



# Nonparametric estimation for a functional-circular regression model

Andrea Meilán-Vila<sup>1</sup> · Rosa M. Crujeiras<sup>2</sup> · Mario Francisco-Fernández<sup>3</sup>

Received: 5 April 2022 / Revised: 18 November 2022 / Published online: 22 March 2023  
© The Author(s) 2023

## Abstract

Changes on temperature patterns, on a local scale, are perceived by individuals as the most direct indicators of global warming and climate change. As a specific example, for an Atlantic climate location, spring and fall seasons should present a mild transition between winter and summer, and summer and winter, respectively. By observing daily temperature curves along time, being each curve attached to a certain calendar day, a regression model for these variables (temperature curve as covariate and calendar day as response) would be useful for modeling their relation for a certain period. In addition, temperature changes could be assessed by prediction and observation comparisons in the long run. Such a model is presented and studied in this work, considering a nonparametric Nadaraya–Watson-type estimator for functional covariate and circular response. The asymptotic bias and variance of this estimator, as well as its asymptotic distribution are derived. Its finite sample performance is evaluated in a simulation study and the proposal is applied to investigate a real-data set concerning temperature curves.

**Keywords** Circular data · Flexible regression · Functional data · Temperature curves

---

Andrea Meilán-Vila, Rosa M. Crujeiras and Mario Francisco-Fernández have contributed equally to this work.

---

✉ Andrea Meilán-Vila  
ameilan@est-econ.uc3m.es

Rosa M. Crujeiras  
rosa.crujeiras@usc.es

Mario Francisco-Fernández  
mariofr@udc.es

- <sup>1</sup> Department of Statistics, Universidad Carlos III de Madrid, Av. de la Universidad 30, 28911 Leganés, Spain
- <sup>2</sup> CITMAga, Galician Centre for Mathematical Research and Technology, Universidade de Santiago de Compostela, San Xerome s/n, 15782 Santiago de Compostela, Spain
- <sup>3</sup> Department of Mathematics, Faculty of Computer Science, Universidade da Coruña. CITIC, Campus de Elviña s/n, 15071 A Coruña, Spain

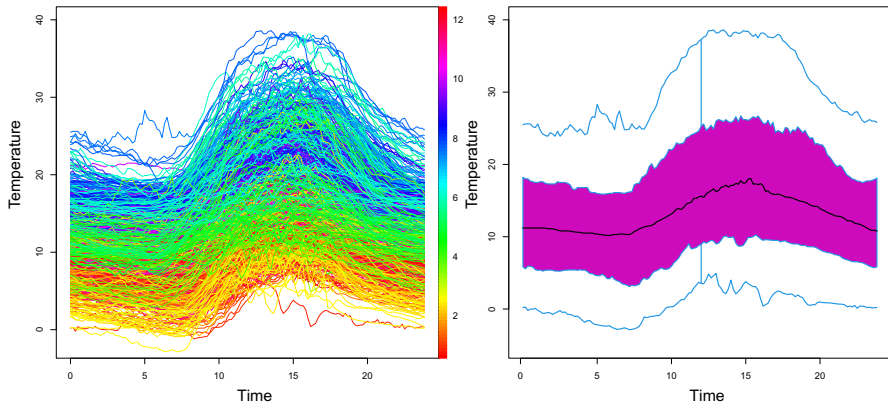
## 1 Introduction

The State of the Climate report by Blunden and Arndt (2020) corresponding to 2019, includes a series of analysis of fully monitored variables on a global scale. One of these variables is surface temperature, revealing that July 2019 was Earth's hottest month on record. In Europe, 2019 was the second hottest year (following 2018), and 2014–2019 are Europe's warmest years on record. The report indicates that the warming of land and ocean surfaces is reflected across the Globe, with lakes and permafrost temperatures increasing, as an evidence of climate change.

However, as pointed out by Bloodhart et al. (2015), it is difficult for individuals to determine if they have experienced the effects of climate change, given that their information refers to a reduced period of time and is usually restricted to a local scale. Nevertheless, understanding how people experience the changes on local weather patterns is important, given that personal experiences are known to affect climate change beliefs (Goebbert et al. 2012) and risk perceptions. Hence, they may condition citizen's support on prevention policies (see Howe 2018; Goebbert et al. 2012; Taylor et al. 2014; Bloodhart et al. 2015, among others). As pointed out by Goebbert et al. (2012) on a global scale, personal perceptions on weather patterns are usually noticed through temperature changes and this indicator is usually taken as an evidence by individuals for making inferences about climate changes. However, these subjective insights should be confirmed with the construction of an appropriate statistical regression model that allows for assessing if relevant changes in temperature patterns could have happened in different periods of time.

As a motivating example, consider daily temperature records in Santiago de Compostela (NW-Spain) for the period 2002–2019. Temperature curves from February 15, 2002 until June 28, 2005 can be seen in Fig. 1 (left), where the color scale indicates the day of the year when each curve was observed (note that the scale-palette preserves the periodicity of the data). In Fig. 1 (right), a functional boxplot (Sun and Genton 2011) is presented, plotting the 50% central region for the observed temperature curves. The vertical segments are the whiskers of the boxplot. Modeling the relation between the temperature curve (as a functional covariate) and the day of the year basis (as a circular response) for a certain period of time would allow to investigate, by comparing the predicted day by such a model for a given temperature curve, in a different period, with the actual day corresponding to that temperature curve, if temperature curve patterns are stable along time or if, on the contrary, observed temperature curves are displaced.

The novel methodology presented in this paper exploits the usage of high frequency temperature records, considering a (flexible) regression model with a functional covariate (the whole daily temperature curve) and a circular response (a day on a year basis). Circular data can be viewed as points on the unit circle, so fixing a direction and a sense of rotation, circular data can be expressed as angles. For a complete introduction to circular data, we refer to Mardia and Jupp (2000) or Ley and Verdebout (2017). This type of data appear in different applied fields, and some examples include wind directions (Fisher 1995), wave directions (Jona-Lasinio et al. 2012; Wang et al. 2015), or animal orientations (Scapini et al. 2002), among others. Regression analysis considering models where the response and/or the covariables are circular variables has been addressed in different papers. Parametric approaches were considered in Fisher



**Fig. 1** Daily temperature curves in Santiago de Compostela (Spain) from February 15, 2002 to June 28, 2005 (left) and its corresponding functional boxplot (right)

and Lee (1992) and Presnell et al. (1998) for regression models with circular response and Euclidean covariates. The authors assumed a parametric distribution model for the circular response variable. An alternative parametric regression model was also analyzed in Kim and SenGupta (2017), considering that a multivariate circular variable depends on several circular covariates. Using nonparametric methods, Di Marzio et al. (2013) introduced a regression estimator for models with circular response, a single real-valued covariate and independent errors and also for errors coming from mixing processes. A similar approach was also applied for time series in Di Marzio et al. (2012). Following similar ideas, Meilán-Vila et al. (2021) proposed and studied local polynomial-type regression estimators considering a model with a circular response and several real-valued covariates for independent data. This class of estimators was also analyzed in Meilán-Vila et al. (2021) in the presence of spatial correlation.

A natural extension of the multiple regression model used in Meilán-Vila et al. (2021) would be to consider a functional covariate belonging to an infinite-dimensional space. Nowadays, with the advances in data collection methods, more and more data are being recorded during an interval time or at several discrete time points with a high frequency, producing functional data. For example, in many fields of applied sciences such as chemometrics (Abraham et al. 2003), biometrics (Gasser et al. 1998) and medicine (Antoniadis and Sapatinas 2007), among others, it is quite common to deal with observations that are curves or functions. The branch of statistics that analyzes this type of data is called functional data analysis (see, for instance, Ramsay and Silverman 2005; Ferraty and Vieu 2006; Kokoszka and Reimherr 2017). The literature on functional regression modeling is really extensive (see, for example, Greven and Scheipl 2017; Morris 2015; Aneiros et al. 2019, 2022; Goia and Vieu 2016, for complete reviews and recent advances in this field), including parametric (in particular, linear) and nonparametric models. Assuming parametric conditions, earlier advances on functional regression were introduced by Ramsay and Silverman (2005), while a recent overview was provided in Febrero-Bande et al. (2017). The nonparametric methodologies became popular in the functional regression context

with the book by Ferraty and Vieu (2006), this topic being widely addressed in the last decade (see Ling and Vieu 2018, for a survey). In this framework, if the explanatory variables are functional, nonparametric regression approaches are essentially based on an adaptation of the Nadaraya–Watson (Ferraty and Vieu 2002, 2006; Ling et al. 2020) or the local linear regression estimators (Aneiros-Pérez et al. 2011; Berlinet et al. 2011; Boj et al. 2008; Baílllo and Grané 2009; Ferraty and Nagy 2022).

As pointed out before the aim of the present work is to propose and study a nonparametric regression estimator for a model with a circular response and a functional covariate. When the response variable is circular, the regression function can be defined as the minimizer of a circular risk function. It can be proved that the minimizer of this risk function is the inverse tangent function of the ratio between the conditional expectation of the sine and the cosine of the response variable. The proposal introduced in this work implicitly considers two regression models, one for the sine and another for the cosine of the response variable. Then, a nonparametric estimator for the regression function is directly obtained by calculating the inverse tangent function of the ratio of appropriate Nadaraya–Watson estimators for the two regression functions of the sine and cosine models. This way of proceeding has already been considered previously. For instance, a similar approach was employed in Meilán-Vila et al. (2021) and Meilán-Vila et al. (2021) for regression models with a circular response, Euclidean covariates and independent or spatially correlated errors, respectively. As in any nonparametric approach, a crucial issue is the selection of an appropriate bandwidth or smoothing parameter. In this paper, cross-validation bandwidth selection methods adapted for the current framework are introduced and analyzed in practice.

This paper is organized as follows. In Sect. 2, the regression model for a circular response and a functional covariate is presented. In Sect. 3, the nonparametric estimator of the circular regression function is proposed. Expressions for its asymptotic bias and variance, as well as its asymptotic distribution are included in Sect. 4. The finite sample performance of the estimator is studied through simulations in Sect. 5, and illustrated in Sect. 6 with the real data example on temperature curves. Finally, a discussion on this paper is included in Sect. 7.

## 2 Functional-circular regression model

Let  $\{(\mathcal{X}_i, \Theta_i)\}_{i=1}^n$  be a random sample from the random vector  $(\mathcal{X}, \Theta)$ , where  $\Theta$  is a circular random variable taking values on  $\mathbb{T} = [0, 2\pi)$ , and  $\mathcal{X}$  is a functional variable supported on  $E$ , a separable Banach space endowed with a norm  $\|\cdot\|$ . This general framework includes  $L^p$ , Sobolev and Besov spaces. Separability condition avoids measurability problems for the random variable  $\mathcal{X}$ .

Assume that the circular random variable  $\Theta$  depends on the functional random variable  $\mathcal{X}$  through the following regression model:

$$\Theta_i = [m(\mathcal{X}_i) + \varepsilon_i]_{(\text{mod } 2\pi)}, \quad i = 1, \dots, n, \quad (1)$$

where  $m$  is a smooth trend or regression function mapping  $E$  onto  $\mathbb{T}$ ,  $\text{mod}$  stands for the modulo operation and  $\varepsilon_i, i = 1, \dots, n$ , is an independent sample of a circular

random variable  $\varepsilon$ , satisfying  $E[\sin(\varepsilon) \mid \mathcal{X} = \chi] = 0$  and having finite concentration. In this setting, the circular regression function  $m$  in model (1) can be defined as the minimizer of the risk function  $E\{1 - \cos[\Theta - m(\mathcal{X})] \mid \mathcal{X} = \chi\}$ . The minimizer of this cosine risk is given by:

$$m(\chi) = \text{atan2}[m_1(\chi), m_2(\chi)], \tag{2}$$

where  $m_1(\chi) = E[\sin(\Theta) \mid \mathcal{X} = \chi]$ ,  $m_2(\chi) = E[\cos(\Theta) \mid \mathcal{X} = \chi]$  and the function  $\text{atan2}(y, x)$  returns the angle between the  $x$ -axis and the vector from the origin to  $(x, y)$ . With this formulation,  $m_1$  and  $m_2$  can be considered as the regression functions of two regression models, being  $\sin(\Theta)$  and  $\cos(\Theta)$  their response variables, respectively, and with functional covariates. Specifically, we assume the models

$$\sin(\Theta_i) = m_1(\mathcal{X}_i) + \xi_i, \quad i = 1, \dots, n, \tag{3}$$

and

$$\cos(\Theta_i) = m_2(\mathcal{X}_i) + \zeta_i, \quad i = 1, \dots, n, \tag{4}$$

where  $m_1$  and  $m_2$  are regression functions mapping  $E$  onto  $[-1, 1]$ . The  $\xi_i$  and the  $\zeta_i$  are independent error terms, absolutely bounded by 1, satisfying  $E(\xi \mid \mathcal{X} = \chi) = E(\zeta \mid \mathcal{X} = \chi) = 0$ . Additionally, for every  $\chi \in E$ , set  $s_1^2(\chi) = \text{Var}(\xi \mid \mathcal{X} = \chi)$ ,  $s_2^2(\chi) = \text{Var}(\zeta \mid \mathcal{X} = \chi)$  and  $c(\chi) = E(\xi\zeta \mid \mathcal{X} = \chi)$ . The assumption that models (3) and (4) simultaneously hold with (1) leads to certain relations between the variances and covariances of the errors in these models, as it is described below.

Set  $\ell(\chi) = E[\cos(\varepsilon) \mid \mathcal{X} = \chi]$ ,  $\sigma_1^2(\chi) = \text{Var}[\sin(\varepsilon) \mid \mathcal{X} = \chi]$ ,  $\sigma_2^2(\chi) = \text{Var}[\cos(\varepsilon) \mid \mathcal{X} = \chi]$  and  $\sigma_{12}(\chi) = E[\sin(\varepsilon)\cos(\varepsilon) \mid \mathcal{X} = \chi]$ . Then, using the sine and cosine addition formulas in model (1), it follows that, for  $i = 1, \dots, n$ ,

$$\sin(\Theta_i) = \sin[m(\mathcal{X}_i)]\cos(\varepsilon_i) + \cos[m(\mathcal{X}_i)]\sin(\varepsilon_i) \tag{5}$$

and

$$\cos(\Theta_i) = \cos[m(\mathcal{X}_i)]\cos(\varepsilon_i) - \sin[m(\mathcal{X}_i)]\sin(\varepsilon_i). \tag{6}$$

Hence, defining  $f_1(\chi) = \sin[m(\chi)]$  and  $f_2(\chi) = \cos[m(\chi)]$  and applying conditional expectations in (5) and (6), it holds that

$$m_1(\chi) = f_1(\chi)\ell(\chi) \quad \text{and} \quad m_2(\chi) = f_2(\chi)\ell(\chi). \tag{7}$$

Note that  $f_1(\chi)$  and  $f_2(\chi)$  correspond to the normalized versions of  $m_1(\chi)$  and  $m_2(\chi)$ , respectively. Indeed, taking into account that  $f_1^2(\chi) + f_2^2(\chi) = 1$ , it can be easily deduced that  $\ell(\chi) = [m_1^2(\chi) + m_2^2(\chi)]^{1/2}$ . Hence, under model (1),  $\ell(\chi)$  amounts to the mean resultant length of  $\Theta$  given  $\mathcal{X} = \chi$ , and taking into account that it is assumed that  $E[\sin(\varepsilon) \mid \mathcal{X} = \chi] = 0$ , also corresponds to the mean resultant length of  $\varepsilon$  given  $\mathcal{X} = \chi$ . The following explicit expressions for the conditional variances of

the error terms involved in models (3) and (4), in terms of the conditional variances and covariance of the Cartesian coordinates of  $\varepsilon$ , can be obtained:

$$s_1^2(\chi) = f_1^2(\chi)\sigma_2^2(\chi) + 2f_1(\chi)f_2(\chi)\sigma_{12}(\chi) + f_2^2(\chi)\sigma_1^2(\chi), \tag{8}$$

$$s_2^2(\chi) = f_2^2(\chi)\sigma_2^2(\chi) - 2f_2(\chi)f_1(\chi)\sigma_{12}(\chi) + f_1^2(\chi)\sigma_1^2(\chi), \tag{9}$$

as well as for the covariance between the error terms in (3) and (4),

$$c(\chi) = f_1(\chi)f_2(\chi)\sigma_2^2(\chi) - f_1^2(\chi)\sigma_{12}(\chi) + f_2^2(\chi)\sigma_{12}(\chi) - f_1(\chi)f_2(\chi)\sigma_1^2(\chi). \tag{10}$$

A nonparametric Nadaraya–Watson-type estimator of the circular regression function  $m$  in model (1) is presented and studied in what follows.

### 3 Nonparametric regression estimator

A Nadaraya–Watson-type estimator for  $m(\chi)$ , in model (1) can be defined by replacing  $m_1(\chi)$  and  $m_2(\chi)$  in (2) by suitable Nadaraya–Watson estimators. Specifically, the following estimator:

$$\hat{m}_h(\chi) = \text{atan2}[\hat{m}_{1,h}(\chi), \hat{m}_{2,h}(\chi)] \tag{11}$$

is considered, where  $\hat{m}_{1,h}(\chi)$  and  $\hat{m}_{2,h}(\chi)$  denote the Nadaraya–Watson estimators of  $m_1(\chi)$  and  $m_2(\chi)$ , respectively. The asymptotic properties (bias, variance and asymptotic normality) of estimator (11) are derived in the following section. For this purpose, assuming that models (3) and (4) hold, the asymptotic properties of  $\hat{m}_{1,h}(\chi)$  and  $\hat{m}_{2,h}(\chi)$  are used.

Considering models (3) and (4), Nadaraya–Watson estimators for  $m_j(\chi)$ ,  $j = 1, 2$ , are respectively defined as:

$$\hat{m}_{j,h}(\chi) = \begin{cases} \frac{\sum_{i=1}^n K(h^{-1}\|\mathcal{X}_i - \chi\|) \sin(\Theta_i)}{\sum_{i=1}^n K(h^{-1}\|\mathcal{X}_i - \chi\|)} & \text{if } j = 1, \\ \frac{\sum_{i=1}^n K(h^{-1}\|\mathcal{X}_i - \chi\|) \cos(\Theta_i)}{\sum_{i=1}^n K(h^{-1}\|\mathcal{X}_i - \chi\|)} & \text{if } j = 2, \end{cases} \tag{12}$$

where  $K$  is a symmetric kernel and  $h = h(n)$  is a strictly positive real-valued bandwidth controlling the smoothness of the estimator. Note that in this functional framework, the support of  $K$  should be contained in  $\mathbb{R}^+$ , since  $\|\chi - \chi'\| \geq 0$ , for all  $\chi, \chi' \in E$ . If the kernel  $K$  is positive with support on  $[0, 1]$ , the estimators given in (12) only consider the observations  $\sin(\Theta_i)$  and  $\cos(\Theta_i)$ , respectively, associated to the curves  $\mathcal{X}_i$  such that  $\|\mathcal{X}_i - \chi\| \leq h$ , since  $K(h^{-1}\|\mathcal{X}_i - \chi\|) = 0$  when the distance between  $\chi$  and  $\mathcal{X}_i$  is larger than  $h$ . The regression estimators given in (12) are the adaptations to the functional context of the classical finite-dimensional Nadaraya–Watson estimators for the sine and cosine regression models (Ferraty et al. 2007).

Although the choice of the kernel function is of secondary importance, the bandwidth parameter plays an important role in the performance of the Nadaraya–Watson estimators (12) and, consequently, of the regression estimator (11). When  $h$  is excessively large, the number of observations involved in the estimation method will also be too large, and an oversmoothed estimator is obtained. Conversely, if the bandwidth parameter is too small, few observations will be considered in the estimation procedure, and an undersmoothed estimator estimator is computed. Therefore, in this case, as in any other kernel-based estimator, in practice, data-driven bandwidth selection methods are needed. A cross-validation approach is used to select the bandwidth parameter  $h$  for (11) in the simulation study and for the real data application. Note that the same bandwidth is used for computing  $\hat{m}_{1,h}$  and  $\hat{m}_{2,h}$  in (12). If two different bandwidths for sine and cosine components were considered, this would imply that possibly different curves were chosen for computing regression estimators of sine and cosine models, which seems quite unnatural, since the Cartesian coordinates are directly related to the the angle itself.

### 4 Asymptotic results

In this section, the asymptotic bias and variance expressions for  $\hat{m}_h(\chi)$ , given in (11), are derived. Moreover, the asymptotic distribution of the proposed estimator is calculated. The proofs of all these theoretical results are collected in Appendix A.

#### 4.1 Asymptotic bias and variance

To derive the asymptotic bias and variance of the estimator given in (11), the asymptotic properties of the Nadaraya–Watson estimators of  $m_j(\chi)$ ,  $j = 1, 2$ , given in (12), are required. Using the asymptotic results given in Ferraty et al. (2007), the asymptotic bias and variance of  $\hat{m}_{j,h}(\chi)$ ,  $j = 1, 2$ , are immediately obtained. These expressions, jointly with the covariance between  $\hat{m}_{1,h}(\chi)$  and  $\hat{m}_{2,h}(\chi)$ , are collected in Lemma 1.

Let  $\varphi_\chi$  and  $\varphi_{j,\chi}$ ,  $j = 1, 2$ , be the functions defined for all  $s \in \mathbb{R}$  by:

$$\varphi_\chi(s) = \mathbb{E} \{ [m(\mathcal{X}) - m(\chi)] \mid \|\mathcal{X} - \chi\| = s \}, \tag{13}$$

$$\varphi_{j,\chi}(s) = \mathbb{E} \{ [m_j(\mathcal{X}) - m_j(\chi)] \mid \|\mathcal{X} - \chi\| = s \}, \tag{14}$$

and denote by  $F_\chi$  the cumulative distribution function of the random variable  $\|\mathcal{X} - \chi\|$ ,

$$F_\chi(t) = \mathbb{P}(\|\mathcal{X} - \chi\| \leq t), \quad t \in \mathbb{R}.$$

In addition, denoting,  $\forall s \in [0, 1]$ ,

$$\tau_{\chi,h}(s) = \frac{F_\chi(hs)}{F_\chi(h)} = \mathbb{P}(\|\mathcal{X} - \chi\| \leq hs \mid \|\mathcal{X} - \chi\| \leq h),$$

the following assumptions are required.

- (C1) The regression functions  $m_j$  and  $s_j^2$ , for  $j = 1, 2$ , are continuous in a neighborhood of  $\chi \in E$ , and  $F_\chi(0) = 0$ .
- (C2)  $\varphi'_{j,\chi}(0)$ , for  $j = 1, 2$ , exist.
- (C3) The bandwidth  $h$  satisfies  $h \rightarrow 0$  and  $nF_\chi(h) \rightarrow \infty$ , as  $n \rightarrow \infty$ .
- (C4) The kernel  $K$  is supported on  $[0, 1]$  and has a continuous derivative on  $[0, 1]$ . Moreover,  $K'(s) \leq 0$  and  $K(1) > 0$ .
- (C5) For all  $s \in [0, 1]$ ,  $\tau_{\chi,h}(s) \rightarrow \tau_{\chi,0}(s)$ , when  $h \rightarrow 0$ .

Assumptions (C1), (C3)–(C4) are similar to those classically employed in the finite-dimensional context. Regarding (C2), as pointed out by Ferraty et al. (2007), this condition avoids the difficulties that could appear when considering differential calculus on Banach spaces. Thus, regularity hypotheses on the regression function  $m$ , such as being twice continuously differentiable, are not required. In assumption (C5), the function  $\tau_{\chi,0}$  is included in the constants that are part of the main terms of the asymptotic expansions of the bias and the variance of the regression estimator (11). The close expression of this function in some particular cases is provided in Ferraty et al. (2007), Proposition 1.

**Lemma 1** *Let  $\{(\mathcal{X}_i, \Theta_i)\}_{i=1}^n$  be a random sample from  $(\mathcal{X}, \Theta)$  supported on  $E \times \mathbb{T}$ . Under assumptions (C1)–(C5), for  $\chi \in E$ , for  $j = 1, 2$ ,*

$$\begin{aligned} \mathbb{E}[\hat{m}_{j,h}(\chi) - m_j(\chi)] &= \varphi'_{j,\chi}(0) \frac{M_{\chi,0}}{M_{\chi,1}} h + \mathcal{O}\left[\frac{1}{nF_\chi(h)}\right] + o(h), \\ \text{Var}[\hat{m}_{j,h}(\chi)] &= \frac{s_j^2(\chi)}{nF_\chi(h)} \frac{M_{\chi,2}}{M_{\chi,1}^2} + o\left[\frac{1}{nF_\chi(h)}\right], \\ \text{Cov}[\hat{m}_{1,h}(\chi), \hat{m}_{2,h}(\chi)] &= \frac{c(\chi)}{nF_\chi(h)} \frac{M_{\chi,2}}{M_{\chi,1}^2} + o\left[\frac{1}{nF_\chi(h)}\right], \end{aligned}$$

where

$$\begin{aligned} M_{\chi,0} &= K(1) - \int_0^1 [sK(s)]' \tau_{\chi,0}(s) ds, \\ M_{\chi,1} &= K(1) - \int_0^1 K'(s) \tau_{\chi,0}(s) ds \end{aligned}$$

and

$$M_{\chi,2} = K^2(1) - \int_0^1 (K^2)'(s) \tau_{\chi,0}(s) ds.$$

**Remark 1** In the case of regression models with Euclidean response and multiple covariates, the main term of the asymptotic bias of the Nadaraya–Watson estimator depends on the gradient and the Hessian matrix of the regression functions (see Härdle and Müller 2012, for further details). Nevertheless, in the present functional setting, the leading term of the asymptotic bias of  $\hat{m}_{j,h}$  depends on the first derivative of the functions  $\varphi_{j,\chi}$  evaluated at zero.



Now, using the previous lemma, the following theorem provides the asymptotic bias and variance of the estimator  $\hat{m}_h(\chi)$ .

**Theorem 1** *Let  $\{(\mathcal{X}_i, \Theta_i)\}_{i=1}^n$  be a random sample from  $(\mathcal{X}, \Theta)$  supported on  $E \times \mathbb{T}$ . Under assumptions (C1)–(C5), the asymptotic bias of estimator  $\hat{m}_h(\chi)$ , for  $\chi \in E$ , is given by:*

$$\mathbb{E}[\hat{m}_h(\chi) - m(\chi)] = \varphi'_\chi(0) \frac{M_{\chi,0}}{M_{\chi,1}} h + \mathcal{O}\left[\frac{1}{nF_\chi(h)}\right] + o(h), \tag{15}$$

and the asymptotic variance is:

$$\text{Var}[\hat{m}_h(\chi)] = \frac{1}{nF_\chi(h)} \frac{\sigma_1^2(\chi)}{\ell^2(\chi)} \frac{M_{\chi,2}}{M_{\chi,1}^2} + o\left[\frac{1}{nF_\chi(h)}\right]. \tag{16}$$

The asymptotic expressions of the bias and variance obtained in Theorem 1 can be used to derive the asymptotic mean squared error (AMSE) of  $\hat{m}_h(\chi)$ . An asymptotically optimal local bandwidth for  $\hat{m}_h(\chi)$  can be selected minimizing this error criterion with respect to  $h$ . A plug-in bandwidth selector could be defined replacing the unknown quantities in the asymptotic optimal parameter by appropriate estimates. This problem is more complex in this functional context than in the finite-dimensional setting (Ferraty et al. 2007). Moreover, the design of global plug-in selectors would require obtaining integrated versions of the AMSE. Making this issue rigorous poses challenges that continue to be topics of current research. Therefore, in this paper, we avoid using plug-in bandwidth selectors and, as pointed out before, we employ a suitable adapted cross-validation technique to select the bandwidth for  $\hat{m}_h$  in the numerical studies (simulations and real data application).

### 4.2 Asymptotic normality

Next, the asymptotic distribution of the proposed nonparametric regression estimator (11) is derived. Apart from the assumptions stated in Sect. 4.1 for computing its asymptotic bias and variance, the following condition is needed:

(C6)  $\varphi'_\chi(0) \neq 0$  and  $M_{\chi,0}$  in Lemma 1 is strictly positive.

Condition (C6) is required to ensure that the leading term of the bias does not vanish. The first part of this assumption is similar to the one stated in the finite-dimensional setting. The second part of (C6) is specific for the infinite-dimensional framework and it is satisfied in some standard situations. For example, if  $\tau_{\chi,0}(s) \neq \mathbb{I}_{(0,1]}(s)$  (being  $\mathbb{I}_{(0,1]}$  the indicator function on the set  $(0, 1]$ ) and  $\tau_{\chi,0}$  is continuously differentiable on  $(0, 1)$ , or if  $\tau_{\chi,0}(s) = \delta_1(s)$  (where  $\delta_1$  stands for the Dirac mass at 1), then  $M_{\chi,0} > 0$  for any kernel  $K$  satisfying (C4). For further details on condition (C6), we refer to Ferraty et al. (2007).

The following theorem provides the asymptotic distribution of the nonparametric regression estimator proposed in (11).

**Theorem 2** Under assumptions (C1)–(C6), it can be proved that , as  $n \rightarrow \infty$ ,

$$\sqrt{nF_\chi(h)} \frac{\ell(\chi)M_{\chi,1}}{\sqrt{\sigma_1^2(\chi)M_{\chi,2}}} [\hat{m}_h(\chi) - m(\chi) - \mathbf{B}_h] \xrightarrow{\mathcal{L}} N(0, 1),$$

where  $\xrightarrow{\mathcal{L}}$  denotes convergence in distribution, with

$$\mathbf{B}_h = \varphi'_\chi(0) \frac{M_{\chi,0}}{M_{\chi,1}} h.$$

A simpler version of the previous theorem can be derived just by considering the following additional assumption:

(C7)  $\lim_{n \rightarrow \infty} h\sqrt{nF_\chi(h)} = 0.$

Condition (C7) allows the cancellation of the bias term  $\mathbf{B}_h$  in Theorem 2. Under this assumption (and also if (C1)–(C6) hold), the pointwise asymptotic Gaussian distribution for the functional nonparametric regression estimate is given in Corollary 1.

**Corollary 1** Under assumptions (C1)–(C7), it can be proved that , as  $n \rightarrow \infty$ ,

$$\sqrt{nF_\chi(h)} \frac{\ell(\chi)M_{\chi,1}}{\sqrt{\sigma_1^2(\chi)M_{\chi,2}}} [\hat{m}_h(\chi) - m(\chi)] \xrightarrow{\mathcal{L}} N(0, 1).$$

Notice that in order to use the asymptotic distribution of  $\hat{m}_h$  in practice, for inferential purposes, it is necessary to estimate the functions  $\ell$  and  $\sigma_1^2$ . Moreover, both constants  $M_{\chi,1}$  and  $M_{\chi,2}$  have to be computed. Assuming a uniform kernel function,  $K(u) = \mathbb{I}_{[0,1]}(u)$ , it is obvious that  $M_{\chi,1} = M_{\chi,2} = 1.$

**Corollary 2** Under assumptions (C1)–(C7), considering  $K(u) = \mathbb{I}_{[0,1]}(u)$  and if  $\hat{\sigma}_1^2$  an  $\hat{\ell}$  are consistent estimators of  $\sigma_1^2$  and  $\ell$ , respectively, it can be proved that

$$\sqrt{n\hat{F}_\chi(h)} \frac{\hat{\ell}(\chi)}{\sqrt{\hat{\sigma}_1^2(\chi)}} [\hat{m}_h(\chi) - m(\chi)] \xrightarrow{\mathcal{L}} N(0, 1) \text{ as } n \rightarrow \infty,$$

where

$$\hat{F}_\chi(h) = \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{\{\|\mathcal{X}_i - \chi\| \leq h\}},$$

being  $\mathbb{I}_{\{A\}} = 1$  if  $A$  is true and 0 otherwise.

**Remark 2** An immediate consequence of Corollary 2 is the possibility of constructing (asymptotic) confidence intervals for the circular regression function. Fixing a

confidence level  $1 - \alpha$ , with  $\alpha \in (0, 1)$ , it can be easily obtained that an asymptotic confidence interval for  $m$  at  $\chi \in E$  is:

$$\left( \hat{m}_h(\chi) \mp z_{\alpha/2} \frac{1}{\sqrt{n\hat{F}_\chi(h)}} \frac{\sqrt{\hat{\sigma}_1^2(\chi)}}{\hat{\ell}(\chi)} \right),$$

where  $z_{\alpha/2}$  denotes the  $(1 - \alpha)$ -quantile of the standard normal distribution and  $\hat{\sigma}_1^2(\chi)$  and  $\hat{\ell}(\chi)$  are (consistent) estimators of  $\sigma_1^2(\chi)$  and  $\ell(\chi)$ , respectively.

Among the possible (consistent) conditional variance estimators for  $\sigma_1^2$ , we suggest to use the following method based on adapting to the functional-circular framework the approach studied in Vilar-Fernández and Francisco-Fernández (2006). In the present context, taking into account that

$$\sigma_1^2(\chi) = \text{Var}[\sin(\varepsilon) \mid \mathcal{X} = \chi] = \mathbb{E}[\sin^2(\varepsilon) \mid \mathcal{X} = \chi] - \{\mathbb{E}[\sin(\varepsilon) \mid \mathcal{X} = \chi]\}^2,$$

and under the assumption that  $\mathbb{E}[\sin(\varepsilon) \mid \mathcal{X} = \chi] = 0$ , only  $\mathbb{E}[\sin^2(\varepsilon) \mid \mathcal{X} = \chi]$  has to be estimated. For this, in a first step, considering model (1), the residuals from a nonparametric fit (using, for instance, the Nadaraya–Watson estimator with a pilot bandwidth) are obtained. Then, in a second step, the estimator of the conditional variance function  $\sigma_1^2$  is defined as the nonparametric estimator (employing again the Nadaraya–Watson estimator with another bandwidth) of the regression function using the squared sine of the residuals as the response variables.

Regarding function  $\ell$ , taking into account that  $\ell(\chi) = [m_1^2(\chi) + m_2^2(\chi)]^{1/2}$ , a consistent estimator for this function is straightforwardly obtained replacing in this equation  $m_j$  by the Nadaraya–Watson estimators  $\hat{m}_{j,h}$ ,  $j = 1, 2$ , defined in (12).

### 5 Simulation study

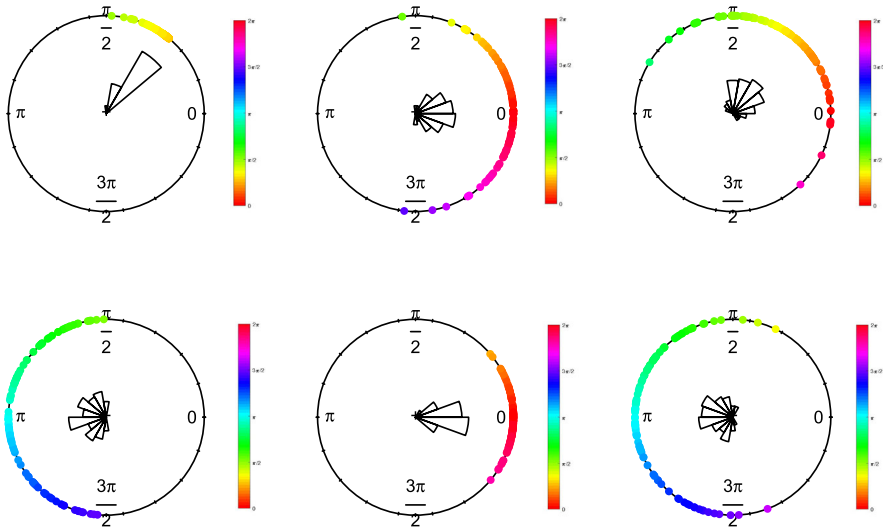
In this section, the regression estimator proposed in (11) is analyzed numerically by simulation, considering different scenarios and model (1). For each scenario, 500 samples of size  $n$  ( $n = 50, 100, 200$  and  $400$ ) are generated using the following regression functions:

$$\begin{aligned} \text{r1: } m(\mathcal{X}) &= \text{atan2}(0.5 + \int_0^{\mathbb{I}} \mathcal{X}(t)dt, \int_0^{\mathbb{I}} \mathcal{X}(t)dt), \\ \text{r2: } m(\mathcal{X}) &= \text{acos}(-0.3 \int_0^{\mathbb{I}} \mathcal{X}(t)dt) + 3/2\text{acos}(0.4 \int_0^{\mathbb{I}} \mathcal{X}(t)dt), \end{aligned}$$

where the random curves are simulated from

$$\mathcal{X}(t) = 30U(1 - t)t^{1+U}, \quad t \in [0, 1], \tag{17}$$

being  $U$  a uniformly distributed variable on  $[0, 1]$ , and the circular errors  $\varepsilon$  in (1) are drawn from a von Mises distribution  $vM(0, \kappa)$ , with different concentrations ( $\kappa = 5, 10$  and  $15$ ).



**Fig. 2** Illustration of model generation (model with regression function r1: top row; model with regression function r2: bottom row). Left: regression function evaluated at a random sample of size  $n = 100$  of curves simulated from (17). Center: independent errors from a von Mises distribution with zero mean and concentration  $\kappa = 5$  (for the model with regression function r1) and  $\kappa = 15$  (for the model with regression function r2). Right: random responses obtained by adding the two previous plots

As an illustration, Fig. 2 shows two realizations of simulated data of size  $n = 100$  (model with regression function r1 in top row and model with regression function r2 in bottom row). Left plots display the circular regression functions evaluated in the curves generated from (17). Central panels present the random errors drawn from a von Mises distribution with zero mean direction and concentration  $\kappa = 5$  (for model with regression function r1) and  $\kappa = 15$  (for model with regression function r2). It can be seen that the errors in the top row, corresponding to  $\kappa = 5$ , present more variability than the ones generated with  $\kappa = 15$ . Right panels show the values of the response variables, obtained adding the mean values in the left panels and the circular errors in the central panels.

For each simulated sample, the Nadaraya–Watson-type estimator (11) is computed using the kernel  $K(u) = 1 - u^2$ ,  $u \in (0, 1)$ , and taking the  $L^2$  metric to calculate the distance between curves. Regarding the bandwidth  $h$ , it is selected by cross-validation, choosing the bandwidth parameter  $h_{CV}$  that minimizes the function:

$$CV(h) = \sum_{i=1}^n \left\{ 1 - \cos \left[ \Theta_i - \hat{m}_h^{(i)}(\mathcal{X}_i) \right] \right\}, \tag{18}$$

where  $\hat{m}_h^{(i)}$  denotes the Nadaraya–Watson-type estimator computed using all observations except for  $(\mathcal{X}_i, \Theta_i)$ .

**Table 1** Average CASE given in (19), over 500 replicates, using a cross-validation bandwidth,  $h_{CV}$ , and an optimal bandwidth,  $h_{CASE}$ , as a benchmark

$\kappa$	$n$	r1		r2	
		$h_{CV}$	$h_{CASE}$	$h_{CV}$	$h_{CASE}$
5	50	0.0096	0.0037	0.0122	0.0087
	100	0.0046	0.0021	0.0096	0.0069
	200	0.0011	0.0005	0.0052	0.0047
	400	0.0004	0.0002	0.0030	0.0023
10	50	0.0020	0.0018	0.0072	0.0058
	100	0.0010	0.0007	0.0062	0.0046
	200	0.0005	0.0003	0.0033	0.0028
	400	0.0004	9e-05	0.0026	0.0021
15	50	0.0019	0.0017	0.0056	0.0048
	100	0.0007	0.0005	0.0056	0.0039
	200	0.0004	0.0003	0.0021	0.0019
	400	0.0002	8e-05	0.0015	0.0013

Data are generated from models with regression functions r1 (left) and r2 (right) and von Mises errors with different concentration parameters ( $\kappa = 5, 10, 15$ )

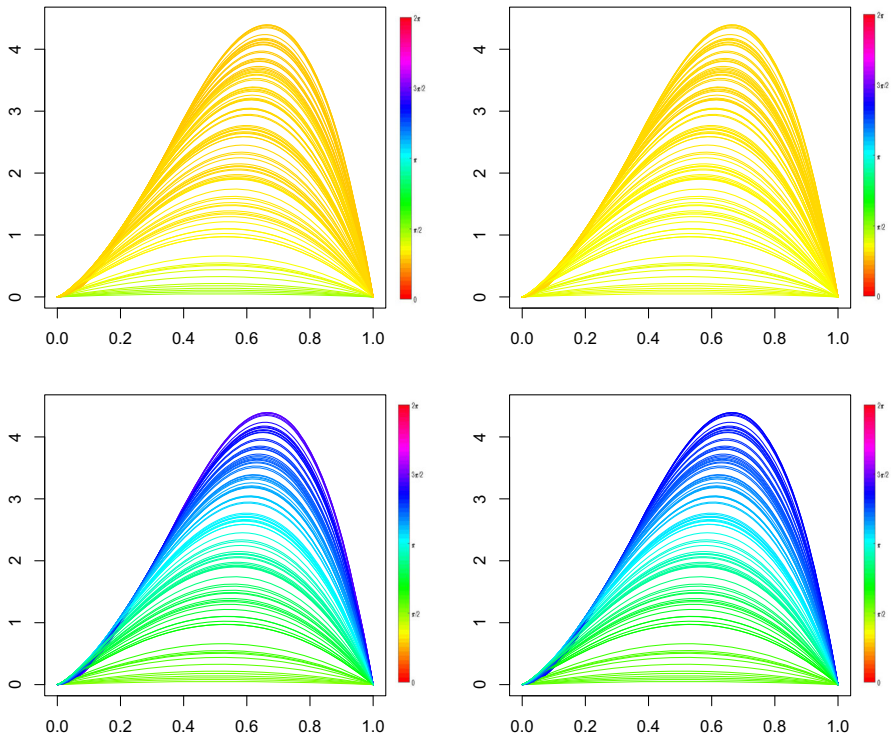
To evaluate the performance of the Nadaraya–Watson-type estimator  $\hat{m}_h$ , the circular average squared error (CASE), defined by Kim and SenGupta (2017):

$$CASE[\hat{m}_h(\chi)] = \frac{1}{n} \sum_{i=1}^n \{1 - \cos [m(\mathcal{X}_i) - \hat{m}_h(\mathcal{X}_i)]\}, \tag{19}$$

is computed. Table 1 shows the average (over 500 replicates) of the CASE when  $h$  is selected by cross-validation. For comparative purposes, in each scenario, the average of the minimum value of the CASE, computed using the bandwidth  $h_{CASE}$  that minimizes (19), is also included in Table 1. Note that the bandwidth  $h_{CASE}$  can not be computed in a practical situation, because the regression function is unknown. The corresponding CASE errors are included in this table as a benchmark. In the different scenarios, it can be seen that the average errors decrease when the sample size increases. In addition, as expected, results are better when the error concentration gets larger.

The appropriate performance of the nonparametric circular regression estimator (11) is also observed in Fig. 3. In this plot, the same scenarios as those considered to obtain Fig. 2 are used. On the left, the theoretical regression functions r1 and r2 are shown, and, on the right, the corresponding estimates using a random sample generated from these models are presented. The representations in the top row correspond to the data simulated using the model with regression function r1 and those in the bottom row with regression function r2.

To complete the study, we repeated the simulations using a  $k$ -nearest neighbor ( $kNN$ ) version of our estimator (see Burba et al. 2009, for a study of the  $kNN$  method in a regression model with scalar response a functional covariate). The results obtained were very similar to those achieved by the Nadaraya–Watson approach, with a slightly



**Fig. 3** Theoretical regression function (left), jointly with the corresponding estimates (right), using the specific scenarios considered in Fig. 2, for the model with regression function  $r_1$  (top row) and  $r_2$  (bottom row)

better performance of the Nadaraya–Watson-type estimator. For reasons of space, only the results with  $r_1$  using the  $k$ NN-type estimator are presented (Table 2). For comparison, we also include in this table the part of Table 1 relative to  $r_1$ .

## 6 Real data illustration

There is a quite extensive literature on the computation of daily temperature averages or profiles, which are useful to summarize temperature patterns (see Huld et al. 2006; Ma and Guttorp 2013; Bernhardt et al. 2018). More recently, there has been some interest in investigating the identification of daily temperature extremes (see Glynis et al. 2021) or the association of diurnal temperature range with mortality burden (see Cai et al. 2021). All these approaches are possible given the high frequency registries of temperatures that are available nowadays, claiming for a proper statistical analysis using functional data methods (see Ruiz-Medina and Espejo 2012, for a spatial autoregressive functional approach, and Aguilera-Morillo et al. (2017), for spatial functional prediction). These works, tailored for a spatial analysis, do not consider the calendar time as a response.

**Table 2** Average CASE given in Eq. (19), over 500 replicates, using the  $k$ NN-type estimator ( $k$ NN)

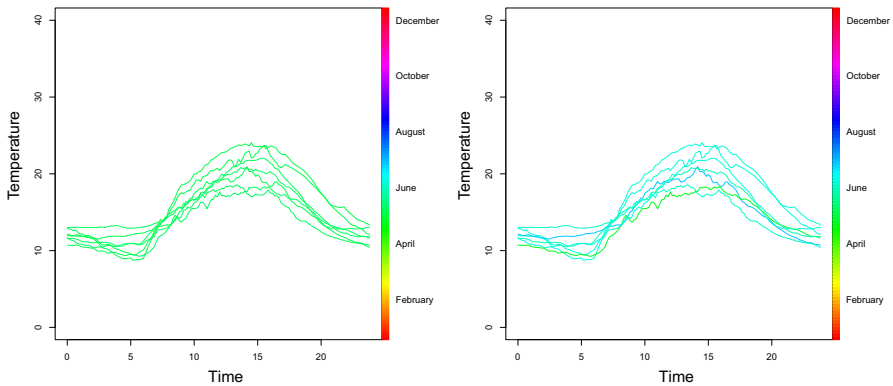
$\kappa$	$n$	$h_{CV}$	$h_{CASE}$	$k$ NN
5	50	0.0096	0.0037	0.0472
	100	0.0046	0.0021	0.0152
	200	0.0011	0.0005	0.0070
	400	0.0004	0.0002	0.0036
10	50	0.0020	0.0018	0.0216
	100	0.0010	0.0007	0.0073
	200	0.0005	0.0003	0.0035
	400	0.0004	9e−05	0.0018
15	50	0.0019	0.0017	0.0143
	100	0.0007	0.0005	0.0048
	200	0.0004	0.0003	0.0024
	400	0.0002	8e−05	0.0013

Data are generated from the model with regression function  $r1$  and von Mises errors with different concentration parameters ( $\kappa = 5, 10, 15$ ). For comparison, results obtained with the Nadaraya–Watson-type estimator in this scenario, presented in Table 1, are also included

Daily temperature curves (obtained by data records every 10 min, so each curve has 144 points) have been collected by Meteogalicia<sup>1</sup> in Santiago de Compostela (NW-Spain) from February 15, 2002 until December 31, 2019. The northern region of Spain presents an oceanic climate, characterized by cool winters and warm summers, with mild temperatures in spring and fall seasons. For the initial period from February 15, 2002 to June 28, 2005, using the available sample  $(\mathcal{X}_i, \Theta_i), i = 1, \dots, n$ , where  $\mathcal{X}_i$  denotes the whole daily temperature curve and  $\Theta_i$  the corresponding day on a year basis, for the day  $i$  in that period (see Fig. 1 in the Introduction), the regression estimator proposed in (11) is computed. The fitted model is obtained using the same kernel and metric as in the simulations and considering the cross-validation bandwidth selector obtained minimizing (18), which gives  $h_{CV} = 8.8139$ . As it has been explained in the Introduction, relevant changes in temperature patterns could be revealed by comparing the predicted days with the fitted model for some temperature curves, in a different period, with the observed days corresponding to those temperature curves. Note that the period selected for estimating the model contains enough data to proceed with our nonparametric method, and it is far enough in time to mitigate temporal correlation from the prediction experiment carried out in what follows.

As a specific example, consider the period from May 21 to May 27, 2019. For this spring week, the 7 observed curves can be seen in Fig. 4 (left), with colors indicating the observed days. In Fig. 4 (right), the same 7 observed curves are now colored according to the predicted days where these curves should be observed with the model fitted from the 2002–2005 data in Fig. 1 (left). According to the model for this early period, the observed curves in the considered week in May 2019 correspond to days in June, July and August. So, this week in May 2019 would be perceived as warmer with respect to the reference period.

<sup>1</sup> Meteogalicia: <https://www.meteogalicia.gal>.



**Fig. 4** Daily temperature curves in Santiago de Compostela (Spain) from May 21, 2019 to May 27, 2019 (left). Colors indicate the observed day. Temperature curves colored by predicted days using the fitted model (right)

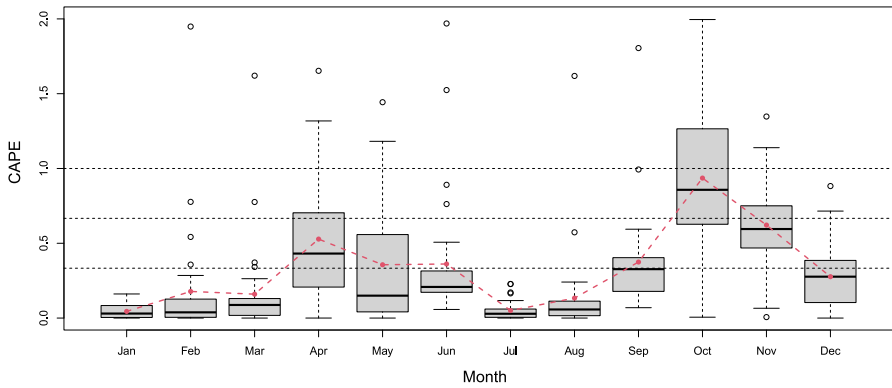
For a more general analysis, consider the temperature curves registered in 2019 and take a division by months. We denote the observed sample by  $(\mathcal{X}_j^b, \Theta_j^b)$ , for  $j = 1, \dots, n_b$  and  $b = 1, \dots, 12$ , being  $n_b$  the number of days of month  $b$ . Similarly to the measurement error considered for evaluating the performance of the estimator in (19), a circular average prediction error (CAPE) is defined for each month as:

$$CAPE_b = \frac{1}{n_b} \sum_{j=1}^{n_b} \left\{ 1 - \cos \left[ \Theta_j^b - \hat{m}_h(\mathcal{X}_j^b) \right] \right\}, \tag{20}$$

where  $\hat{m}_h(\mathcal{X}_j^b)$ , for  $j = 1, \dots, n_b$  and  $b = 1, \dots, 12$ , is the predicted day in the year 2019 with the fitted model using the 2002–2005 data. Note that this prediction error is given in an easily interpretable scale since a perfect match between observed and predicted days for a sample of temperature curves yields a null value if predictions are located 3-months later/before than observations, then the prediction error takes value 1 (so a month time lag gives one third); if predictions and observations are 6-months apart, then the prediction error takes value 2. Boxplots for monthly prediction errors for 2019 are depicted in Fig. 5.

Note that there are seven months (April, May, June, September, October, November and December) that present prediction errors larger than  $1/3$ , indicating that the mismatch between the observed day for a certain curve and the predicted day (with the model fitted for the early period) is larger than one month. On the other hand, it is observed in Fig. 5 that January/February and July/August in 2019 behave, more or less, as “usual” Januaries/Februarys or Julys/Augusts in the years considered in the training sample, perhaps with only a small “shift” in days for the temperature curves of these months. The experts from the meteorological service consulted by us indicated that this is the expected behavior for these months in a context of global warming. In the Atlantic climate (the one corresponding to Galicia, where the data were recorded), one of the manifestations of climate change is the disappearance of





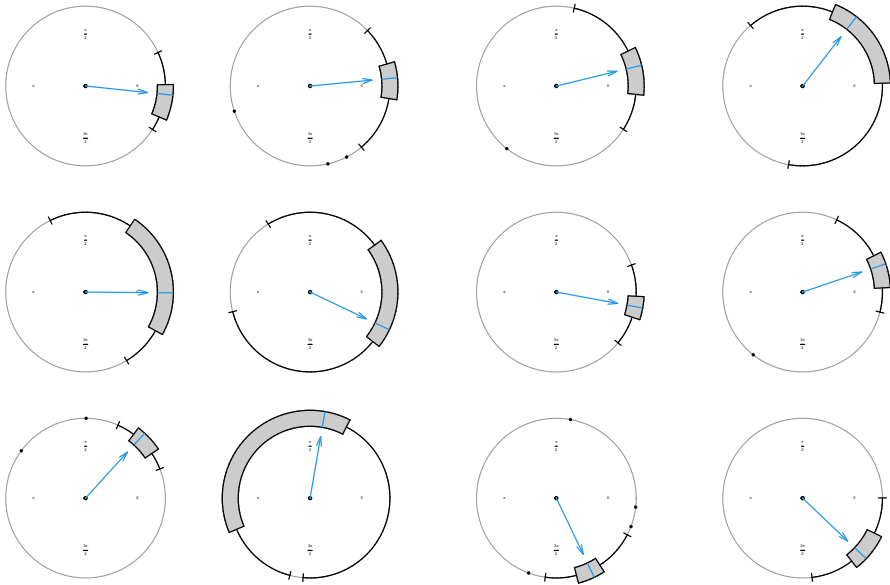
**Fig. 5** Boxplots for the circular prediction errors, by months, for 2019. Red dots: circular average prediction error for each month. (Color figure online)

seasons, specifically, spring and fall. January/February and July/August are still the coolest and warmest months, respectively, in the region considered in this research and they would be the months that would distinguish the “periods” of the year.

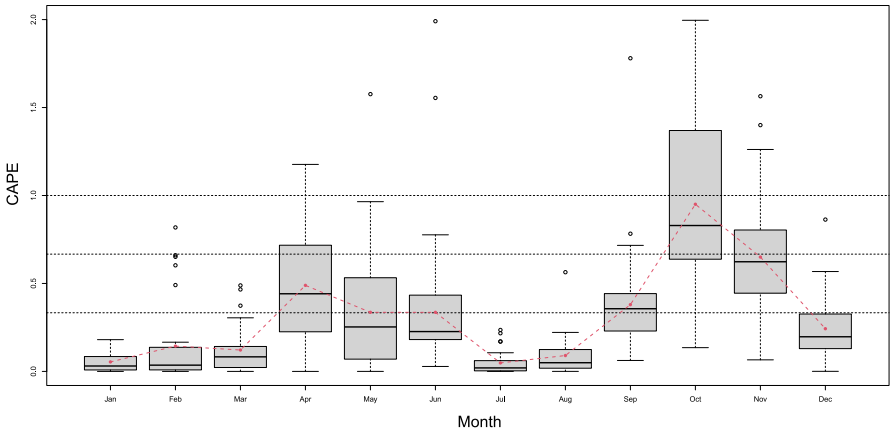
Moreover, to analyze more deeply the direction of the displacements observed in Fig. 5, circular boxplots (Buttarazzi et al. 2018) of the differences between the observed days and those predicted by our nonparametric estimator are computed. Figure 6 shows these circular boxplots, for all the months, that is, the circular boxplots of  $\Theta_j^b - \hat{m}_h(\mathcal{X}_j^b)$ , for  $j = 1, \dots, n_b$  and  $b = 1, \dots, 12$ . Note the large variability in predictions for April, May, June and October, as well as the small variability but with a clear shift (prediction errors are not centered at zero) for September (warmer than expected according to the fitted model), November and December (cooler than expected). Additionally, to rule out that year 2019 was a singular year, we replicated this experiment for years 2017, 2018 and 2020, showing a very similar pattern in all cases.

We also repeated this study using the  $k$ NN-type estimator compared in the simulations. The results and conclusions derived are practically the same as those obtained with the Nadaraya–Watson-type estimator. Figure 7 shows the same information as that presented in Fig. 5, but using the  $k$ NN-type estimator.

As a final comment, it should be highlighted the potential use of the proposed method in order to explain climate change effects on agriculture. The IPCC 2022 (Pörtner HO et al. 2022, Chap. 5) indicates that temperature trend changes have modified the life cycle of crops (both shortening or prolonging them, depending on latitude). In some areas, these effects have lead farmers to change their agricultural practices. In addition, the IPCC 2022 (Pörtner HO et al. 2022) also notices that temperature variability directly affects both the characteristics of some harvests (e.g. acidity or colour of fruits) and the risk of pests and diseases. Understanding temperature variability throughout the year can be useful for adapting agricultural practices.



**Fig. 6** Circular boxplots for the circular prediction errors, by months, for 2019. The months are in order from left to right and from top to bottom



**Fig. 7** Boxplots for the circular prediction errors, by months, for 2019. Red dots: circular average prediction error for each month, using  $k$ NN-type estimator. (Color figure online)

### 7 Discussion

Regression analysis of models involving non-Euclidean data represents a challenging problem given the need of developing new statistical approaches for its study. In this paper, we focus on one of these non-standard settings. In particular, we consider a regression model with a circular response and a functional covariate. A nonparametric estimator of Nadaraya–Watson type of the corresponding regression function is

proposed and analyzed from a theoretical and a practical point of view. Asymptotic bias and variance expressions, as well as asymptotic normality of the nonparametric approach are derived. The finite sample performance of the estimator is assessed by simulations and illustrated using a real data set. From a practical perspective, our methodology allows modeling the relation between a temperature curve (as a covariate) and a day (as a response), enabling to illustrate how temperature patterns change on a local scale.

As in any kernel-based method an important tuning parameter, called the bandwidth or smoothing parameter, has to be selected by the user. In this paper, a cross-validation approach was proposed for this task and used in numerical studies. Other possible selectors of plug-in type or based on bootstrap resampling methods could also be designed. While the use of global plug-in bandwidths in the infinite-dimensional context can be very intricate, given the difficulty of dealing with integrated versions of the AMSE, a bootstrap bandwidth selection method could be easily formulated. The idea would consist in approximating the CASE, given in (19) (or the mean of the CASE), by a bootstrap version of this error criterion using bootstrap samples obtained from bootstrap residuals. For this, two pilot bandwidths are needed, one to obtain the residuals and another to compute the bootstrap samples. This second pilot bandwidth can also be used to compute the Nadaraya–Watson-type estimator of  $m$  employed instead of the theoretical regression function in the bootstrap expression of the CASE (19). Repeating this process many times, the bootstrap bandwidth would be obtained by minimizing the averaged bootstrap CASE. Theoretical justification for this functional-circular bootstrapping procedure as well as the design of practical rules to obtain both pilot bandwidths are open problems out of the scope of the present paper, but of interest for a future research.

In this paper, we have focused on a Nadaraya–Watson-type estimator for  $m$  in model (1). However, a local linear-type estimator could be also defined. In that case, local linear regression estimators for models with a Euclidean response and a functional covariate are required. Moreover, using the asymptotic theory on these estimators (Baíllo and Grané 2009; Ferraty and Nagy 2022) and with similar arguments to those employed in the present paper for (11), asymptotic results for the local linear-type regression estimator could also be derived.

A possible extension of the model assumed in this paper would consist in including additional type of covariates. In a practical situation, this kind of complex models could help to obtain better predictions. The most natural way to address this problem (following our kernel methodology) would be to use product kernels defined on possibly different spaces for the estimators of the sine and cosine components of the response. For instance, the ideas in García-Portugués et al. (2013) for cylindrical density estimation, and the ones in Racine and Li (2004) or in Li and Racine (2004) for categorical data can be employed to define the corresponding weights in this context. This type of product kernel estimators have been recently studied in nonparametric functional regression in Shang (2014) and in Selk and Gertheiss (2022), for models with scalar response and different types of covariates (functional, real-valued or discrete-valued), and they could be directly employed in our circular-functional framework as regression estimators of the sine and cosine models. The important bandwidth selection problem in this context could be accomplished by a cross-validation procedure or using a

Bayesian approach (Shang 2014). We think that a deep study of these extended models is an interesting point that could be addressed in further research.

It should be noticed that the temperature curves considered in our data illustration exhibit temporal dependence. The nonparametric approach for estimating the regression function is not affected by data dependencies, but such an issue should be accounted for when obtaining the asymptotic properties and a more appropriate bandwidth parameter. Considering that the proposed estimator is computed as the `atan2` of nonparametric regression estimators for sine and cosine models, results about nonparametric regression for scalar responses and functional covariate with temporal dependence could be employed to address this issue (Attouch et al. 2010, 2013; Ferraty and Vieu 2006; Masry 2005; Ling et al. 2019; Kurisu 2022). Note also that our real data application considers a single location, but if the data collection presents also a high spatial frequency, it would be possible to proceed with a spatial analysis. For this, an extension to the infinite-dimensional context of the regression model assumed in Meilán-Vila et al. (2021) could be considered, using the results obtained in Ruiz-Medina and Espejo (2012) or in Aguilera-Morillo et al. (2017) for spatial functional data as part of the analysis.

The numerical analysis carried out in this research was performed with the statistical environment R (Development Core Team 2022), using the functions supplied with the `fd.a.usc` package (Febrero-Bande and Oviedo de la Fuente 2012).

**Acknowledgements** The authors acknowledge the support of Meteogalicia for providing the data used in this research. The authors thank Prof. Manuel Oviedo, from Universidade da Coruña, for his help in the use of the `fd.a.usc` package of R. The authors also thank two anonymous referees for numerous useful comments that significantly improved this article.

**Author Contributions** All authors contributed equally to this work.

**Funding** Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. Research of A. Meilán-Vila and M. Francisco-Fernández has been supported by MINECO (Grant MTM2017-82724-R), MICINN (Grant PID2020-113578RB-I00), and by Xunta de Galicia (Grupos de Referencia Competitiva ED431C-2020-14 and Centro de Investigación del Sistema Universitario de Galicia ED431G 2019/01), all of them through the ERDF. Research of R. M. Crujeiras has been supported by MICINN (Grant PID2020-116587GB-I00), and by Xunta de Galicia (Grupos de Referencia Competitiva ED431C-2021-24), all of them through the ERDF.

## Declarations

**Conflict of interest** The authors have no relevant financial or non-financial interests to disclose.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

### Appendix A: Proofs of the theoretical results

**Proof of Lemma 1** For a fixed  $\chi$ , the asymptotic bias and variance of  $\hat{m}_{j,h}(\chi)$ , for  $j = 1, 2$ , can be directly obtained using the asymptotic properties of the Nadaraya–Watson estimator for models with a Euclidean response and a functional covariate (Ferraty et al. 2007).

Regarding the covariance between estimators  $\hat{m}_{1,h}(\chi)$  and  $\hat{m}_{2,h}(\chi)$ , a decomposition can be obtained following similar arguments to those used by Friedlander (1980) (see also Liu 1999) in the finite-dimensional case. In that setting, analytic arguments about the Taylor expansion of the function  $1/z$  around zero were employed and, therefore, the result can be extended to the functional framework. Denoting

$$\begin{aligned}
 R_1 &= \frac{1}{nF_\chi(h)} \sum_{i=1}^n K(h^{-1}\|\mathcal{X}_i - \chi\|) \sin(\Theta_i), \\
 R_2 &= \frac{1}{nF_\chi(h)} \sum_{i=1}^n K(h^{-1}\|\mathcal{X}_i - \chi\|) \cos(\Theta_i) \quad \text{and} \\
 S &= \frac{1}{nF_\chi(h)} \sum_{i=1}^n K(h^{-1}\|\mathcal{X}_i - \chi\|),
 \end{aligned}$$

it follows that

$$\begin{aligned}
 &\text{Cov}[\hat{m}_{1,h}(\chi), \hat{m}_{2,h}(\chi)] \\
 &= \text{Cov} \left[ \frac{\frac{1}{nF_\chi(h)} \sum_{i=1}^n K(h^{-1}\|\mathcal{X}_i - \chi\|) \sin(\Theta_i)}{\frac{1}{nF_\chi(h)} \sum_{i=1}^n K(h^{-1}\|\mathcal{X}_i - \chi\|)}, \right. \\
 &\quad \left. \frac{\frac{1}{nF_\chi(h)} \sum_{i=1}^n K(h^{-1}\|\mathcal{X}_i - \chi\|) \cos(\Theta_i)}{\frac{1}{nF_\chi(h)} \sum_{i=1}^n K(h^{-1}\|\mathcal{X}_i - \chi\|)} \right] \\
 &= \frac{\mu_{R_1}}{\mu_S} \frac{\mu_{R_2}}{\mu_S} \left[ \frac{\sigma_S^2}{\mu_S^2} + \frac{\text{Cov}(R_1, R_2)}{\sigma_{R_1} \sigma_{R_2}} \frac{\sigma_{R_1}}{\mu_{R_1}} \frac{\sigma_{R_2}}{\mu_{R_2}} \right. \\
 &\quad \left. - \frac{\text{Cov}(R_2, S)}{\sigma_{R_2} \sigma_S} \frac{\sigma_{R_2}}{\mu_{R_2}} \frac{\sigma_S}{\mu_S} - \frac{\text{Cov}(R_1, S)}{\sigma_{R_1} \sigma_S} \frac{\sigma_{R_1}}{\mu_{R_1}} \frac{\sigma_S}{\mu_S} \right] + o \left[ \frac{1}{nF_\chi(h)} \right] \\
 &= \frac{\mu_{R_1}}{\mu_S} \frac{\mu_{R_2}}{\mu_S} \left[ \frac{\sigma_S^2}{\mu_S^2} + \frac{\text{Cov}(R_1, R_2)}{\mu_{R_1} \mu_{R_2}} - \frac{\text{Cov}(R_2, S)}{\mu_{R_2} \mu_S} - \frac{\text{Cov}(R_1, S)}{\mu_{R_1} \mu_S} \right] \\
 &\quad + o \left[ \frac{1}{nF_\chi(h)} \right], \tag{A1}
 \end{aligned}$$

where  $\mu_{R_1}$  and  $\sigma_{R_1}^2$ ,  $\mu_{R_2}$  and  $\sigma_{R_2}^2$ , and  $\mu_S$  and  $\sigma_S^2$  denote the expectation and variance of  $R_1$ ,  $R_2$  and  $S$ , respectively. Now, using Lemma 4 and 5 of (Ferraty et al. 2007), it

can be obtained that

$$\mu_{R_1} \rightarrow m_1(\chi)M_{\chi,1}, \tag{A2}$$

$$\mu_{R_2} \rightarrow m_2(\chi)M_{\chi,1}, \tag{A3}$$

$$\mu_S \rightarrow M_{\chi,1}, \tag{A4}$$

and

$$\sigma_S^2 = \frac{M_{\chi,2}}{nF_{\chi}(h)}[1 + o(1)], \tag{A5}$$

$$\text{Cov}(R_1, S) = m_1(\chi) \frac{M_{\chi,2}}{nF_{\chi}(h)}[1 + o(1)], \tag{A6}$$

$$\text{Cov}(R_2, S) = m_2(\chi) \frac{M_{\chi,2}}{nF_{\chi}(h)}[1 + o(1)]. \tag{A7}$$

Moreover, taking into account the continuity of  $m_1, m_2$  and  $c$ , and using the results by Ferraty et al. (2007), p. 283, it follows that

$$\begin{aligned} & \text{Cov}(R_1, R_2) \\ &= \frac{1}{nF_{\chi}^2(h)} \left\{ \mathbb{E} \left[ \sin \Theta \cos \Theta K^2(h^{-1} \|\mathcal{X} - \chi\|) \right] \right. \\ & \quad \left. - \mathbb{E} \left[ \sin \Theta K(h^{-1} \|\mathcal{X} - \chi\|) \right] \mathbb{E} \left[ \cos \Theta K(h^{-1} \|\mathcal{X} - \chi\|) \right] \right\} \\ &= \frac{1}{nF_{\chi}^2(h)} \left\{ [c(\chi) + m_1(\chi)m_2(\chi) + o(1)] \mathbb{E} \left[ K^2(h^{-1} \|\mathcal{X} - \chi\|) \right] \right. \\ & \quad \left. - [m_1(\chi) + o(1)] \mathbb{E} \left[ K(h^{-1} \|\mathcal{X} - \chi\|) \right] [m_2(\chi) + o(1)] \mathbb{E} \left[ K(h^{-1} \|\mathcal{X} - \chi\|) \right] \right\} \\ &= [c(\chi) + m_1(\chi)m_2(\chi)] \frac{M_{\chi,2}}{nF_{\chi}(h)} [1 + o(1)], \tag{A8} \end{aligned}$$

since

$$\begin{aligned} \{\mathbb{E}[K(h^{-1} \|\mathcal{X} - \chi\|)]\}^2 &= \mathcal{O}[F_{\chi}^2(h)], \\ \mathbb{E}[K^2(h^{-1} \|\mathcal{X} - \chi\|)] &= F_{\chi}(h) \left[ K^2(1) - \int_0^1 (K^2)'(s) \tau_{\chi,h}(s) ds \right] \\ &= F_{\chi}(h) M_{\chi,2}. \end{aligned}$$

Therefore, considering (A1) and using (A2)–(A8), it can be directly obtained that

$$\begin{aligned} & \text{Cov}[\hat{m}_{1,h}(\chi), \hat{m}_{2,h}(\chi)] \\ &= \frac{c(\chi)}{nF_{\chi}(h)} \frac{M_{\chi,2}}{M_{\chi,1}^2} + o \left[ \frac{1}{nF_{\chi}(h)} \right]. \end{aligned}$$

□

**Proof of Theorem 1** First, to obtain the bias of  $\hat{m}_h(\chi)$ , given in (11), the function  $\text{atan2}(\hat{m}_{1,h}, \hat{m}_{2,h})$  is expanded in Taylor series around  $(m_1, m_2)$ , where for simplicity,  $\hat{m}_{j,h}$  and  $m_j$  denote  $\hat{m}_{1,h}(\chi)$  and  $m(\chi)$ , respectively, for  $j = 1, 2$ , to get

$$\begin{aligned} & \text{atan2}(\hat{m}_{1,h}, \hat{m}_{2,h}) \\ &= \text{atan2}(m_1, m_2) + \frac{m_2}{m_1^2 + m_2^2}(\hat{m}_{1,h} - m_1) \\ & \quad - \frac{m_1}{m_1^2 + m_2^2}(\hat{m}_{2,h} - m_2) + \frac{m_1 m_2}{(m_1^2 + m_2^2)^2}(\hat{m}_{2,h} - m_2)^2 \\ & \quad - \frac{m_1 m_2}{(m_1^2 + m_2^2)^2}(\hat{m}_{1,h} - m_1)^2 - \frac{m_1^2 - m_2^2}{(m_1^2 + m_2^2)^2}(\hat{m}_{1,h} - m_1)(\hat{m}_{2,h} - m_2) \\ & \quad + \mathcal{O}[(\hat{m}_{1,h} - m_1)^3] + \mathcal{O}[(\hat{m}_{2,h} - m_2)^3]. \end{aligned} \tag{A9}$$

Hence, noting that  $\ell(\chi) = [m_1^2(\chi) + m_2^2(\chi)]^{1/2}$  and taking expectations in the above expression, it follows that

$$\begin{aligned} & \mathbb{E}[\hat{m}_h(\chi)] - m(\chi) \\ &= \frac{m_2(\chi)}{\ell^2(\chi)} \mathbb{E}[\hat{m}_{1,h}(\chi) - m_1(\chi)] - \frac{m_1(\chi)}{\ell^2(\chi)} \mathbb{E}[\hat{m}_{2,h}(\chi) - m_2(\chi)] \\ & \quad + \frac{m_1(\chi)m_2(\chi)}{\ell^4(\chi)} \mathbb{E}\{[\hat{m}_{2,h}(\chi) - m_2(\chi)]^2\} \\ & \quad - \frac{m_1(\chi)m_2(\chi)}{\ell^4(\chi)} \mathbb{E}\{[\hat{m}_{1,h}(\chi) - m_1(\chi)]^2\} \\ & \quad - \frac{m_1^2(\chi) - m_2^2(\chi)}{\ell^4(\chi)} \mathbb{E}\{[\hat{m}_{1,h}(\chi) - m_1(\chi)][\hat{m}_{2,h}(\chi) - m_2(\chi)]\} \\ & \quad + \mathcal{O}\{[\hat{m}_{1,h}(\chi) - m_1(\chi)]^3\} + \mathcal{O}\{[\hat{m}_{2,h}(\chi) - m_2(\chi)]^3\}. \end{aligned}$$

Noting that  $\mathbb{E}[(\hat{m}_{j,h} - m_j)^2] = \text{Var}(\hat{m}_{j,h}) + [\mathbb{E}(\hat{m}_{j,h} - m_j)]^2$ , and using the results in Lemma 1, it is obtained that

$$\begin{aligned} & \mathbb{E}[\hat{m}_h(\chi) - m(\chi)] \\ &= \frac{m_2(\chi)}{\ell^2(\chi)} \varphi'_{1,\chi}(0) \frac{M_{\chi,0}}{M_{\chi,1}} h - \frac{m_1(\chi)}{\ell^2(\chi)} \varphi'_{2,\chi}(0) \frac{M_{\chi,0}}{M_{\chi,1}} h + \mathcal{O}\left[\frac{1}{nF_\chi(h)}\right] + o(h) \\ &= \frac{1}{\ell^2(\chi)} \frac{M_{\chi,0}}{M_{\chi,1}} h [m_2(\chi) \varphi'_{1,\chi}(0) - m_1(\chi) \varphi'_{2,\chi}(0)] + \mathcal{O}\left[\frac{1}{nF_\chi(h)}\right] + o(h). \end{aligned} \tag{A10}$$

Now, expression (A10) can be further simplify, expanding the function  $\text{atan2}$  at  $[m_1(\mathcal{X}), m_2(\mathcal{X})]$  in Taylor series around  $[m_1(\chi), m_2(\chi)]$ , as in (A9). Taking conditional expectations in that expansion, one gets that

$$\begin{aligned}
 & \mathbb{E}[m(\mathcal{X}) - m(\chi) \mid \|\mathcal{X} - \chi\| = s] \\
 &= \frac{m_2(\chi)}{\ell^2(\chi)} \mathbb{E}[m_1(\mathcal{X}) - m_1(\chi) \mid \|\mathcal{X} - \chi\| = s] \\
 &\quad - \frac{m_1(\chi)}{\ell^2(\chi)} \mathbb{E}[m_2(\mathcal{X}) - m_2(\chi) \mid \|\mathcal{X} - \chi\| = s] \\
 &\quad + \frac{m_1(\chi)m_2(\chi)}{\ell^4(\chi)} \mathbb{E}\{[m_2(\mathcal{X}) - m_2(\chi)]^2 \mid \|\mathcal{X} - \chi\| = s\} \\
 &\quad - \frac{m_1(\chi)m_2(\chi)}{\ell^4(\chi)} \mathbb{E}\{[m_1(\mathcal{X}) - m_1(\chi)]^2 \mid \|\mathcal{X} - \chi\| = s\} \\
 &\quad - \frac{m_1^2(\chi) - m_2^2(\chi)}{\ell^4(\chi)} \mathbb{E}\{[m_1(\mathcal{X}) - m_1(\chi)][m_2(\mathcal{X}) - m_2(\chi)] \mid \|\mathcal{X} - \chi\| = s\} \\
 &\quad + \mathcal{O}\{[m_1(\mathcal{X}) - m_1(\chi)]^3\} + \mathcal{O}\{[m_2(\mathcal{X}) - m_2(\chi)]^3\}. \tag{A11}
 \end{aligned}$$

Noting that  $\varphi_\chi(s) = \mathbb{E}[m(\mathcal{X}) - m(\chi) \mid \|\mathcal{X} - \chi\| = s]$  and  $\varphi_{j,\chi}(s) = \mathbb{E}[m_j(\mathcal{X}) - m_j(\chi) \mid \|\mathcal{X} - \chi\| = s]$ , for  $j = 1, 2, \forall s \in \mathbb{R}$ , deriving expression (A11) with respect to  $s$  and replacing the obtained result in (A10), it follows that

$$\mathbb{E}[\hat{m}_h(\chi) - m(\chi)] = \varphi'_\chi(0) \frac{M_{\chi,0}}{M_{\chi,1}} h + \mathcal{O}\left[\frac{1}{nF_\chi(h)}\right] + o(h).$$

In order to derive the variance of the estimator  $\hat{m}_h(\chi)$ , the function  $\text{atan}^2(\hat{m}_{1,h}, \hat{m}_{2,h})$  is expanded in Taylor series around  $(m_1, m_2)$  to obtain that

$$\begin{aligned}
 & \text{atan}^2(\hat{m}_{1,h}, \hat{m}_{2,h}) \\
 &= \text{atan}^2(m_1, m_2) + \frac{2\text{atan}^2(m_1, m_2)m_2}{m_1^2 + m_2^2} (\hat{m}_{1,h} - m_1) \\
 &\quad - \frac{2\text{atan}^2(m_1, m_2)m_1}{m_1^2 + m_2^2} (\hat{m}_{2,h} - m_2) \\
 &\quad + \frac{2\text{atan}^2(m_1, m_2)m_1m_2}{(m_1^2 + m_2^2)^2} (\hat{m}_{2,h} - m_2)^2 \\
 &\quad - \frac{2\text{atan}^2(m_1, m_2)m_1m_2}{(m_1^2 + m_2^2)^2} (\hat{m}_{1,h} - m_1)^2 \\
 &\quad - \frac{2\text{atan}^2(m_1, m_2)(m_1^2 - m_2^2)}{(m_1^2 + m_2^2)^2} (\hat{m}_{1,h} - m_1)(\hat{m}_{2,h} - m_2) \\
 &\quad + \frac{m_1^2}{(m_1^2 + m_2^2)^2} (\hat{m}_{2,h} - m_2)^2 + \frac{m_2^2}{(m_1^2 + m_2^2)^2} (\hat{m}_{1,h} - m_1)^2 \\
 &\quad - \frac{2m_1m_2}{(m_1^2 + m_2^2)^2} (\hat{m}_{1,h} - m_1)(\hat{m}_{2,h} - m_2) \\
 &\quad + \mathcal{O}\left[(\hat{m}_{1,h} - m_1)^3\right] + \mathcal{O}\left[(\hat{m}_{2,h} - m_2)^3\right]. \tag{A12}
 \end{aligned}$$



Taking expectations in the Taylor expansions given in (A9) and (A12), one gets that

$$\begin{aligned}
 & \text{Var}[\hat{m}_h(\chi)] \\
 &= \text{atan}^2[m_{1,h}(\chi), m_{2,h}(\chi)] \\
 &+ \frac{2\text{atan}2[m_1(\chi), m_2(\chi)]m_2(\chi)}{\ell^2(\chi)} \mathbb{E}[\hat{m}_{1,h}(\chi) - m_1(\chi)] \\
 &- \frac{2\text{atan}2[m_1(\chi), m_2(\chi)]m_1(\chi)}{\ell^2(\chi)} \mathbb{E}[\hat{m}_{2,h}(\chi) - m_2(\chi)] \\
 &+ \frac{2\text{atan}2[m_1(\chi), m_2(\chi)]m_1(\chi)m_2(\chi)}{\ell^4(\chi)} \mathbb{E}\{[\hat{m}_{2,h}(\chi) - m_2(\chi)]^2\} \\
 &- \frac{2\text{atan}2[m_1(\chi), m_2(\chi)]m_1(\chi)m_2(\chi)}{\ell^4(\chi)} \mathbb{E}\{[\hat{m}_{1,h}(\chi) - m_1(\chi)]^2\} \\
 &- \frac{2\text{atan}[m_1(\chi), m_2(\chi)][m_1^2(\chi) - m_2^2(\chi)]}{\ell^4(\chi)} \mathbb{E}\{[\hat{m}_{1,h}(\chi) - m_1(\chi)][\hat{m}_{2,h}(\chi) - m_2(\chi)]\} \\
 &+ \frac{m_1^2(\chi)}{\ell^4(\chi)} \mathbb{E}\{[\hat{m}_{2,h}(\chi) - m_2(\chi)]^2\} + \frac{m_2^2(\chi)}{\ell^4(\chi)} \mathbb{E}\{[\hat{m}_{1,h}(\chi) - m_1(\chi)]^2\} \\
 &- \frac{2m_1(\chi)m_2(\chi)}{\ell^4(\chi)} \mathbb{E}\{[\hat{m}_{1,h}(\chi) - m_1(\chi)][\hat{m}_{2,h}(\chi) - m_2(\chi)]\} \\
 &- \left( \text{atan}2[m_1(\chi), m_2(\chi)] + \frac{m_2(\chi)}{\ell^2(\chi)} \mathbb{E}[\hat{m}_{1,h}(\chi) - m_1(\chi)] \right. \\
 &- \frac{m_1(\chi)}{\ell^2(\chi)} \mathbb{E}[\hat{m}_{2,h}(\chi) - m_2(\chi)] + \frac{m_1(\chi)m_2(\chi)}{\ell^4(\chi)} \mathbb{E}\{[\hat{m}_{2,h}(\chi) - m_2(\chi)]^2\} \\
 &- \frac{m_1(\chi)m_2(\chi)}{\ell^4(\chi)} \mathbb{E}\{[\hat{m}_{1,h}(\chi) - m_1(\chi)]^2\} \\
 &- \left. \frac{m_1^2(\chi) - m_2^2(\chi)}{\ell^4(\chi)} \mathbb{E}\{[\hat{m}_{1,h}(\chi) - m_1(\chi)][\hat{m}_{2,h}(\chi) - m_2(\chi)]\} \right)^2 \\
 &+ \mathcal{O}\{[\hat{m}_{1,h}(\chi) - m_1(\chi)]^3\} + \mathcal{O}\{[\hat{m}_{2,h}(\chi) - m_2(\chi)]^3\} \\
 &+ \mathcal{O}\{[\hat{m}_{1,h}(\chi) - m_1(\chi)]^3\} + \mathcal{O}\{[\hat{m}_{2,h}(\chi) - m_2(\chi)]^3\}.
 \end{aligned}$$

By straightforward calculations, it can be obtained that

$$\begin{aligned}
 & \text{Var}[\hat{m}_h(\chi)] \\
 &= \frac{m_1^2(\chi)}{\ell^4(\chi)} \mathbb{E}\{[\hat{m}_{2,h}(\chi) - m_2(\chi)]^2\} + \frac{m_2^2(\chi)}{\ell^4(\chi)} \mathbb{E}\{[\hat{m}_{1,h}(\chi) - m_1(\chi)]^2\} \\
 &- \frac{2m_1(\chi)m_2(\chi)}{\ell^4(\chi)} \mathbb{E}\{[\hat{m}_{1,h}(\chi) - m_1(\chi)][\hat{m}_{2,h}(\chi) - m_2(\chi)]\} \\
 &- \frac{m_2^2(\chi)}{\ell^4(\chi)} \{\mathbb{E}[\hat{m}_{1,h}(\chi) - m_1(\chi)]^2\} - \frac{m_1^2(\chi)}{\ell^4(\chi)} \{\mathbb{E}[\hat{m}_{2,h}(\chi) - m_2(\chi)]^2\} \\
 &+ \frac{2m_1(\chi)m_2(\chi)}{\ell^4(\chi)} \mathbb{E}[\hat{m}_{1,h}(\chi) - m_1(\chi)]\mathbb{E}[\hat{m}_{2,h}(\chi) - m_2(\chi)] \\
 &+ \mathcal{O}\{[\hat{m}_{1,h}(\chi) - m_1(\chi)]^3\} + \mathcal{O}\{[\hat{m}_{2,h}(\chi) - m_2(\chi)]^3\}.
 \end{aligned}$$

So, noting that  $\mathbb{E}[(\hat{m}_{j,h} - m_j)^2] = \text{Var}(\hat{m}_{j,h}) + [\mathbb{E}(\hat{m}_{j,h} - m_j)]^2$ , it can be obtained that the conditional variance is:

$$\begin{aligned} &\text{Var}[\hat{m}_h(\chi)] \\ &= \frac{m_1^2(\chi)}{\ell^4(\chi)} \text{Var}[\hat{m}_{2,h}(\chi)] + \frac{m_2^2(\chi)}{\ell^4(\chi)} \text{Var}[\hat{m}_{1,h}(\chi)] \\ &\quad - \frac{2m_1(\chi)m_2(\chi)}{\ell^4(\chi)} \text{Cov}[\hat{m}_{1,h}(\chi), \hat{m}_{2,h}(\chi)] \\ &\quad + \mathcal{O}\left\{[\hat{m}_{1,h}(\chi) - m_1(\chi)]^3\right\} + \mathcal{O}\left\{[\hat{m}_{2,h}(\chi) - m_2(\chi)]^3\right\}. \end{aligned}$$

Therefore, using Lemma 1, one gets that

$$\begin{aligned} \text{Var}[\hat{m}_h(\chi)] &= \frac{m_1^2(\chi)}{\ell^4(\chi)} \frac{s_2^2(\chi)}{nF_\chi(h)} \frac{M_{\chi,2}}{M_{\chi,1}^2} + \frac{m_2^2(\chi)}{\ell^4(\chi)} \frac{s_1^2(\chi)}{nF_\chi(h)} \frac{M_{\chi,2}}{M_{\chi,1}^2} \\ &\quad - \frac{2m_1(\chi)m_2(\chi)}{\ell^4(\chi)} \frac{c(\chi)}{nF_\chi(h)} \frac{M_{\chi,2}}{M_{\chi,1}^2} + o\left[\frac{1}{nF_\chi(h)}\right]. \end{aligned}$$

Considering equations (7)–(10), and taking into account that  $f_1^2(\chi) + f_2^2(\chi) = 1$ , it follows that

$$\begin{aligned} &m_1^2(\chi)s_2^2(\chi) + m_2^2(\chi)s_1^2(\chi) - 2m_1(\chi)m_2(\chi)c(\chi) \\ &= f_1^2(\chi)f_2^2(\chi)\ell^2(\chi)\sigma_2^2(\chi) - 2f_2(\chi)f_1^3(\chi)\ell^2(\chi)\sigma_{12}(\chi) \\ &\quad + f_1^4(\chi)\ell^2(\chi)\sigma_1^2(\chi) + f_2^2(\chi)f_1^2(\chi)\ell^2(\chi)\sigma_2^2(\chi) \\ &\quad + 2f_2^3(\chi)f_1(\chi)\ell^2(\chi)\sigma_{12}(\chi) + f_2^4(\chi)\ell^2(\chi)\sigma_1^2(\chi) \\ &\quad - 2f_1^2(\chi)f_2^2(\chi)\ell^2(\chi)\sigma_2^2(\chi) + 2f_1^3(\chi)f_2(\chi)\ell^2(\chi)\sigma_{12}(\chi) \\ &\quad - 2f_2^3(\chi)f_1(\chi)\ell^2(\chi)\sigma_{12}(\chi) + 2f_1^2(\chi)f_2^2(\chi)\ell^2(\chi)\sigma_1^2(\chi) \\ &= \ell^2(\chi)\sigma_1^2(\chi). \end{aligned}$$

Therefore,

$$\text{Var}[\hat{m}_h(\chi)] = \frac{1}{nF_\chi(h)} \frac{\sigma_1^2(\chi)}{\ell^2(\chi)} \frac{M_{\chi,2}}{M_{\chi,1}^2} + o\left[\frac{1}{nF_\chi(h)}\right].$$

□

**Proof of Theorem 2** In order to derive the asymptotic distribution of  $\hat{m}_h$ , given in (11), we compute the asymptotic distribution of  $\hat{M} = \hat{m}_{1,h}/\hat{m}_{2,h}$  and apply Theorem A of Serfling (1980). First, using similar arguments to those in Lemma 6 of Ferraty et al. (2007), it can be obtained that

$$\hat{M}(\chi) - \mathbb{E}[\hat{M}(\chi)] = \frac{\hat{m}_{1,h}(\chi)}{\hat{m}_{2,h}(\chi)} - \frac{\mathbb{E}[\hat{m}_{1,h}(\chi)]}{\mathbb{E}[\hat{m}_{2,h}(\chi)]} + o\left[\frac{1}{\sqrt{nF_\chi(h)}}\right].$$

However, the following decomposition holds:

$$\frac{\hat{m}_{1,h}(\chi) - \mathbb{E}[\hat{m}_{1,h}(\chi)]}{\hat{m}_{2,h}(\chi) - \mathbb{E}[\hat{m}_{2,h}(\chi)]} = \frac{\{\hat{m}_{1,h}(\chi) - \mathbb{E}[\hat{m}_{1,h}(\chi)]\}\mathbb{E}[\hat{m}_{2,h}(\chi)]}{\hat{m}_{2,h}(\chi)\mathbb{E}[\hat{m}_{2,h}(\chi)]} + \frac{\{\mathbb{E}[\hat{m}_{2,h}(\chi)] - \hat{m}_{2,h}(\chi)\}\mathbb{E}[\hat{m}_{1,h}(\chi)]}{\hat{m}_{2,h}(\chi)\mathbb{E}[\hat{m}_{2,h}(\chi)]}$$

Therefore,  $\hat{M}(\chi) - \mathbb{E}[\hat{M}(\chi)]$  has the same asymptotic distribution as

$$N(\chi) = \frac{\{\hat{m}_{1,h}(\chi) - \mathbb{E}[\hat{m}_{1,h}(\chi)]\}\mathbb{E}[\hat{m}_{2,h}(\chi)]}{\hat{m}_{2,h}(\chi)\mathbb{E}[\hat{m}_{2,h}(\chi)]} + \frac{\{\mathbb{E}[\hat{m}_{2,h}(\chi)] - \hat{m}_{2,h}(\chi)\}\mathbb{E}[\hat{m}_{1,h}(\chi)]}{\hat{m}_{2,h}(\chi)\mathbb{E}[\hat{m}_{2,h}(\chi)]}.$$

Note that

$$[\hat{m}_{1,h}(\chi)/\hat{m}_{2,h}(\chi)] - \{\mathbb{E}[\hat{m}_{1,h}(\chi)]/\mathbb{E}[\hat{m}_{2,h}(\chi)]\}$$

can be expressed as an array of independent centered random variables and, consequently, the Central Limit Theorem can be applied. Moreover, using the Theorem of Slutsky, it can be obtained that  $N(\chi)$  also follows a normal distribution.

The asymptotic bias and variance of  $\hat{M}$  could be computed by expanding the function  $\hat{m}_{1,h}/\hat{m}_{2,h}$  in Taylor series around  $m_1/m_2$  and using similar steps to those employed to derive  $\mathbb{E}[\hat{m}_h(\chi)]$  and  $\text{Var}[\hat{m}_h(\chi)]$ . Therefore, denoting by  $\mathbf{b}_h$  and  $v$  the leading terms of the asymptotic bias and variance of  $\hat{M}$ , respectively, it follows that, as  $n \rightarrow \infty$ ,

$$\frac{\hat{M}(\chi) - m_1(\chi)/m_2(\chi) - \mathbf{b}_h}{v^{1/2}(\chi)} \xrightarrow{\mathcal{L}} N(0, 1).$$

Finally, applying Theorem A of Serfling (1980) and Theorem 1, it can be obtained that, as  $n \rightarrow \infty$ ,

$$\sqrt{nF_\chi(h)} \frac{\ell(\chi)M_{\chi,1}}{\sqrt{\sigma_1^2(\chi)M_{\chi,2}}} [\hat{m}_h(\chi) - m(\chi) - \mathbf{B}_h] \xrightarrow{\mathcal{L}} N(0, 1).$$

□

**Proof of Corollary 1** This result can be obtained by combining Theorem 2 and assumption (C7). □

**Proof of Corollary 2** Using the Glivenko–Cantelli Theorem, it follows that

$$\frac{\hat{F}_\chi(h)}{F_\chi(h)} \xrightarrow{\mathbb{P}} 1.$$

Therefore, using this result, assumption (C7), and the consistency of estimators  $\hat{\sigma}_1^2$  and  $\hat{l}$ , as well as Theorem 2, Corollary 2 can be directly obtained. □

## References

- Abraham C, Cornillon PA, Matzner-Løber E et al (2003) Unsupervised curve clustering using B-splines. *Scand J Stat* 30(3):581–595
- Aguilera-Morillo MC, Durbán M, Aguilera AM (2017) Prediction of functional data with spatial dependence: a penalized approach. *Stoch Environ Res Risk Assess* 31(1):7–22
- Aneiros-Pérez G, Cao R, Vilar-Fernández JM (2011) Functional methods for time series prediction: a nonparametric approach. *J Forecast* 30(4):377–392
- Aneiros G, Cao R, Fraiman R et al (2019) Recent advances in functional data analysis and high-dimensional statistics. *J Multivar Anal* 170:3–9. Special Issue on Functional Data Analysis and Related Topics
- Aneiros G, Horová I, Hušková M et al (2022) On functional data analysis and related topics. *J Multivar Anal* 189(104):861
- Antoniadis A, Sapatinas T (2007) Estimation and inference in functional mixed-effects models. *Comput Stat Data Anal* 51(10):4793–4813
- Attouch M, Laksaci A, Saïd EO (2010) Asymptotic normality of a robust estimator of the regression function for functional time series data. *J Korean Stat Soc* 39(4):489–500
- Attouch MK, Laksaci A, Saïd EO (2013) Robust regression for functional time series data. *J Jpn Stat Soc* 42(2):125–143
- Baíllo A, Grané A (2009) Local linear regression for functional predictor and scalar response. *J Multivar Anal* 100(1):102–111
- Berlinet A, Elamine A, Mas A (2011) Local linear regression for functional data. *Ann Inst Stat Math* 63:1047–1075
- Bernhardt J, Carleton AM, LaMagna C (2018) A comparison of daily temperature-averaging methods: spatial variability and recent change for the conus. *J Clim* 31(3):979–996
- Bloodhart B, Maibach E, Myers T et al (2015) Local climate experts: the influence of local tv weather information on climate change perceptions. *PLoS ONE* 10(11):e0141526
- Blunden J, Arndt D (2020) State of the climate in 2019. *Bull Am Meteorol Soc* 101(8):S1–S429
- Boj E, Delicado P, Fortiana J (2008) Local linear functional regression based on weighted distance-based regression. In: *Functional and operatorial statistics*. Springer, New York, pp 57–64
- Burba F, Ferraty F, Vieu P (2009) k-nearest neighbour method in functional nonparametric regression. *J Nonparametr Stat* 21:453–469
- Buttarazzi D, Pandolfo G, Porzio GC (2018) A boxplot for circular data. *Biometrics* 74(4):1492–1501
- Cai M, Hu J, Zhou C et al (2021) Mortality burden caused by diurnal temperature range: a nationwide time-series study in 364 Chinese locations. *Stoch Environ Res Risk Assess* 35:1605–1614
- Di Marzio M, Panzera A, Taylor CC (2012) Non-parametric smoothing and prediction for nonlinear circular time series. *J Time Ser Anal* 33(4):620–630
- Di Marzio M, Panzera A, Taylor CC (2013) Non-parametric regression for circular responses. *Scand J Stat* 40(2):238–255
- Febrero-Bande M, Oviedo de la Fuente M (2012) Statistical computing in functional data analysis: the R package fda.usc. *J Stat Softw* 51(4):1–28
- Febrero-Bande M, Galeano P, González-Manteiga W (2017) Functional principal component regression and functional partial least-squares regression: an overview and a comparative study. *Int Stat Rev* 85(1):61–83
- Ferraty F, Nagy S (2022) Scalar-on-function local linear regression and beyond. *Biometrika* 109(2):439–455
- Ferraty F, Vieu P (2002) The functional nonparametric model and application to spectrometric data. *Comput Stat* 17(4):545–564
- Ferraty F, Vieu P (2006) *Nonparametric functional data analysis: theory and practice*. Springer, New York
- Ferraty F, Mas A, Vieu P (2007) Nonparametric regression on functional data: inference and practical aspects. *Aust N Z J Stat* 49(3):267–286
- Fisher NI (1995) *Statistical analysis of circular data*. Cambridge University Press, Cambridge
- Fisher NI, Lee AJ (1992) Regression models for an angular response. *Biometrics* 48(3):665–677
- Friedlander LJ (1980) A study of correlation between ratio variables. PhD thesis, University of Washington
- García-Portugués E, Crujeiras RM, González-Manteiga W (2013) Kernel density estimation for directional-linear data. *J Multivar Anal* 121:152–175
- Gasser T, Hall P, Presnell B (1998) Nonparametric estimation of the mode of a distribution of random curves. *J R Stat Soc Ser B* 60(4):681–691

- Glynis KG, Iliopoulou T, Dimitriadis P et al (2021) Stochastic investigation of daily air temperature extremes from a global ground station network. *Stoch Environ Res Risk Assess* 35:1585–1603
- Goebbert K, Jenkins-Smith HC, Klockow K et al (2012) Weather, climate, and worldviews: the sources and consequences of public perceptions of changes in local weather patterns. *Weather Clim Soc* 4(2):132–144
- Goia A, Vieu P (2016) An introduction to recent advances in high/infinite dimensional statistics. *J Multivar Anal* 146:1–6. Special Issue on Statistical Models and Methods for High or Infinite Dimensional Spaces
- Greven S, Scheipl F (2017) A general framework for functional regression modelling. *Stat Model* 17(1–2):1–35
- Härdle W, Müller M (2012) Multivariate and semiparametric kernel regression, chap 12. Wiley, New York, pp 357–391
- Howe PD (2018) Perceptions of seasonal weather are linked to beliefs about global climate change: evidence from Norway. *Clim Change* 148(4):467–480
- Huld TA, Šúri M, Dunlop ED et al (2006) Estimating average daytime and daily temperature profiles within Europe. *Environ Modell Softw* 21(12):1650–1661
- Jona-Lasinio G, Gelfand A, Jona-Lasinio M (2012) Spatial analysis of wave direction data using wrapped Gaussian processes. *Ann Appl Stat* 6(4):1478–1498
- Kim S, SenGupta A (2017) Multivariate-multiple circular regression. *J Stat Comput Simul* 87(7):1277–1291
- Kokoszka P, Reimherr M (2017) Introduction to functional data analysis, 1st edn. Chapman and Hall/CRC, New York
- Kurusu D (2022) Nonparametric regression for locally stationary functional time series. *Electron J Stat* 16(2):3973–3995
- Ley C, Verdebout T (2017) Modern directional statistics. Chapman and Hall/CRC, Boca Raton
- Li Q, Racine J (2004) Cross-validated local linear nonparametric regression. *Stat Sin* 14(2):485–512
- Ling N, Vieu P (2018) Nonparametric modelling for functional data: selected survey and tracks for future. *Statistics* 52(4):934–949
- Ling N, Meng S, Vieu P (2019) Uniform consistency rate of knn regression estimation for functional time series data. *J Nonparametr Stat* 31(2):451–468
- Ling N, Wang L, Vieu P (2020) Convergence rate of kernel regression estimation for time series data when both response and covariate are functional. *Metrika* 83(6):713–732
- Liu Y (1999) The statistical validity of using ratio variables in human kinetics research. PhD thesis, The University of British Columbia
- Ma Y, Guttorp P (2013) Estimating daily mean temperature from synoptic climate observations. *Int J Climatol* 33(5):1264–1269
- Mardia KV, Jupp PE (2000) Directional statistics. Wiley, Chichester
- Masry E (2005) Nonparametric regression estimation for dependent functional data: asymptotic normality. *Stoch Process Their Appl* 115(1):155–177
- Meilán-Vila A, Crujeiras RM, Francisco-Fernández M (2021a) Nonparametric estimation of circular trend surfaces with application to wave directions. *Stoch Environ Res Risk Assess* 35:923–939
- Meilán-Vila A, Francisco-Fernández M, Crujeiras RM et al (2021b) Nonparametric multiple regression estimation for circular response. *TEST* 30:650–672
- Morris JS (2015) Functional regression. *Annu Rev Stat Appl* 2:321–359
- Pörtner HO, Roberts D, Tignor M et al (2022) Climate change 2022: impacts, adaptation and vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge. <https://doi.org/10.1017/9781009325844>
- Presnell B, Morrison SP, Littell RC (1998) Projected multivariate linear models for directional data. *J Am Stat Assoc* 93(443):1068–1077
- R Development Core Team (2022) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna <http://www.R-project.org>
- Racine J, Li Q (2004) Nonparametric estimation of regression functions with both categorical and continuous data. *J Econom* 119(1):99–130
- Ramsay JO, Silverman B (2005) Functional data analysis. Springer, New York
- Ruiz-Medina M, Espejo R (2012) Spatial autoregressive functional plug-in prediction of ocean surface temperature. *Stoch Environ Res Risk Assess* 26(3):335–344

- Scapini F, Aloia A, Bouslama MF et al (2002) Multiple regression analysis of the sources of variation in orientation of two sympatric sandhoppers, *Talitrus saltator* and *Talorchestia brito*, from an exposed mediterranean beach. *Behav Ecol Sociobiol* 51(5):403–414
- Selk L, Gertheiss J (2022) Nonparametric regression and classification with functional, categorical, and mixed covariates. *Adv Data Anal Classif*. <https://doi.org/10.1007/s11634-022-00513-7>
- Serfling RJ (1980) Approximation theorems of mathematical statistics, vol 162. Wiley, New York
- Shang HL (2014) Bayesian bandwidth estimation for a functional nonparametric regression model with mixed types of regressors and unknown error density. *J Nonparametr Stat* 26(3):599–615
- Sun Y, Genton MG (2011) Functional boxplots. *J Comput Graph Stat* 20(2):316–334
- Taylor A, de Bruin WB, Dessai S (2014) Climate change beliefs and perceptions of weather-related changes in the united kingdom. *Risk Anal* 34(11):1995–2004
- Vilar-Fernández J, Francisco-Fernández M (2006) Nonparametric estimation of the conditional variance function with correlated errors. *J Nonparametr Stat* 18:375–391
- Wang F, Gelfand AE, Jona-Lasinio G (2015) Joint spatio-temporal analysis of a linear and a directional variable: space-time modeling of wave heights and wave directions in the Adriatic Sea. *Stat Sin* 25(1):25–39

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.