Sun and Genton (2011) to summarize the estimation of the functional effects, $\hat{\mu}(t)$, $\hat{\alpha}_i(t)$ and $\hat{\beta}_j(t)$, for functional median polish and the mean version of functional ANOVA, shown in Figures 1, 2, 3 for models 1, 2, 3, respectively.

Under the three models, functional median polish gives unbiased estimation of the functional grand, row and column effects, whereas the mean version of functional ANOVA is clearly affected by outliers in models 2 and 3. As can be seen in Figure 2 for the mean version of functional ANOVA, the median curves of the functional grand, row and column effects all shift away from the truth due to the magnitude outliers in model 2. In Figure 3

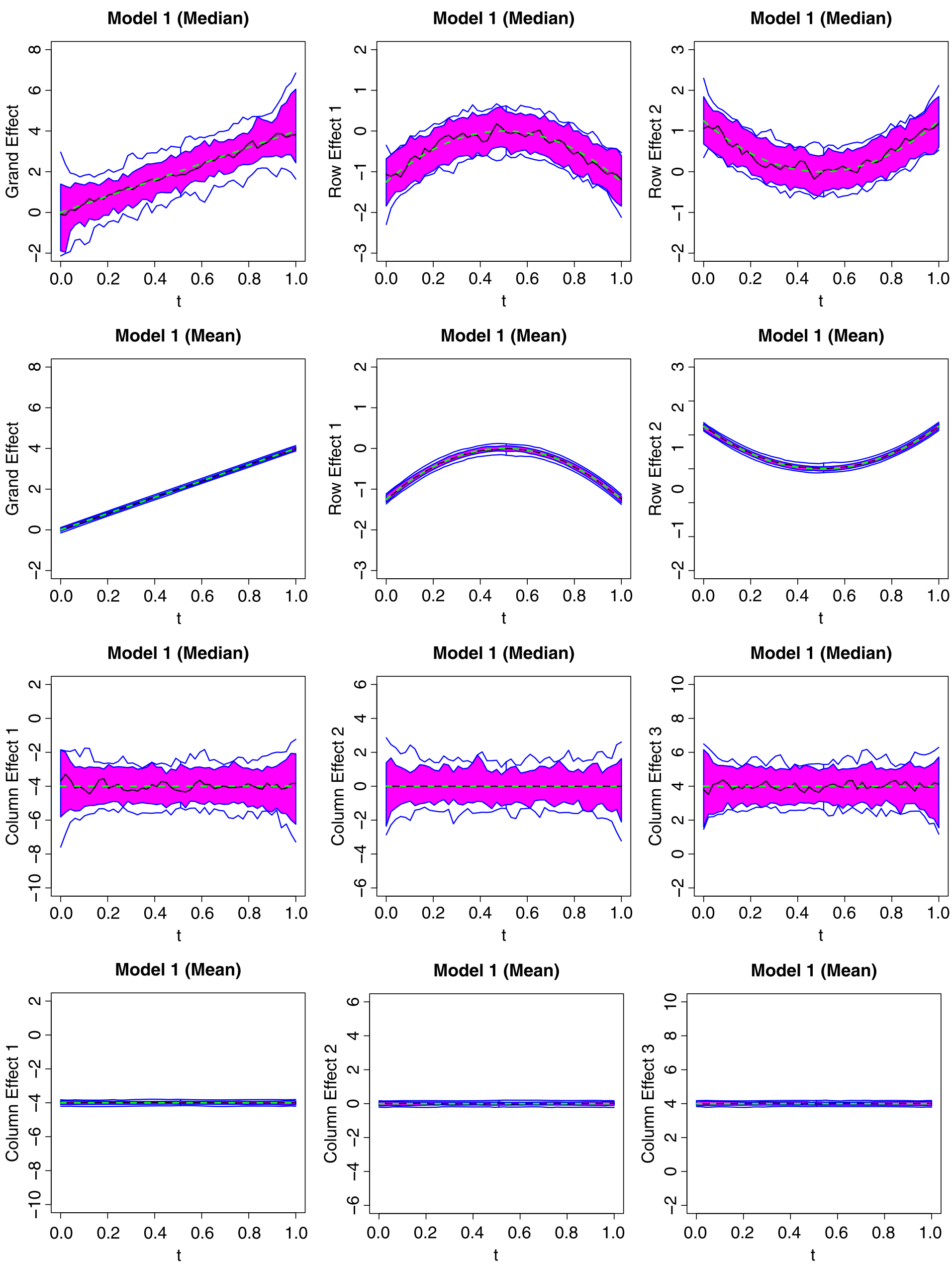

Figure 1. Functional boxplots of the functional effects estimation for median polish and the mean version of functional ANOVA for model 1 with blue curves denoting envelopes, and a black curve representing the median curve. True values are represented by dashed green curves. (Color figure online.)

for the mean version of functional ANOVA, because of the shape of outliers in model 3, the median curves of the functional grand, row and column effects all depart from the truth with an increasing trend in magnitude through time. The spread is increasing through time as well. Moreover, the substantially biased mean version of functional ANOVA in the pres-

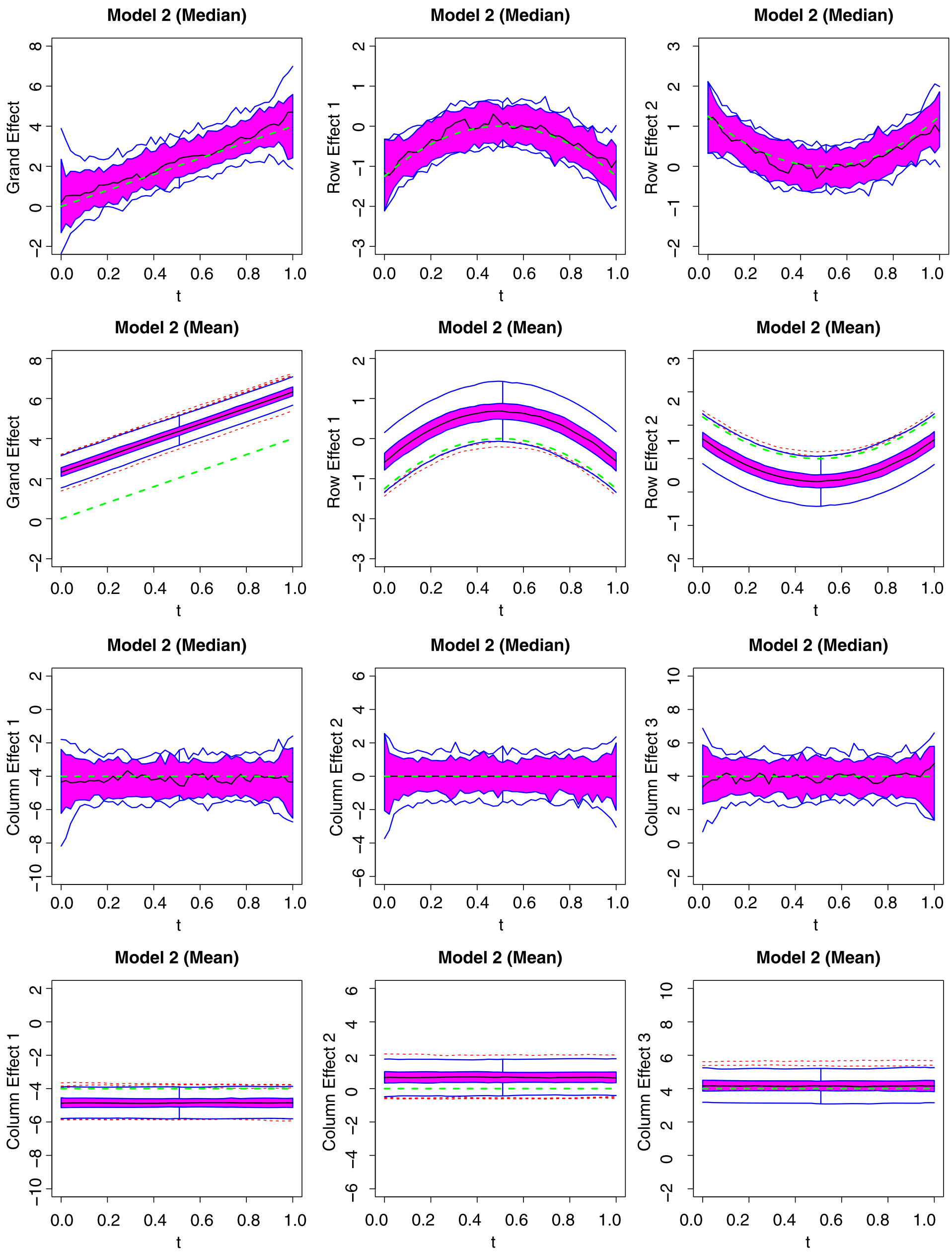

Figure 2. Functional boxplots of the functional effects estimation for median polish and the mean version of functional ANOVA for model 2. The red dashed curves are the outlier candidates detected by the 1.5 times the 50 % central region rule. True values are represented by dashed green curves. (Color figure online.)

ence of outliers has yet a considerably lower variance which makes the wrong inferences more misleading.

However, for model 1 in Figure 1 where no outlier is present, the mean version of functional ANOVA is better with unbiased estimates and small variability while the functional

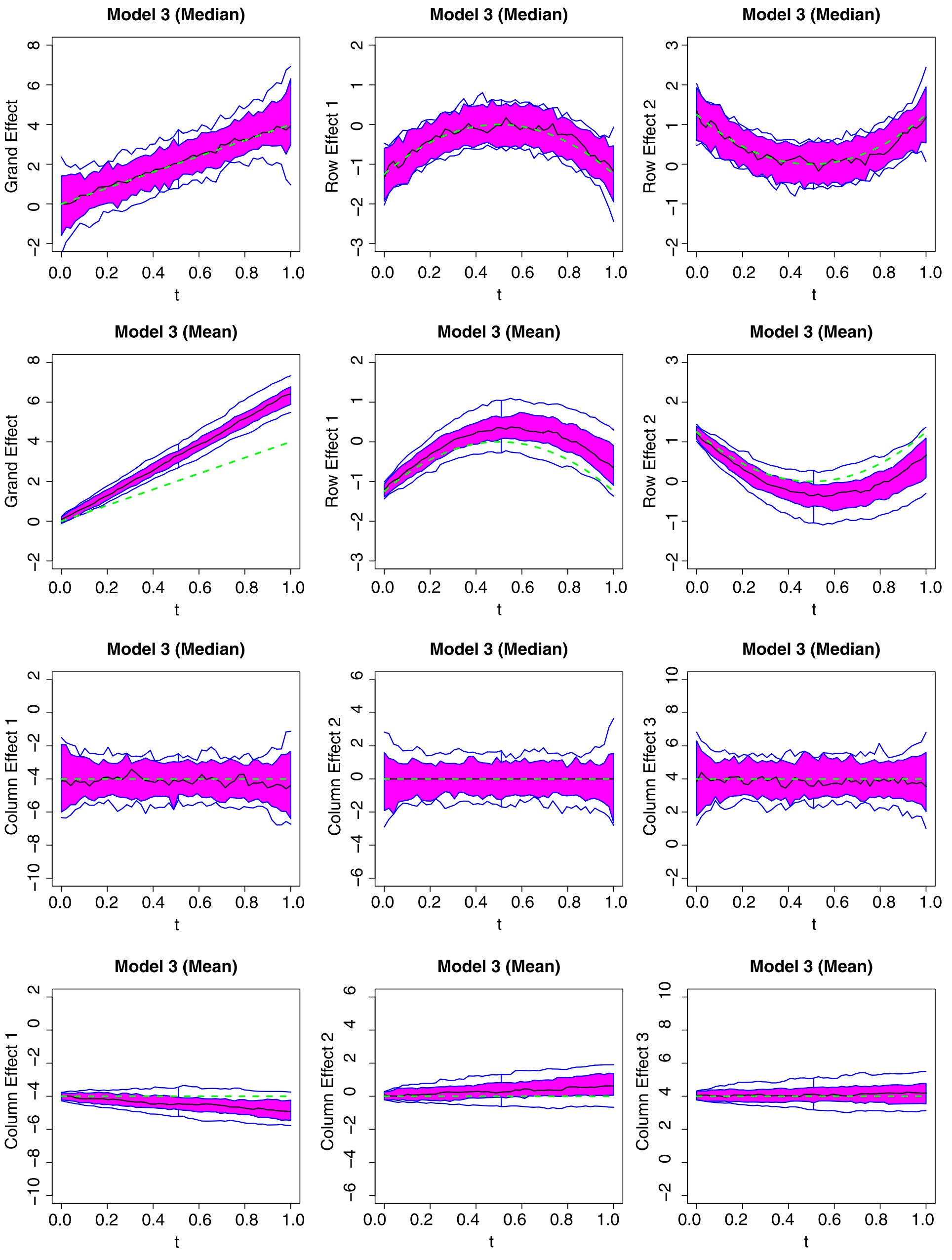

Figure 3. Functional boxplots of the functional effects estimation for median polish and the mean version of functional ANOVA for model 3. True values are represented by dashed green curves. (Color figure online.)

median polish is unbiased but has a relatively larger variability. Just like other median related method, the functional median polish is more variable than the mean version for normal data but is robust and performs much better for contaminated data. Therefore, the functional median polish is a robust method to fit the functional ANOVA model and a better choice when outliers are present.

4. APPLICATIONS

Researchers have found that there is a close link between local climate and the occurrence or severity of some diseases and other threats to human health. For example, extreme heat contributes directly to the mortality rate of cardiovascular and respiratory diseases, high temperatures raise the levels of pollutants in the air, and increasingly variable rainfall patterns are likely to affect the supply of fresh water. Moreover, climate models are used to study the future climate by running under different scenarios and project climate changes in the intensity and range of climate sensitive diseases. For instance, a continued increase of atmospheric green house gases scenario is likely to generate more frequent, more intensive or longer-lasting heat waves threatening human health. A better understanding of the climate characteristics and the model structures can facilitate a potential health risk study as well as inform the policy related analysis.

In this section, we present four applications in climate science and comparisons between the functional median polish and the mean version of functional ANOVA. The Canadian temperature in Section 4.1 and the U.S. precipitation in Section 4.2 are examples of one-way functional median polish where temporal curves are observed at weather stations from different climatic regions and we aim at estimating the functional grand and region effects. The applications in Section 4.3 are examples of one-way functional median polish where observations are spatio-temporal outputs from climate models and we treat the data either as temporal curves at different locations or spatial surfaces at different time points. Therefore, the functional median polish is applied to both curves and images. An example of a two-way functional median polish is provided in Section 4.4.

4.1. CANADIAN TEMPERATURE

We consider the Canadian weather data introduced by Ramsay and Silverman (2005) and studied by Kaufman and Sain (2010) with different mean-based functional ANOVA models. The data shown in Figure 4 consist of monthly average temperatures for 35 Canadian weather stations from four climatic regions: Atlantic, Continental, Pacific and Arctic. We perform one-way functional median polish as well as the traditional mean version of functional ANOVA to estimate the grand and the region effects shown in Figure 5 for the model in (2.1). The functional grand effect shows the overall temperature trend through seasons of the year, and the functional region effects represent the different temperature

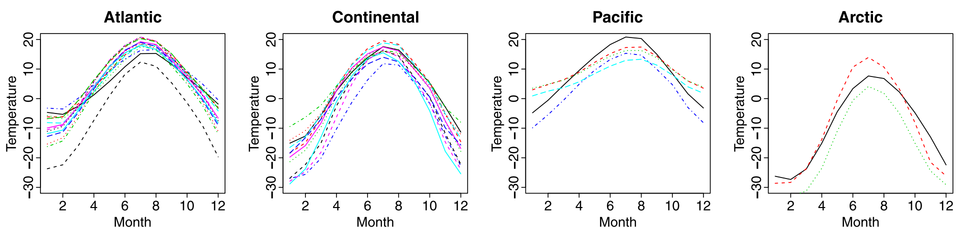

Figure 4. Averaged monthly temperature (°C) data at 35 Canadian weather stations from four climatic regions.

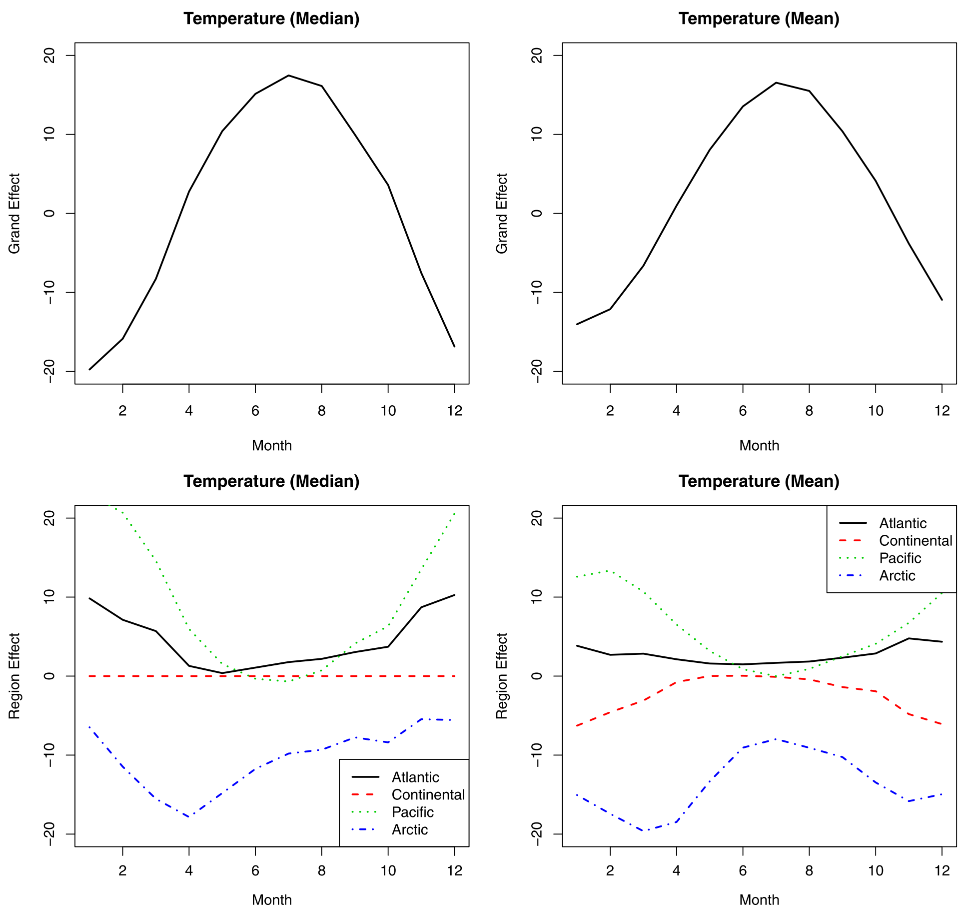

Figure 5. Canadian temperature (°C): estimation of the functional grand and region effects for the functional median polish and the traditional mean version of functional ANOVA.

features of the four climatic regions. For example, the Continental region has the same pattern as the functional grand effect, the Atlantic region tends to be warmer overall, the Pacific region is warmer, particularly so during the winter months and the Arctic region is always colder, especially in the winter and spring.

Figure 5 shows that the Atlantic and Arctic effects in the mean version of functional ANOVA are smaller than those in the functional median polish. By looking at the original curves in Figure 4, we can see that one curve for Atlantic and one for Arctic are below the majority, which affects the estimation in the mean version. Similarly, in the mean version of functional ANOVA, the Continental and Pacific effects are smaller than those in the functional median polish at the beginning and the end of the year, which can be explained by the four original curves for Continental and two for Pacific below the majority at the beginning and the end of the year. Therefore, the functional median polish provides a more robust estimation of the various effects.

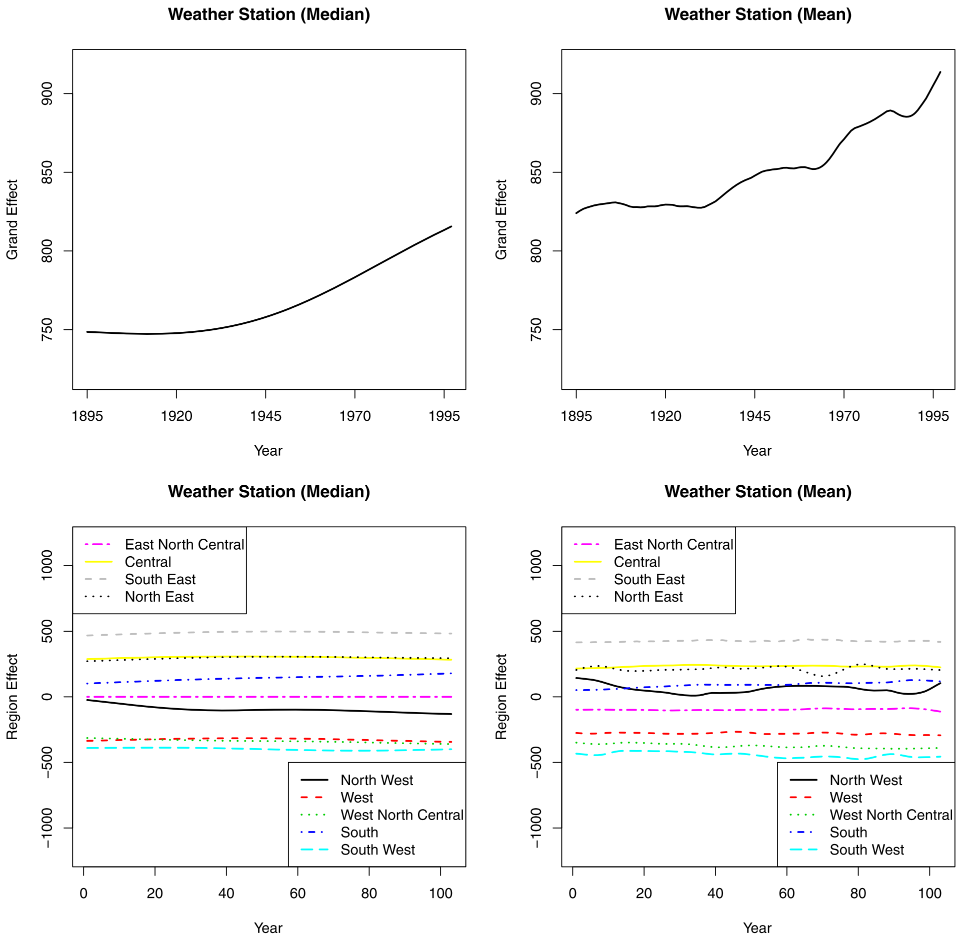

Figure 6. U.S. precipitation (mm/year): estimation of the functional grand and region effects for the functional median polish and the traditional mean version of functional ANOVA.

4.2. U.S. PRECIPITATION

Sun and Genton (2011) proposed functional boxplots to visualize and compare the annual precipitation for nine different climatic regions in the U.S. defined by the National Climatic Data Center (NCDC). The observed annual total precipitation data for the coterminous U.S. from 1895 to 1997 at 11,918 weather stations were smoothed by a spline smoothing approach. We consider the same dataset and the one-way functional ANOVA of the smoothed curves is shown in Figure 6 for the functional median polish and the traditional mean version of functional ANOVA. The functional grand effect shows that there is an increasing overall trend of the precipitation in the U.S. and the functional region effects represent the different annual precipitation levels in different climatic regions with respect to the overall pattern. For example, the South East, Central and North East have higher precipitation while the South West, West and West North Central have drier years.

Comparing the functional median polish with the mean version of functional ANOVA, we can see that the grand effect in the mean version is much higher than that in the functional median polish due to the extreme large outliers detected by the functional boxplots in Sun and Genton (2012). The region effects are also affected by the outliers in the mean version, especially for the North West which has higher precipitation than the overall pattern whereas the functional median polish actually shows the opposite. Since the mean version is not resistant to outliers, the region effect is likely (based on simulation results in Section 3) to be overestimated because of the outlying locations with high precipitation along the west coast (Sun and Genton 2012).

4.3. Global Climate Model

A global climate model (GCM), also referred to as general circulation model, represents physical processes in the atmosphere and ocean. It uses complex computer programs to simulate the response of the global climate system. Sun and Genton (2012) proposed the adjusted functional boxplot to visualize the U.S. precipitation data from weather stations studied in Section 4.2 and the data generated from the National Center for Atmospheric Research-Community Climate System Model Version 3.0 under A2 scenario which considers a continued increase of atmospheric green house gases and the associated warming throughout the 21st century; see IPCC (2000), Collins et al. (2006), Ammann et al. (2010) and references therein. The weather station data were matched with GCM data by averaging the observations within each cell for the coterminous U.S., the two functional boxplots were then compared. The functional median polish and the traditional mean version of functional ANOVA for this example estimate the functional grand effect and data source (weather station and GCM) effect by treating the output either as temporal curves at different locations (Figure 7), or images over the spatial region at different time points (Figure 8).

As shown in Figure 7, the effects in the functional median polish have more annual oscillations than those in the mean version, and the GCM tends to have higher precipitation than the weather stations for most of the years. Moreover, the functional median polish shows more clearly that the GCM has lower precipitation for the last several years. Both the rank test and the two-sample functional t-test reject H_0 in (2.3) at the 5 % significance level, indicating that the GCM or weather station effect is significant. In Figure 8, from the functional median polish, we can see that the GCM has higher precipitation than the weather stations in some locations in the North but lower precipitation in the South and South East. The effect estimation from the mean version is similar to the functional median polish but smaller in magnitude. The rank test fails to reject H_0 for images at the 5 % significance level while the functional t-test rejects it.

The GCM also simulates precipitation for the future. We apply the functional median polish to estimate the segment effect for GCM past (1900–1999) and GCM future (2000–2099), and estimate the functional grand effect and segment effect by treating the output either as temporal curves at different locations (Figure 9), or images over the spatial region at different time points (Figure 10) for the functional median polish and the traditional mean version of functional ANOVA. As shown in Figure 9, the effects in the

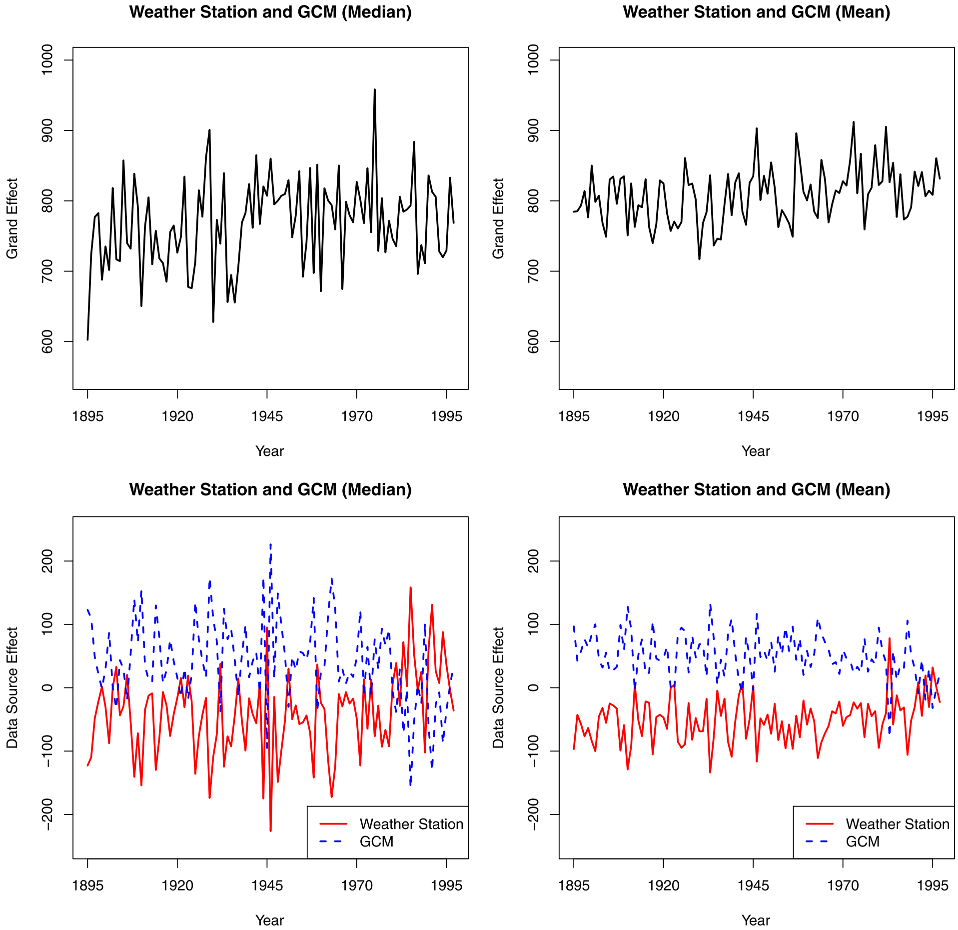

Figure 7. Weather station and GCM precipitation (mm/year) curves: estimation of the functional grand and data source effects for the functional median polish and the traditional mean version of functional ANOVA.

functional median polish have more annual oscillations than those in the mean version, and the functional median polish shows more clearly that the GCM future runs have higher precipitation than the past runs for the last 20 years. In Figure 10, the GCM future runs have higher precipitation than the past in the South, but lower precipitation in the North. Similar to Figure 9, the GCM past and future effects in the mean version are less significant. For both curves and images, the rank test fails to reject H_0 in (2.3) at the 5 % significance level which does not provide evidence that the GCM past or future effect is significant. In contrast, the functional t-test rejects it in both cases. By decomposition of the variability, the functional median polish helps us better understand the impacts of the emission scenario on climate variables, and make further connections between various scenarios and climate change related public health problems possible.

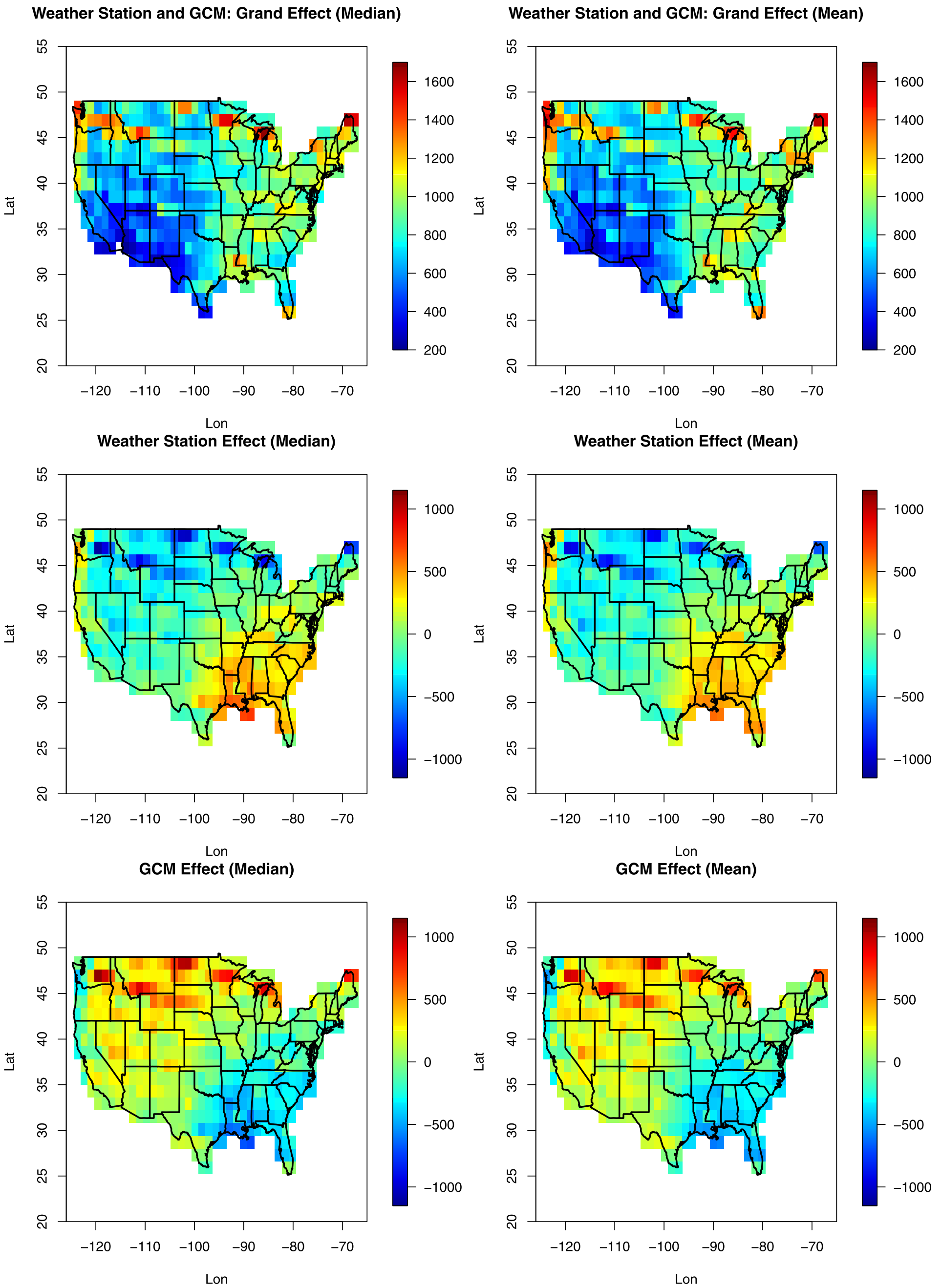

Figure 8. Weather station and GCM precipitation (mm/year) images: estimation of the functional grand and data source effects for the functional median polish and the traditional mean version of functional ANOVA.

4.4. Regional Climate Model

A regional climate model (RCM) is a comprehensive physical model representing the important components of the climate. It has a higher resolution and provides finer spatial

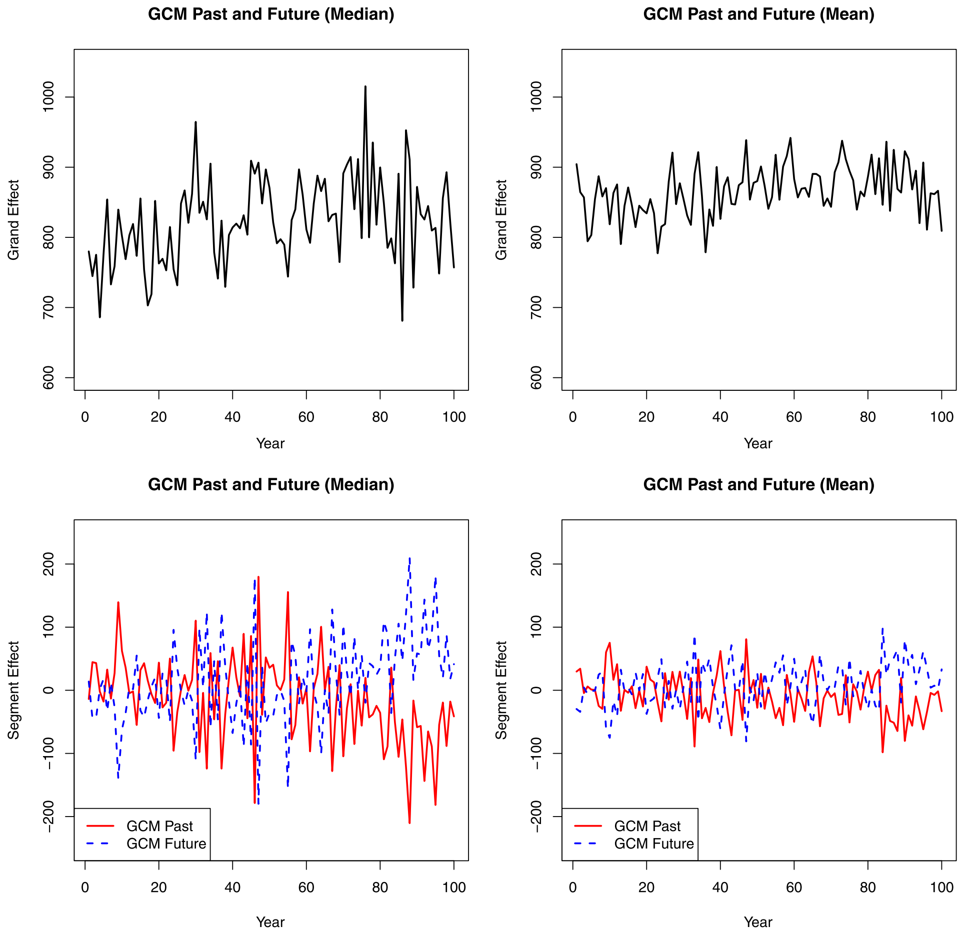

Figure 9. GCM past and future precipitation (mm/year) curves: estimation of the functional grand and segment effects for the functional median polish and the traditional mean version of functional ANOVA.

and temporal details than a GCM. A RCM only covers a limited area of the globe and its lateral boundaries are driven by variables output, for instance, winds, temperatures and humidity, from a GCM. Functional ANOVA is one way to estimate how much variability in the model output is from RCM and how much is due to the boundary conditions from GCM. This problem has been studied by Kaufman and Sain (2010) under a mean-based Bayesian framework for the data from the PRUDENCE project (Christensen, Carter, and Giorgi 2002). We apply our functional median polish to the same subset of the data in Kaufman and Sain (2010), consisting of control runs (1961–1990) for two RCMs (HIRHAM and RCAO) crossed with two GCMs (ECHAM4 and HadAm3H), and analyze the output of summer temperatures for these 30 years over the United Kingdom and Ireland for the four combinations of the RCM and GCM. The RCM High-Resolution Atmospheric Model (HIRHAM) was developed in collaboration between the Danish Meteorological Institute, the Royal Netherlands Meteorological Institute, and the Max Planck Institute for Meteorology. The RCM Rossby Centre Atmosphere-Ocean Model (RCAO) was developed at

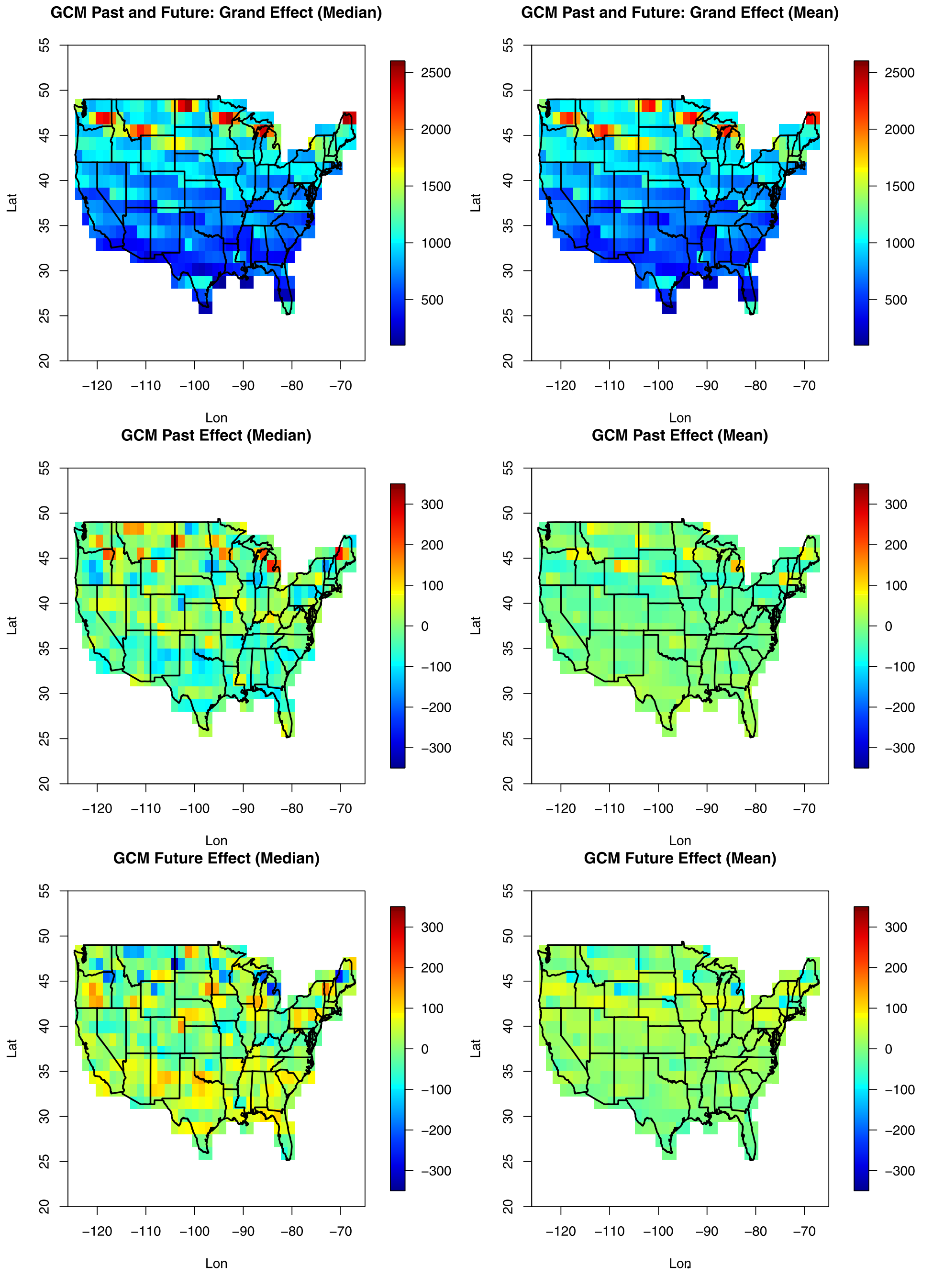

Figure 10. GCM past and future precipitation (mm/year) images: estimation of the functional grand and segment effects for the functional median polish and the traditional mean version of functional ANOVA.

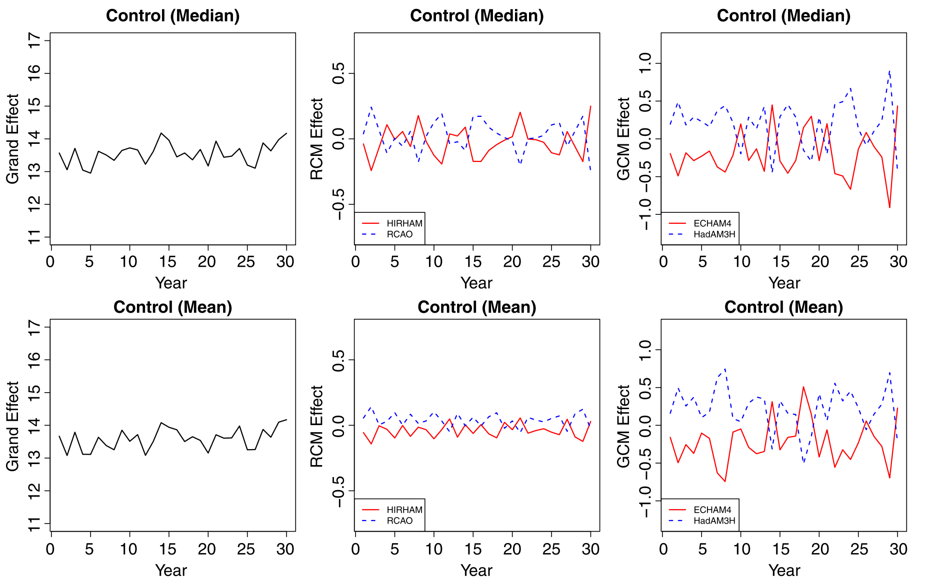

Figure 11. RCM and GCM temperature (°C) curves: estimation of the functional grand and RCM/GCM effects for the functional median polish and the traditional mean version of functional ANOVA.

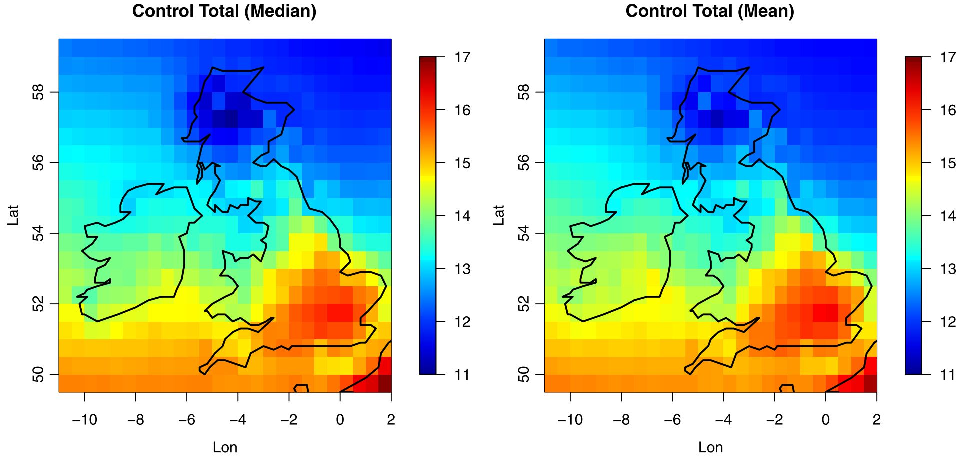

Figure 12. RCM and GCM temperature (°C) images: estimation of the functional grand effect for the functional median polish and the traditional mean version of functional ANOVA.

the Rossby Centre at the Swedish Meteorological and Hydrological Institute. The GCM European Center HAMburg 4 (ECHAM4) was from the Max Planck Institute, and the GCM Hadley Atmospheric Model (HadAm3H) was from the Hadley Centre in the United Kingdom. Details can be found at *http://prudence.dmi.dk/*.

We estimate the functional grand effect and RCM/GCM effects by treating the output either as temporal curves at different locations (Figure 11), or images over the spatial region at different time points (Figures 12, 13, 14) for the functional median polish and the traditional mean version of functional ANOVA.

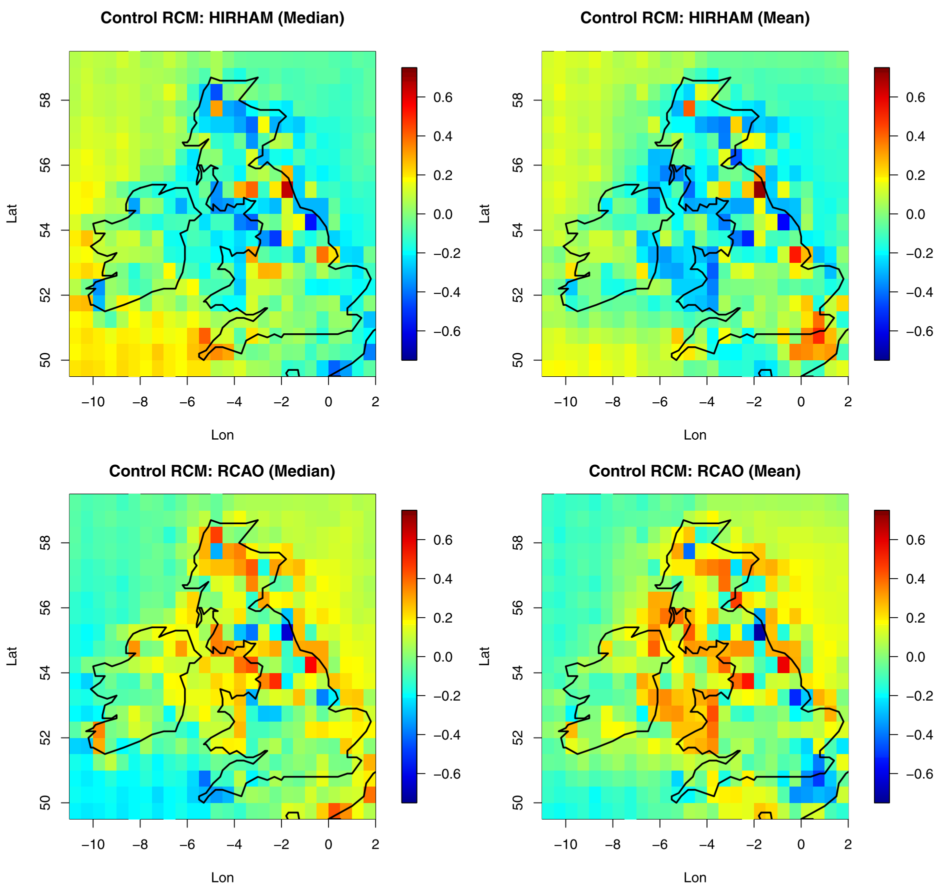

Figure 13. RCM temperature (°C) images: estimation of the functional RCM effect for the functional median polish and the traditional mean version of functional ANOVA.

Because there are only two levels for each factor, we can only look at the curve or the image for one factor effect by symmetry. For both curves and images, it is shown that most of the variability is due to the choice of GCM with a large effect in magnitude over time or space, especially for the North Sea, and with boundary conditions HadAm3H from the GCM, the output tends to be warmer. The choice of RCM has a smaller effect in magnitude for most of the locations, but does have some significant local effects, and the RCM HIRHAM output tends to be warmer in the west and cooler in the east. For curves, the rank test fails to reject H_0 in (2.3) at the 5 % significance level for the GCM effect. Since the functional median represents a typical location, the test shows that the local effect is not significant for different GCM boundary conditions. For the RCM effect, the rank test rejects H_0 at the 5 % significance level showing the significant RCM local effect. For images, the rank test rejects H_0 at the 5 % significance level for the GCM effect indicating its significant effect over the whole spatial region, but fails to reject it for the RCM effect. Since the functional median represents a typical map in this case, the test shows that the

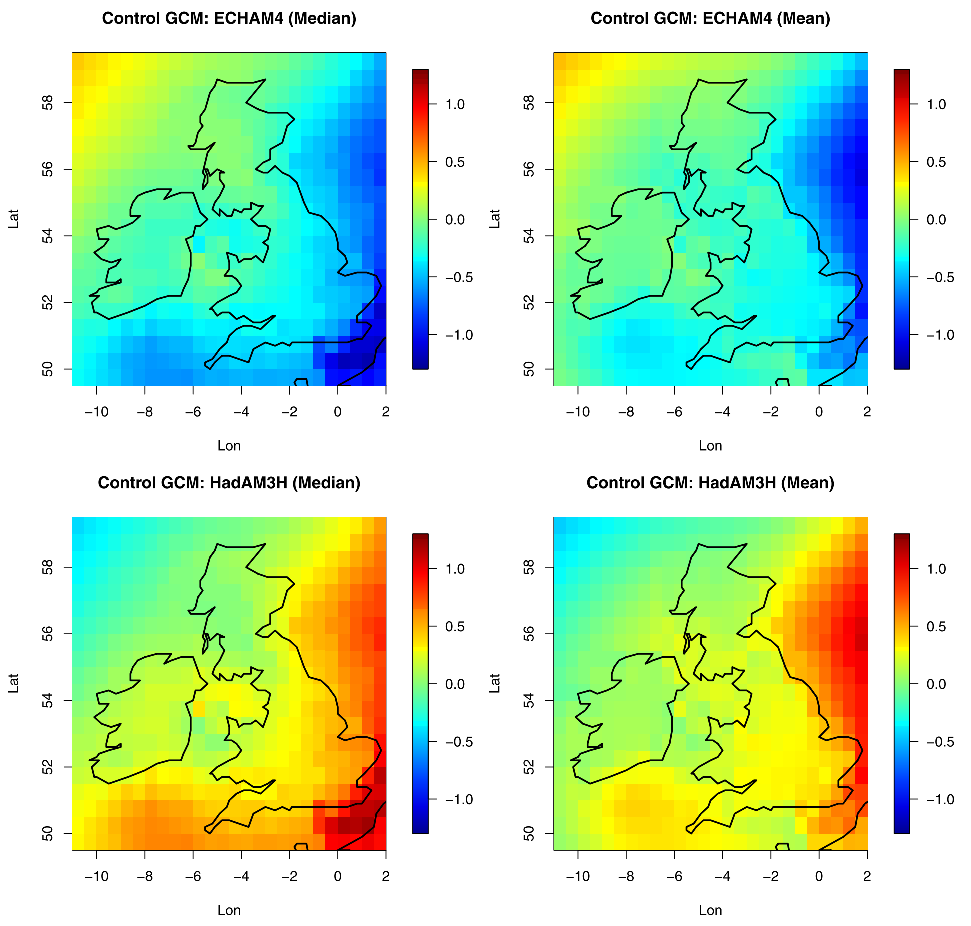

Figure 14. GCM temperature (°C) images: estimation of the functional GCM effect for the functional median polish and the traditional mean version of functional ANOVA.

two RCMs only differ at certain locations but overall the effect is not significant. However, the two-sample functional t-test rejects H_0 at the 5 % significance level for all cases.

The information we have obtained from the functional median polish is consistent with the conclusion in Kaufman and Sain (2010). However, their model is based on means and provides more similar estimation with the traditional mean version of functional ANOVA. To better compare the median and the mean version, we look at the two averaged residual maps in Figure 15. It shows high temperature values at some locations in Ireland and Scotland which indicate outliers remaining in the residuals for the functional median polish, while the residuals in the mean version of functional ANOVA are close to zero everywhere which implies that the estimation of the effects is affected by outliers as demonstrated in our simulations in Section 3. We can also see the influence of the outliers in Figure 12 where the grand effect in the mean version is warmer over the area in Ireland and Scotland than that in the functional median polish. Thus, further investigation for the GCM and RCM at these outlying locations is needed.

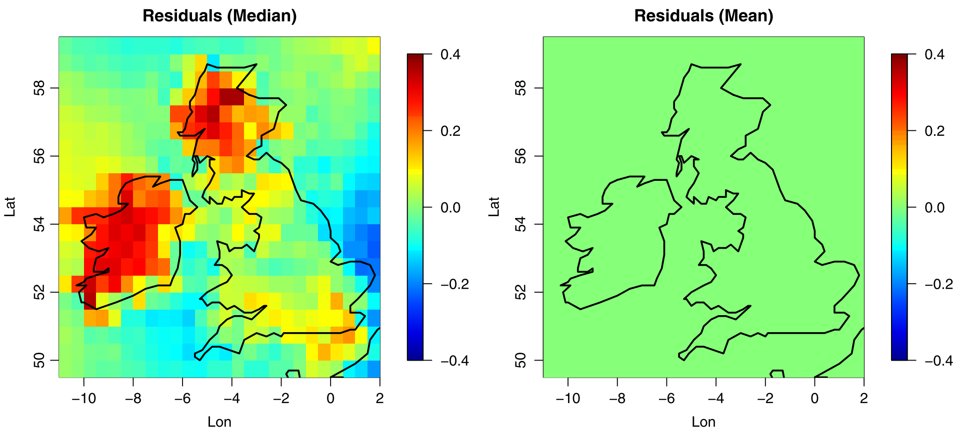

Figure 15. Temperature (°C) residual maps: averaged residual maps for the functional median polish and the traditional mean version of functional ANOVA.

5. DISCUSSION

This article has focused on the ANOVA for functional data. The functional median polish we have proposed is a robust method to fit an additive functional ANOVA model by functional medians, assuming no interaction among factors. As an extension to the univariate median polish, we have described the iterative algorithm for one-way and two-way functional ANOVA, and our simulation studies have shown that the functional median polish is robust under different outlier models while the traditional mean version of functional ANOVA is not resistant to outliers. Our applications focused on climate data collected from weather stations or generated from climate models, where the functional data can be either treated as curves in time or images over space. In these examples, compared to the mean version of functional ANOVA, the functional median polish provided robust estimation of climatic region effects for weather station data, as well as various factor effects in climate models which play an important role in understanding the model scenarios and interpreting model output.

If two factors are taken to be geographic coordinates, e.g. the latitude and longitude of gridded data, then we obtain a functional version of the median polish algorithm described by Cressie (1993) in the context of spatial statistics. Instead of just one number at each spatial locations, we have a whole function of time. Therefore, a large scale trend in space and time can then be removed.

The functional median polish algorithm is based on functional median defined by modified band depth (López-Pintado and Romo 2009), which allows us to order functional data. With this measure, it is natural to extend rank-based tests to functional data. Thus, in the functional ANOVA framework, besides the estimation problem solved by the functional median polish, hypothesis tests can be set up and rank-based tests can be generalized. Moreover, functional median polish can be straightforwardly extended to the setting of sparse functional data by means of the band depth in that context recently defined by López-Pintado and Wei (2011).

Like the classical median polish procedure for univariate data, our functional median polish algorithm does not guarantee to yield the least L_1-norm residuals. Fink (1988) proposed a rather complex modification of the classical procedure that converges to the least L_1-norm residuals. The generalization of his modifications to our functional setting appears to be very challenging and remains an open problem.

ACKNOWLEDGEMENTS

This research was partially supported by NSF grants DMS-1007504, DMS-1100492, and Award No. KUS-C1-016-04, made by King Abdullah University of Science and Technology (KAUST). The authors thank Bo Li, Guest Editor, and two referees for valuable comments, as well as Caspar M. Ammann for providing the GCM data analyzed in Section 4.3.

[Published Online August 2012.]

REFERENCES

Ammann, C. M., Washington, W. M., Meehl, G. A., Buja, L., and Teng, H. (2010), "Climate Engineering Through Artificial Enhancement of Natural Forcings: Magnitudes and Implied Consequences," *Journal of Geophysical Research*, 115, D22109. doi:*10.1029/2009JD012878*.

Christensen, J., Carter, T., and Giorgi, F. (2002), "PRUDENCE Employs New Methods to Assess European Climate Change," *EOS, Transactions—American Geophysical Union*, 83.

Collins, W. D. et al.(2006), "The Community Climate System Model Version 3 (CCSM3)," *Journal of Climate*, 19, 2122–2143. doi:*10.1175/JCLI3761.1*.

Cressie, N. A. C. (1993), *Statistics for Spatial Data*, New York: Wiley.

Emerson, J. D., and Hoaglin, D. C. (1983), "Analysis of Two-Way Tables by Medians," in *Understanding Robust and Exploratory Data Analysis*, eds. D. C. Hoaglin, F. Mosteller, and J. W. Tukey, New York: Wiley, pp. 166–210.

Fink, A. M. (1988), "How to Polish off Median Polish," *SIAM Journal on Scientific and Statistical Computing*, 9, 932–940.

Hoaglin, D. C., Mosteller, F., and Tukey, J. W. (eds.) (1983), *Understanding Robust and Exploratory Data Analysis*, New York: Wiley.

——— (1985), *Exploring Data Tables, Trends, and Shapes*, New York: Wiley.

Intergovernmental Panel on Climate Change (IPCC) (2000), *Special Report on Emission Scenarios*. A special report of Working Group III of the Intergovernmental Panel on Climate Change, eds. N. Nakicenovic and R. Swart, 612, Cambridge: Cambridge Univ. Press.

Kaufman, C. G., and Sain, S. R. (2010), "Bayesian Functional ANOVA Modeling Using Gaussian Process Prior Distributions," *Bayesian Analysis*, 5, 123–150.

López-Pintado, S., and Romo, J. (2009), "On the Concept of Depth for Functional Data," *Journal of the American Statistical Association*, 104, 718–734.

López-Pintado, S., and Wei, Y. (2011), "Depth for Sparse Functional Data," in *Recent Advances in Functional Data Analysis and Related Topics*, ed. F. Ferraty, Berlin: Springer, pp. 209–212.

Mosteller, F., and Tukey, J. W. (1977), *Data Analysis and Regression*, Cambridge: Addison-Wesley.

Ramsay, J. O., Hooker, G., and Graves, S. (2009), *Functional Data Analysis With R and Matlab*, New York: Springer.

Ramsay, J. O., and Silverman, B. W. (2005), *Functional Data Analysis* (2nd ed.), New York: Springer.

Sain, S. R., Nychka, D., and Mearns, L. (2011), "Functional ANOVA and Regional Climate Experiments: A Statistical Analysis of Dynamic Downscaling," *Environmetrics*, 22, 700–711.

Sun, Y., and Genton, M. G. (2011), "Functional Boxplots," *Journal of Computational and Graphical Statistics*, 20, 316–334.

——— (2012), "Adjusted Functional Boxplots for Spatio-temporal Data Visualization and Outlier Detection," *Environmetrics*, 23, 54–64.

Sun, Y., Genton, M. G., and Nychka, D. W. (2012), "Exact Fast Computation of Band Depth for Large Functional Datasets," manuscript.

Tukey, J. W. (1970), *Exploratory Data Analysis*, Vol. 1, Reading: Addison-Wesley, Chapter 5 (Limited Preliminary Edition).

——— (1977), *Exploratory Data Analysis*, Reading: Addison-Wesley.

Velleman, P. F., and Hoaglin, D. C. (1981), *Applications, Basics, and Computing of Exploratory Data Analysis*, N. Scituate: Duxbury Press.