Journal of Information and Computing Science
J. Info. Comput. Sci., Vol. 18, No. 1, pp. 1-14

DOI: 10.4208/JICS-2023-001
June 2023

Weighted Supervised Functional Principal Components Analysis

Zewen Zhang[1], Chunzheng Cao[1,*] and Shuren Cao[1]

[1] *School of Mathematics and Statistics, Nanjing University of Information Science and Technology, Nanjing 210044, China*

Abstract. In functional linear regression, a supervised version of functional principal components analysis (FPCA) can automatically estimate the leading functional principal components (FPCs), which not only represent the major source of variation of the functional predictor but also are simultaneously correlated with the response. However, the existing supervised FPCA (sFPCA) is only applicable to single modal functional data. In this paper, we propose a weighted version of supervised FPCA (w-sFPCA) by considering the adaptive weighting of multi-modal functional predictors. The new w-sFPCA not only assigns corresponding weights to each modal of functional predictors, but also automatically estimates the leading FPCs associated with response variables, representing the main sources of variation in functional predictors. The method is demonstrated to have a better prediction accuracy than the conventional sFPCA method by using one real application on meteorological data and two carefully designed simulation studies.

AMS subject classifications: 62H12, 62G05

Key words: Functional data analysis, Functional regression, Supervised learning, Classification.

1 Introduction

Functional data become widely used in various fields such as chemometrics, climatology, economy, image analysis, linguistics, meteorology and other areas. As a consequence, there is a recent interest in methods dealing with functional data that include functional principal component analysis and functional linear models [1–3]. In the study of meteorological problems, functional statistical downscaling method is a new tool to catch the statistical relationship between some regional predictable factors and the large-scale circulation data from Global climate model (GCM) [4,5].

*Corresponding author. *Email addresses:* caochunzheng@nuist.edu.cn (C. Cao)

We study the following functional linear model [6]

$$Y=\beta_0+\int\beta(t)[X(t)-\mu(t)]\mathrm{d}t+\varepsilon, \tag{1.1}$$

where Y is a regional scalar response, β_0 is the intercept, $\mu(t)$ is the mean function, and $X(t)$ stands for the functional large-scale predict factors. Model (1.1) has been used to predict climate change scenarios in the future by linking a scalar response with an integral form of a functional predictor [7].

Due to the presence of response variables, functional predictor and response variables should have correlations [8, 9]. Functional partial least-squares regression is a important method to solve this problem [10]. [11] proposed supervised sparse functional principal component, which incorporate supervision information to recover the true functional principal components (FPCs). An supervised functional principal components analysis (sFPCA) has been proposed [12], which is able to borrow the information from the response variable Y to estimate the leading FPCs. The estimated FPCs have a better performance in functional regression work. Supervised FPCA can be applied to two-dimensional functional data [13] as well.

However, a common limitation of the above methods is that when the functional predictors have different levels, the information form all these levels are totally separated from each other [14]. For example, when predicting precipitation, many studies may consider adding one or several humidity factors instead of considering the atmospheric circulation as a predictor [15]. They may adding the specific humidity (hus) of 850hPa or 500hPa only, when the factor has different observations on 17 pressure level [16, 17]. This may lead to the loss of information about this preparatory factor. For another example, for predictors which are observed at only one pressure level, it may have simulated data from different global climate models. Existing studies often use only a single global climate model [18, 19]. There is great uncertainty in the use of climate models for future scenario projections, so statistical downscaling simulations and projections using a single climate model can no longer meet the needs of climate impact assessment. Ensemble methods are beginning to be applied by meteorologists to statistical downscaling to reduce the uncertainty of the projections, thereby improving the level of regional climate prediction [20–22].

In this paper, we proposed a weighted supervised functional principle components analysis (w-sFPCA) method. Our target is to study different levels of predictors at the same time. We use weight parameters to find the relationships between these levels to obtain better prediction performance. The novelty of the paper is twofold. Firstly, we propose a framework to calculate weights to consider two or more types of functional predictors and utilize the scalar response variable, to obtain better prediction performance. Secondly, our estimation algorithm is based on eigenvalue decomposition which is much easier to implement.

The rest of the paper is organized as follows. Details of our method is described in Section 2. Two carefully designed simulation studies are used to evaluate the finite sample performance of our proposed method in Section 3. Then we show one real data application in Section 4. Section 5 provides concluding remarks.

2 Method

2.1 The w-sFPCA model

Consider n independent samples $\{X_i(t)|i=1,\dots,n\}$, generated from an L^2 stochastic process $X(t)$ at a sequence of random points on the domain $[0,\ T]$ with measurement errors. Under the classical FPCA, there is

$$X(t)=\mu(t)+\sum_{k=1}^{\infty}\tau_k\phi_k(t), \tag{2.1}$$

where $\mu(t)=\mathrm{E}[X(t)]$ is the mean function, and $\{\phi_k(t)|k=1,\dots\}$ are the eigenfunctions of covariance function $G(s,t)=\mathrm{Cov}(X(s),X(t))$ $(t,\ s\in[0,\ T])$, which are called FPCs. The FPC score τ_k are uncorrelated random variables with mean 0 and variance λ_k. Assuming that there is an orthogonal expansion of G in terms of eigenfunctions ϕ_k and decreasing eigenvalues λ_k:

$$G(s,t)=\sum_{k=1}^{\infty}\lambda_k\phi_k(s)\phi_k(t). \tag{2.2}$$

The corresponding τ_k's are called FPC scores, which are uncorrelated random variables with mean 0 and variance λ_k, where $\lambda_k\geq\lambda_{k+1}$ and $\sum_k\lambda_k<\infty$. Usually, $X(t)$ is approximated by the first K leading FPCs:

$$X(t)=\mu(t)+\sum_{k=1}^{K}\tau_k\phi_k(t). \tag{2.3}$$

It can be calculated as $\tau_k=\int_{\mathcal{T}}(X(t)-\mu(t))\phi_k(t)\mathrm{d}t$. For the rest of this paper, we assume $\mu(t)\equiv 0$ without loss of generality. Then it reduces to

$$X(t)=\sum_{k=1}^{K}\tau_k\phi_k(t)=\boldsymbol{\tau}^\top\boldsymbol{\phi}(t), \tag{2.4}$$

in which $\boldsymbol{\tau}=(\tau_1,...,\tau_K)^\top$ and $\boldsymbol{\phi}(t)=(\phi_1(t),...,\phi_K(t))^\top$.

When each subject has q observations, i.e., $\boldsymbol{X}_i(t)=(X_i^{(1)}(t),\dots,X_i^{(q)}(t))^\top$, a weighted stochastic processes $W_i(t)$ can be expressed as

$$W_i(t)=\sum_{j=1}^{q}w_jX_i^{(j)}(t)=\boldsymbol{w}^\top\boldsymbol{X}_i(t), \tag{2.5}$$

in which $\boldsymbol{w}=(w_1,...,w_q)^\top$.

Substituting (2.5) into (1.1) forms an weighted scalar-on-function linear regression

$$Y=\beta_0+\int_{\mathscr{T}}\beta(t)W(t)\mathrm{d}t+\varepsilon. \tag{2.6}$$

The functional slope $\beta(t)$ can be written in terms of $\phi_1,...,\phi_K$ as

$$\beta(t)=\sum_{k=1}^{K}b_k\phi_k(t)=\boldsymbol{b}^\top\boldsymbol{\phi}(t), \tag{2.7}$$

where $\boldsymbol{b}=(b_1,...,b_K)^\top$ is an unknown coefficient vector to be estimated from the data.

The leading p functional principal components (FPCs) $\hat{\phi}_1(t),\ldots,\hat{\phi}_p(t)$ can be estimated by maximizing

$$Q(\phi)=\frac{\theta\langle\phi,\hat{\mathscr{C}}\phi\rangle+(1-\theta)\mathrm{cov}^2(Y,\langle W(t),\phi\rangle)}{\|\phi\|_\lambda^2}, \tag{2.8}$$

subject to $\|\phi\|_\lambda=1$, $\langle\phi_i,\hat{\phi}_j\rangle=0$, for $i\neq j$, and $0\leq\theta\leq 1$. Here $\|\phi\|_\lambda=\sqrt{\|\phi\|^2+\lambda\|\mathscr{D}^2\phi\|^2}$ with $\mathscr{D}^2\phi=\int_{\mathscr{T}}\phi''(t)\mathrm{d}t$, the norm $\|\phi\|=\sqrt{\|\phi\|^2}=\sqrt{\langle\phi,\phi\rangle}$, $\langle f,g\rangle$ denotes the usual $\mathscr{L}^2$ inner product $\langle f,g\rangle=\int_{\mathscr{T}}f(t)g(t)\mathrm{d}t$ and $\hat{\mathscr{C}}$ denotes the empirical covariance operator [12].

The w-sFPCA is suitable for both functional regression and classification. Specifically, the value of $\mathrm{cov}^2(Y,\langle W(t),\phi\rangle)$ in equation (2.8) is computed as $\langle W(t),\phi\rangle\boldsymbol{L}\langle W(t),\phi\rangle^\top$. For scalar-on-function regression, the matrix $\boldsymbol{L}$ is equal to $\boldsymbol{Y}\boldsymbol{Y}^\top$. For multi-classification, the entries of the matrix $\boldsymbol{L}$ are computed as follows: $L_{ij}=0$ when $Y_i\neq Y_j$, $L_{ij}=1/n_k$ when $Y_i=Y_j=k$, where n_k is the sample number of the k-th class.

2.2 Parameter estimation

Refer to sFPCA, we can estimate the smooth supervised FPCs $\boldsymbol{\phi}(t)$ given a set of value for $(\theta,\,\lambda)$. In this subsection, we give the computational details on how to estimate the parameters $\boldsymbol{w}$. We consider to minimize

$$(\boldsymbol{Y}-\int_{\mathscr{T}}\boldsymbol{W}(t)\beta(t)\mathrm{d}t)^\top(\boldsymbol{Y}-\int_{\mathscr{T}}\boldsymbol{W}(t)\beta(t)\mathrm{d}t). \tag{2.9}$$

We update $\boldsymbol{w}$ and $\beta(t)$ based on Lagrange multiplier method, which can be obtained by minimizing

$$(\boldsymbol{Y}-\boldsymbol{Z}\boldsymbol{w})^\top(\boldsymbol{Y}-\boldsymbol{Z}\boldsymbol{w}), \tag{2.10}$$

subject to $\mathbf{1}_q^\top\boldsymbol{w}=1$, where

$$\boldsymbol{Z}=\int_{\mathscr{T}}\boldsymbol{X}(t)\beta(t)dt. \tag{2.11}$$

The detailed algorithm is outlined as follows:

(1) Set the initial value of $\boldsymbol{w}^{(0)}$, which satisfies $\mathbf{1}_q^\top\boldsymbol{w}^{(0)}=1$;

(2) Given the current value of $\boldsymbol{X}(t)$, Y, we can obtain the estimated $\beta^{(l)}(t)$ and $\boldsymbol{\phi}^{(l)}(t)$ by sFPCA [12].

(3) Calculate $\boldsymbol{Z}$ (Equation (2.11)) to update $\boldsymbol{w}^{(l)}$ to $\boldsymbol{w}^{(l+1)}$ by minimizing

$$(\boldsymbol{Y}-\boldsymbol{Z}\boldsymbol{w}^{(l+1)})^\top(\boldsymbol{Y}-\boldsymbol{Z}\boldsymbol{w}^{(l+1)}), \tag{2.12}$$

subject to $\mathbf{1}_q^\top \boldsymbol{w}^{(l+1)}=1$. Considering Lagrange multiplier method, there is

$$f(\boldsymbol{w},\zeta)=1/2(\boldsymbol{Y}-\boldsymbol{Z}\boldsymbol{w})^\top(\boldsymbol{Y}-\boldsymbol{Z}\boldsymbol{w})+\zeta(\mathbf{1}_q^\top\boldsymbol{w}-1). \tag{2.13}$$

Calculate the partial derivative of $\boldsymbol{w}$ according equation (2.13) and make it be zero:

$$\frac{\partial f}{\partial \boldsymbol{w}}=\boldsymbol{Z}^T(Y-\boldsymbol{Z}\boldsymbol{w})+\zeta^{(l+1)}\mathbf{1}_q=0, \tag{2.14}$$

then we have

$$\boldsymbol{w}^{(l+1)}=\left(\boldsymbol{Z}^\top\boldsymbol{Z}\right)^{-1}\left(\boldsymbol{Z}^\top\boldsymbol{Y}+\zeta^{(l+1)}\mathbf{1}_q\right). \tag{2.15}$$

Multiply both sides of the equation by $\mathbf{1}_q^\top$ to update the value of $\zeta^{(l+1)}$:

$$\mathbf{1}_q^\top\left(\boldsymbol{Z}^\top\boldsymbol{Z}\right)^{-1}\left(\boldsymbol{Z}^\top Y+\zeta^{(l+1)}\mathbf{1}_q\right)=1. \tag{2.16}$$

Then we obtain the updated value of $\boldsymbol{w}$ by

$$\boldsymbol{w}^{(l+1)}=\left(\boldsymbol{Z}^\top\boldsymbol{Z}\right)^{-1}\left(\boldsymbol{Z}^\top\boldsymbol{Y}+\frac{1-\mathbf{1}_q^\top\left(\boldsymbol{Z}^\top\boldsymbol{Z}\right)^{-1}\boldsymbol{Z}^\top\boldsymbol{Y}}{\mathbf{1}_q^\top\left(\boldsymbol{Z}^\top\boldsymbol{Z}\right)^{-1}\mathbf{1}_q}\mathbf{1}_q\right). \tag{2.17}$$

(4) Repeat Steps 2 and 3 until convergence.

We transform multimodal functional covariates into a unimodal functional covariate through weighting. It is because both the weighting parameter and functional principal components (FPCs) are all need to be estimated. We obtain their estimated values by alternating iterations. We estimate the FPCs of the weighting functional covariate. Then we use the estimated FPCs to update the weights. These two steps are implemented alternately.

3 Simulation studies

Different simulations are designed to evaluate the proposed w-sFPCA. We briefly introduce the generation mechanism for the functional predictor $X(t)$ in the beginning of this section. Then we discuss each simulation in details. We use FPCs to generate samples.

They are the leading FPCs estimated based on Canadian weather data, which includes daily temperature measurements from 35 weather stations across Canada [14].

The functional predictor $X_i(t), i=1,\ldots,n$, is simulated as: $X_i(t_q)=w_1 x_{i1}(t_q)+w_2 x_{i2}(t_q)$, q=1, …, 365. Then each functional predictor $X_i(t), i=1,...,n$ is simulated as $X_{ij}(t_q)=\tau_{1ij}\phi_1(t_q)+\tau_{2ij}\phi_2(t_q)+\tau_{3ij}\phi_3(t_q)+\tau_{4ij}\phi_4(t_q)$, where $\phi_k(t_q)$ is the k-th true FPCs, $k=1,\ldots,4$. The simulated FPC score is simulated as $\boldsymbol{\tau}_{ij}^{\top}=(\tau_{1ij},\tau_{2ij},\tau_{3ij},\tau_{4ij})^{\top}$, where $\boldsymbol{\tau}_{ij}\sim$MVN($\mathbf{0}$, Σ_j), in which Σ_1 = diag(100, 80, 50, 30) and Σ_2 = diag(120, 70, 45, 20).

The response variable Y is generated using the functional linear regression model (1.1) with $\beta(t)$ being specified as $\beta(t)=b_1\phi_1(t)+b_2\phi_2(t)\phi_j(t)+b_3\phi_3(t)+b_4\phi_4(t)=\boldsymbol{b}^{\top}\boldsymbol{\phi}(t)$, where $\boldsymbol{b}=(b_1,b_2,b_3,b_4)^{\top}$ and $\boldsymbol{\phi}(t)$=($\phi_1(t)$, $\phi_2(t)$, $\phi_3(t)$, $\phi_4(t))^{\top}$. Without loss of generality we set $\beta_0=0$.

3.1 The first simulation study

The first simulation study is designed to evaluate the proposed method when the response variable is binary. We generate 1000 sample curves, $X_i(t), i=1,...,1000$. The response variable Y is generated as:

$$\begin{aligned} &Y_i \sim Bernoulli(p_i), \\ &logit(p_i)=\int_{\mathscr{T}}\beta(t)X_i(t)\mathrm{d}t, i=1,...,1000, \end{aligned} \tag{3.1}$$

in which $\beta(t)=\phi_4(t)$. That is the binary response Y is only related to the fourth FPC $\phi_4(t)$. We randomly select 800 samples as the training set and used the other 200 samples as test set. For un-weighted sFPCA, we let one of the predictors have a weight of 0. For weighted sFPCA, set the weight to $w^0=(0.25,0.75)$ or $w^0=(0.75,0.25)$. For un-weighted sFPCA, weight parameters w^0 can be considered as (0, 1) or (1, 0). The smoothing and weighting parameters of the sFPCA are both selected by five-fold cross-validation.

We compare the prediction performance of w-sFPCA with un-weighted sFPCA in terms of classification accuracy on the test data in 100 simulation. Figure 1 summarizes the classification accuracy estimated by weighted sFPCA and un-weighted sFPCA. As shown in Figure 1(a), when the weight parameters of the weighted sFPCA are set to (0.25, 0.75), its average classification accuracy reaches 0.671, and the highest of classification accuracy is more than 0.75. For the un-weighted sFPCA, when we only use the first predictor, the maximum classification accuracy is slightly higher than 0.725, and the average classification accuracy is 0.66. Also, the deviation range is large. When we only use the second predictor, neither the average classification accuracy nor the highest classification accuracy is not more than 0.6. As can be seen from Figure 1(b), when the weight parameters of the weighted sFPCA are set to (0.75, 0.25), its average classification accuracy reaches 0.675, and the highest is more than 0.735. For the un-weighted sFPCA, when only the first predictor is used, the maximum is slightly higher than 0.735, and the average classification accuracy is 0.655, which is a large deviation range. When only the

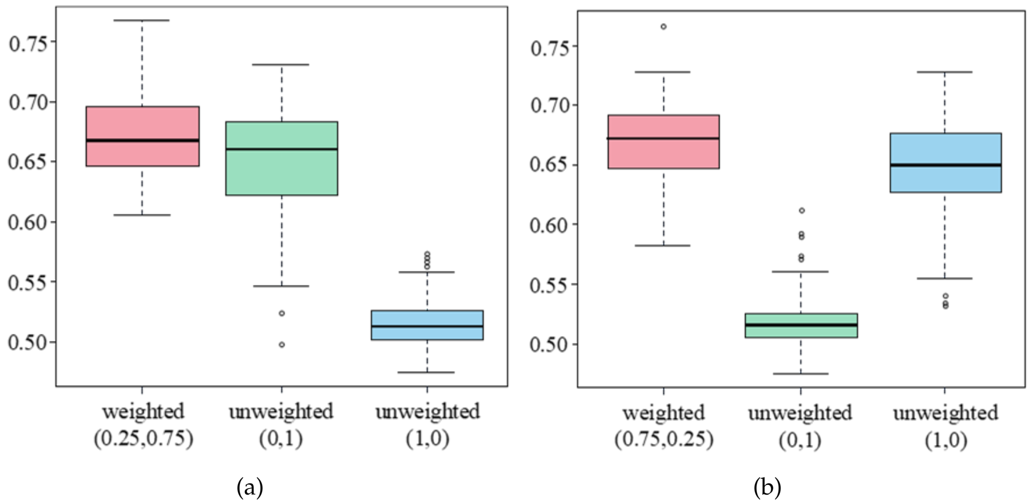

Figure 1: Boxplots of the predicted classification accuracy for 100 simulation runs when using the first 2 FPC scores estimated by weighted sFPCA and un-weighted sFPCA. (a):w^0=(0.25, 0.75), (0, 1) and (1, 0). (b):w^0=(0.75, 0.25), (0, 1) and (1, 0).

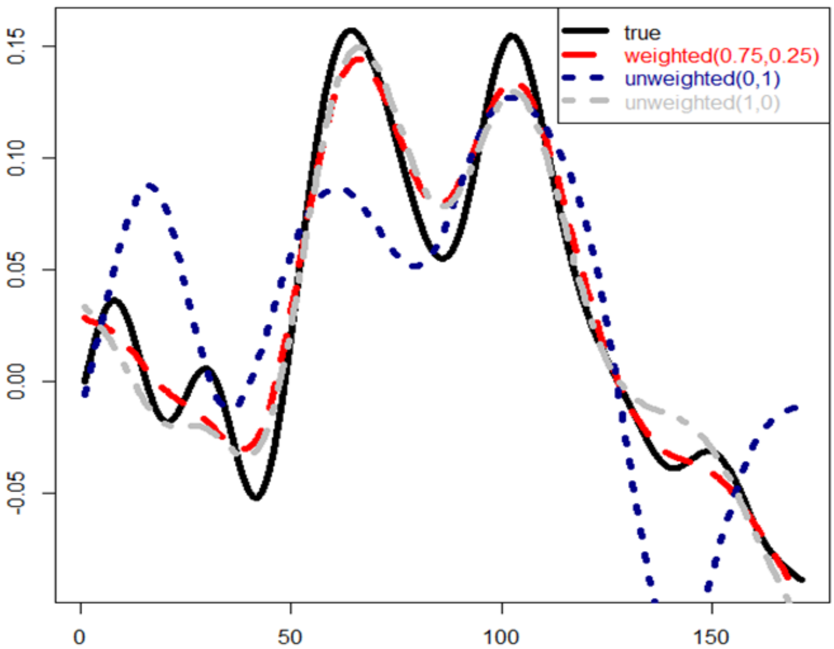

Figure 2: The leading FPC estimated by the weighted and un-weighted sFPCA of the first study.

second predictor is used, neither the average classification accuracy nor the maximum prediction accuracy is not more than 0.575.

To gain some insight, in Figure 2, we compare the leading FPC estimated by the weighted and un-weighted sFPCA in one simulation run, along with the true FPC used to simulate the response variable. We can see that the first FPC estimated by w-sFPCA method is much closer to the true FPC.

3.2 The second simulation study

The simulation data $X(t)$ is generated in the same way as the first simulation study. This section discusses the performance of the proposed method when the response variables are continuous. 800 samples are randomly selected as the training set, and the remaining 200 samples are used as the test set. We compare the prediction performance of weighted and un-weighted sFPCA through 100 simulation runs. The prediction error is using the mean square error (RMSE). It is defined as:

$$\mathrm{RMSE} = \frac{\sum_{\ell=1}^{n} (\hat{y}_\ell - y_\ell)^2}{\sum_{\ell=1}^{n} (\bar{y} - y_\ell)^2}. \tag{3.2}$$

Here y_ℓ and $\hat{y}_\ell$ denote the observed ℓ th response in the test set, respectively, and $\bar{y}$ represents the average of those observed responses the training set.

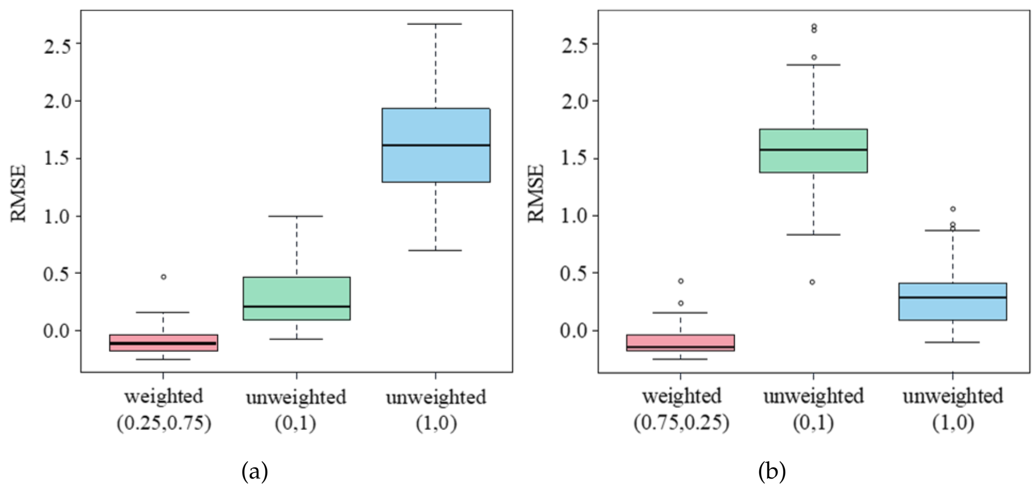

Figure 3: Boxplots of RMSE for 100 simulation runs when using the first 2 FPC scores estimated by weighted sFPCA and un-weighted sFPCA. (a):w^0=(0.25, 0.75), (0, 1) and (1, 0). (b):w^0=(0.75, 0.25), (0, 1) and (1, 0).

Figure 3 summarizes the prediction RMSEs for 100 repeated runs. It can be found that the prediction mean square error of the weighted sFPCA is significantly smaller than that of the un-weighted sFPCA, regardless of the weight setting. As can be seen from

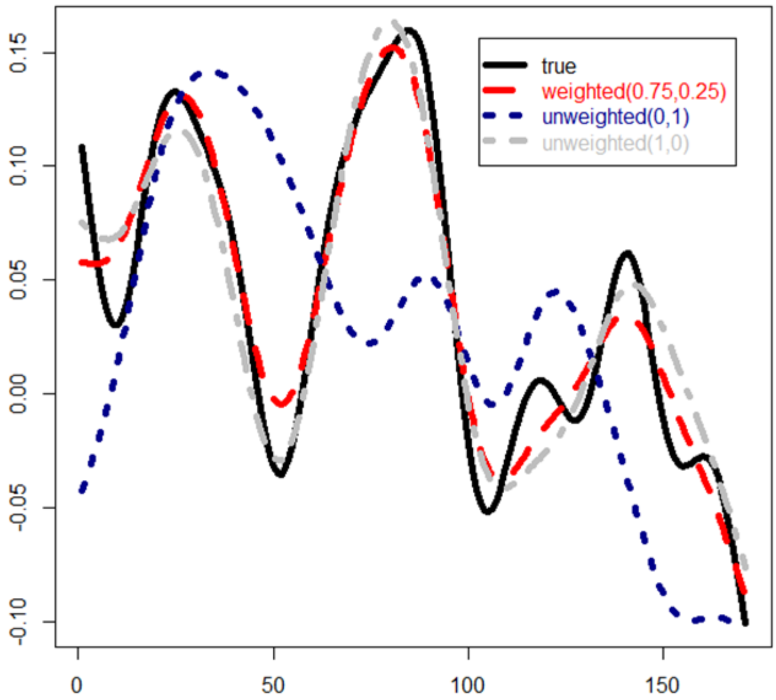

Figure 4: The leading FPC estimated by the weighted and un-weighted sFPCA of the second study.

Figure 3(a), when the weight parameters of the weighted sFPCA are set to (0.25, 0.75), the maximum prediction mean square error is less than 0.5, and the average prediction mean square error is less than 0.25. For the un-weighted sFPCA, when only the first predictor is used, the minimum prediction mean square error is also more than 0.6, and the average prediction mean square error is more than 1.7. Also, the deviation range is large. When only the second predictor is used, the average prediction mean square error is close to 0.5, and the deviation range is also large. As shown in Figure 3(b), when the weight parameters of weighted sFPCA are set to (0.75, 0.25), the prediction mean square error is less than 0.5, and the average prediction mean square error is less than 0.2. For un-weighted sFPCA, when only the first predictor is used, the mean square error of the average prediction exceeds 0.5, which is a large deviation range. When only the second predictor is used, the lowest mean square error of the prediction is also more than 1, and the average mean square error of the prediction is higher than 1.5, with a wide range of deviations.

To gain some insight, Figure 4 displays the true FPC estimated by weighted and un-weighted sFPCA. We can see the first FPC estimated by w-sFPCA is much more closer to the true FPC compared with the first FPC estimated by un-weighted sFPCA. This indicates that w-sFPCA is able to detect the FPC that is truly related to the continuous response variable.

4 Application

In this section, we use the proposed method to analyze the summer temperature in the Yangtze-Huaihe region. When we do data predictions, the selection of climate prediction factors has an extremely important impact on the construction of the model, and it has the ability to determine the feather of future climate prediction scenarios in the studied area. Therefore, the selected factors should have an important physical relationship with the climate under study, be strongly correlated with the local climate, and be able to be accurately simulated by the model. In addition, there must be a weak correlation between factors. Firstly, the meteorological prediction factors are selected to construct a weighted sFPCA of monthly temperature in the Yangtze-Huaihe region, and the robustness of the model is tested based on independent verification.

In order to determine the combination of prediction factors with the best statistical down-scaling effect, multiple groups of factor combination down-scaling experiments are carried out. According to Figure 5, four factors with significant and strong correlation between summer temperature in the Yangtze-Huaihe region are selected, they are 850hPa air temperature, 500hPa geopotential altitude, 500hPa zonal wind and sea level pressure. The prediction performance of each group of experiments will be compared by the root mean square error (RMSE) and the time correlation coefficient (Cor) between the prediction results and the observations. Table 1 and 2 shows the effects of different factor combination model from 1961 to 2007. Except for the 850hPa air temperature, the RMSE is greater than or equal to 0.51°C in all other one-factor tests, and the correlation coefficient is also relatively small, all of which are below 0.8, and fail the test with a significance level of 0.01. When the estimation factor is a combination of multiple factors, the corresponding RMSE decreases, and when the combination of 850hPa air temperature and 500hPa geopotential height field is used as the estimation factor, the minimum RMSE is 0.32°C and the large correlation coefficient is 0.9. Therefore, the combination of 500hPa height field and 850hPa temperature field is used as the prediction factor for the statistical down-scaling of summer temperature in the Yangtze-Huaihe region.

Table 1: The performance of different factor combination model from 1961 to 2007. a, h, u, and s represent the site's average results of the 850hPa air temperature field, 500hPa geopotential height field, 500hPa zonal wind field, and the sea level pressure field.

factor combination	a	h	u	s	a+h	a+u
Cor	0.76	0.72	0.68	0.70	0.90	0.81
RMSE	0.51	0.60	0.55	0.56	0.32	0.42

We use a 500hPa height field and an 850hPa temperature field as prediction factors. Considering that the 850hPa height field and 850hPa temperature field also have a high correlation with the summer temperature in the Yangtze-Huaihe region (Figure 5), for the weighted sFPCA, the 500hPa height field, the 850hPa height field, the 500hPa temperature field and the 850hPa temperature field are used as prediction factors. We calculate

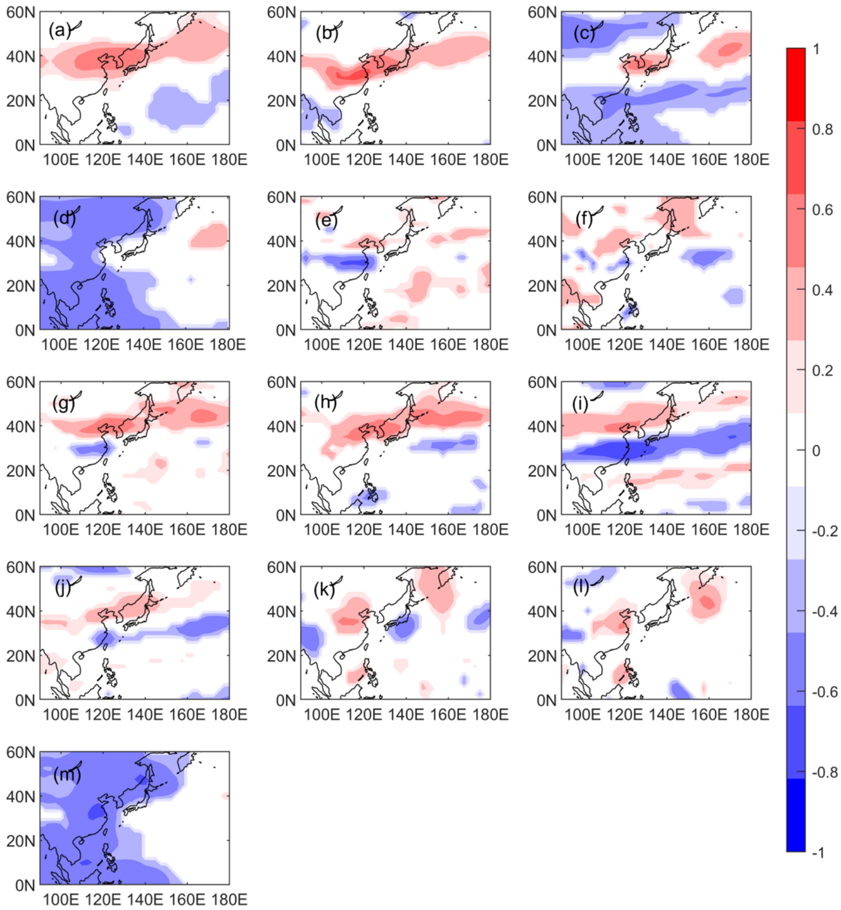

Figure 5: Spatial distribution of the temporal correlation between regional mean time series and influencing factors. (a) involving a 500 hPa air temperature field, (b) involving an 850 hPa air temperature field, (c) involving a 500 hPa geopotential height field, (d) involving a 850 hPa geopotential height field, (e) involving a 500 hPa relative humidity field, (f) involving an 850 hPa relative humidity field, (g) involving a 500 hPa specific humidity field, (h) involving an 850 hPa specific humidity field, (i) involving a 500 hPa zonal wind field, (j) involving an 850 hPa zonal wind field, (K.)) involves a 500 hPa meridional wind field, (l) involves an 850 hPa meridional wind field, and (m) involves a sea level pressure field

Table 2: The performance of different factor combination model from 1961 to 2007.

factor combination	a+s	a+h+u	a+h+s	a+u+s	a+h+u+s
Cor	0.80	0.87	0.86	0.91	0.92
RMSE	0.46	0.41	0.39	0.42	0.44

Table 3: Results of BMA, FPCA, sFPCA, and w-sFPCA. Cor: correlation coefficient; r(%): RMSE error ratio.

method	BMA		FPCA		sFPCA		w-sFPCA	
	Cor	r(%)	Cor	r(%)	Cor	r(%)	Cor	r(%)
train	0.65	5.72	0.48	7.21	0.59	6.18	0.67	5.68
test	0.60	8.35	0.52	8.85	0.57	7.63	0.60	7.92

their weights to obtain the optimal combination of them to estimate the summer temperature in the Yangtze-Huaihe region. For the un-weighted sFPCA, without loss of generality, we combine 500hPa height field, 850hPa height field, 500hPa temperature field and 850hPa temperature field in pairs, and select the combination with the best performance as the prediction factors of un-weighted sFPCA.

We use data from 1961-2007 as the training set and data from 2008-2019 as the test set. Table 3 shows the average fitting results of Bayesian Model Averaging (BMA) [23], FPCA, sFPCA, and w-sFPCA. The correlation coefficient between the prediction and the observed and the RMSE error ratio are listed in the table. The number of functional principal components is 2. The results of BMA, FPCA, and sFPCA are the averages of the optimal factor combination for each site. The fitting result of w-sFPCA is better than that of un-weighted sFPCA and traditional FPCA model. The result of test set, compared with other methods, the correlation coefficient of w-sFPCA maintains a leading advantage, and the RMSE error ratio also has a certain advantages.

5 Conclusions

In this paper, we consider the problem of predicting scalar response variables using multiple types of function predictors. FPC scores estimated by sFPCA are related to the response variable, but only one type of predictors are used. We propose a weighted version that it considers the effects of different types of predictors on the response variables. Numerical simulations and application of summer temperature data in the Yangtze-Huaihe region confirm that the proposed method has better performance than the un-weighted sFPCA method. Although our example demonstrates that the proposed weighted method is better than the Bayesian model average, more research is still needed to compare them. Our experience is that the new weighted method may achieve better interpretability in some applications compared to Bayesian model averaging.

The weighting method that we proposed focuses only on revealing the correlation between predictors and regional climate at two levels. In fact, traditional predictors have 17 or more levels, and in fact the levels themselves can be considered as functional data. In addition, future temperature prediction will involve a variety of levels and models for weighted predictors. Incorporating more layers and more patterns into the model will greatly improve the predictability of the model, but it can also be a great challenge to calculate in this case.

Acknowledgments

This work is supported by the National Social Science Foundation of China (Grant No. 22BTJ035).

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] G. I. Allen, Multi-way functional principal components analysis, In *Proc. 5th IEEE Int. Workshop Comput. Adv. Multi-Sensor Adapt. Process. (CAMSAP)*, 2013, 220–223.

[2] C. Cao, M. He, J. Q. Shi, and X. Liu, Estimation and prediction of a generalized mixed-effects model with t-process for longitudinal correlated binary data, *Comput. Statist.*, 2021, 36: 1461–1479.

[3] H. Shi and J. Cao, Robust functional principal component analysis based on a new regression framework, *J. Agric. Biol. Environ. Stat.*, 2022, 27: 523–543.

[4] S. Cao, C. Cao, Y. Li, and L. Zhu, A statistical downscaling model based on multi-way functional principal component analysis for Southern Australia winter rainfall, *J. Appl. Meteorol. Climatol.*, 2023, 62: 677–689.

[5] A. D. Assamnew and G. M. Tsidu, The performance of regional climate models driven by various general circulation models in reproducing observed rainfall over East Africa, *Theor. Appl. Climatol.*, 2020, 142(3): 1169–1189.

[6] J. O. Ramsay and B. W. Silverman, *Functional Data Analysis*, Springer, New York, 2005.

[7] M. A. Hael, Modeling of rainfall variability using functional principal component method: A case study of Taiz region, Yemen, *Model. Earth Syst. Environ.*, 2021, 7(1): 17–27.

[8] U. Beyaztas, H. Shang, and A. Alin, A function-on-function partial quantile regression, *J. Agric. Biol. Environ. Stat.*, 2022, 27: 149–174.

[9] T. Cai, L. Zhang, and H. Zhou, Adaptive functional linear regression via functional principal component analysis and block thresholding, *Statist. Sinica*, 2018, 28(10): 2455–2468.

[10] M. Febrero-Bande, P. Galeano, and W. González-Manteiga, Functional principal component regression and functional partial least-squares regression: An overview and a comparative study, *Int. Statist. Rev.*, 2017, 85(1): 61–83.

[11] G. Li, H. Shen, and J. Z. Huang, Supervised sparse and functional principal component analysis, *J. Comput. Graph. Stat.*, 2016, 25(3): 859–878.
[12] Y. Nie, L. Wang, B. Liu, et al., Supervised functional principal component analysis, *Statist. Comput.*, 2018, 28: 713–723.
[13] S. Jiang, J. Cao, B. Rosner, and G. A. Colditz, Supervised two-dimensional functional principal component analysis with time-to-event outcomes and mammogram imaging data, *Biometrics*, 2023, 79(2): 1359–1369.
[14] G. Chen, W. Wang, Q. Bao, et al., Evaluation of simulated cloud diurnal variation in CMIP6 climate models, *J. Geophys. Res.: Atmos.*, 2022, 127(6): e2021JD036422.
[15] S. Schubert, Downscaling local extreme temperature changes in South-Eastern Australia from the CSIRO MARK2 GCM, *Int. J. Climatol.*, 1998, 18(13): 1419–1438.
[16] G. Biau, E. Zorita, H. Von-Storch, et al., Estimation of precipitation by kriging in EOF space, *J. Climate*, 1999, 12: 1070–1085.
[17] J. Winkler, J. Palutikof, J. Andresen, et al., The simulation of daily temperature time series from GCM output: Part II: Sensitivity analysis of an empirical transfer function methodology, *J. Climate*, 1997, 10: 2514–2535.
[18] J. Murphy, Predictions of climate change over Europe using statistical and dynamical downscaling techniques, *Int. J. Climatol.*, 2000, 20(5): 489–501.
[19] E. Zorita and H. Storch, The analog method as a simple statistical downscaling technique: Comparison with more complicated methods, *J. Climate*, 1999, 12(8): 2474–2489.
[20] J. Ahn, Y. Kim, K. Shim, et al., Climatic yield potential of japonica-type rice in the Korean Peninsula under RCP scenarios using the ensemble of multi-GCM and multi-RCM chains, *Int. J. Climatol.*, 2021, 41(S1): E1287–1302.
[21] C. Zhu, C. Park, W. Lee, et al., Statistical downscaling for multi-model ensemble prediction of summer monsoon rainfall in the Asia-Pacific region using geopotential height field, *Adv. Atmos. Sci.*, 2008, 25(5): 867–884.
[22] S. Kang, J. Hur, and J. Ahn, Statistical downscaling methods based on APCC multi-model ensemble for seasonal prediction over South Korea, *Int. J. Climatol.*, 2014, 34(14): 3801–3810.
[23] A. E. Raftery, T. Gneiting, F. Balabdaoui, and M. Polakowski, Using Bayesian model averaging to calibrate forecast ensembles, *Mon. Weather Rev.*, 2005, 133: 1155–1174.