Research on the Development of Urban E-commerce Based on Functional Data

Tiansui Wu*

College of Professional Studies, Northeastern University, Vancouver, Canada

*Corresponding author: wu.tians@northeastern.edu

Abstract. With the advancement of the digital economy, the volume of data generated daily is growing exponentially. Traditional statistical analysis methods struggle to handle such large amounts of data. As a significant portion of the collected data is presented in time series, these data demonstrate certain characteristics of functional data as the data volume increases. By expanding these data into functional data using B-spline basis functions, they can be transformed into infinite-dimensional data with function properties. This study utilizes the transaction data of an e-commerce platform throughout a day, applies functional principal component analysis, and analyzes the development of e-commerce in major cities across the country. Future marketing recommendations for these cities are also proposed.

Keywords: functional data, functional principal component analysis, B-spline basis functions.

1. Introduction

With the rapid development of computer technology, the amount of data collected by researchers is growing exponentially. More complex types and forms of data are constantly emerging, and the storage format of data is also changing. One of these formats is functional data, such as order volume in the e-commerce industry, stock transaction prices, and weather conditions in a certain location throughout the year. Although these data may not appear in the form of functions on a scatter plot, over time they begin to exhibit the characteristics and properties of functions. This type of data is dynamically unfolded. If traditional vector-based statistical analysis methods are still used for research, some characteristics of functional data may be overlooked, resulting in the loss of important information that cannot be restored, or even leading to model failure and other issues.

The concept of functional data was first proposed by J.O.Ramsay[1][7], who viewed data as dynamic information based on functions and proposed corresponding research methods. In 1991, he further developed and popularized research methods for functional data in collaboration with Dalzell. Together, they published the book "Functional Data Analysis" in 2005, which summarized various analytical methods for functional data such as functional principal components and functional linear regression, and applied them to the analysis of precipitation and temperature data collected at meteorological stations. Philip T. Reiss, Meng Xu [2][3] utilized B-spline basis functions to represent time-series data. However, this type of data often exhibits autocorrelation. Additionally, several time-series data with periodicity were collected and analyzed using Fourier basis functions for functional principal component analysis. The principal component weight functions and mean perturbation functions show the temporal changes in the data, and the characteristics of the data variations are interpreted in relation to the practical context. The development of the population in major cities in China over the past 20 years has been studied using the functional principal component analysis method by Jingxiang Liu, Junghui Chen, Dan Wang [4][5][6]. They used B-spline basis functions to transform the discrete time series data into continuous functional data. The population situation in major cities in China over the past 20 years was analyzed using functional principal component analysis. Mi Tian and Youxi Luo extended previous research by considering a univariate functional regression model with covariates as functions and response variables as scalars. They proposed an adaptive weighted truncation method to select the number of terms in the expansion of principal components, based on the percentage of variance explained and the percentage of correlated variation explained. Yan conducted research on various cases of functional linear regression and functional

principal component analysis. A method for estimating a varying coefficient function model suitable for irregular data was obtained. This method does not require assumptions about the structure of the covariance function, thus avoiding the risk of incorrect assumptions. J.R.Berrendero, A.Justel, and M.Svarc proposed a new method for multivariate functional principal component analysis in their study. The data can be arranged in a matrix, where the elements of the matrix will become functions. This allows the original data to be simplified into a small number of transformation functions while retaining as much information as possible. Trevor Hastie, Robert Tibshirani, and Jerome Friedman discuss the basis function expansion method for multivariate functional data in their book "The Elements of Statistical Learning: Data Mining". They point out that univariate B-spline basis functions can be extended to the multivariate case, allowing a univariate curve to become a multivariate surface or hyperplane.

2. Multivariate Function Principal Component Analysis

For a column of discrete data $y=(y_1,y_2,\cdots,y_n),$, a mathematical model is established:

$$y_i = X_i(t)+\varepsilon_i, i=1,2,\cdots,n.$$

Where $X_i(t)$ represents the data to be fitted in the form of a time series, and ε represents the disturbance term in the original data. The basis functions refer to a set of functions that constitute a function set, where the functions in the system are mutually independent, and any function in the system can be represented as a linear combination of the basis functions. The method of fitting the curve using a linear combination of known K basis functions in the basis function system is also known as the basis function smoothing method. If we select a suitable set of basis functions: $\varphi_1(t),\varphi_2(t),\cdots,\varphi_K(t)$, there exists a coefficient vector $\xi^T=(\xi_1,\xi_2,\cdots,\xi_K)$ that makes

$$X(t)=\sum_{k=1}^{K}\xi_k\varphi_k(t)$$

Where $\xi_k=\int X(t)\varphi_k(t)dt$ is the projection of $X(t)$ on $\varphi_k(t)$, that allows us to approximate infinite-dimensional data as a finite sum of expansions. The more basis functions we have, the higher the accuracy of the approximation, but it may also lead to overfitting. The information contained in the basis functions is represented by a finite-dimensional vector, achieving dimensionality reduction. The basis function parameters are often estimated using least squares method, i.e.

$$SMSSE(y|\xi)=\min\{\sum_{i=1}^{n}[y_i-\sum_{k=1}^{K}\xi_{ik}\varphi_k]^2\}$$

In this article, B-spline basis functions are chosen.

Functional principal component analysis can effectively reduce the dimensionality of the original data. The dimensionality reduction of functional data is achieved by expanding and truncating the functional basis. Let continuous curve $X_i(t),i=1,2,\cdots,n.$ Let's define the principal components for each curve:

$$\xi_i=\int\varphi(t)X_i(t)\,dt, i=1,2,\cdots,n.$$

Where $\varphi(t)$ is the weight function. The extraction of functional principal components is similar to the classical principal component case, solving the maximization problem under certain constraints:

$$\max\frac{1}{n}\sum_{i=1}^{n}\xi_i^2=\max\frac{1}{n}\sum_{i=1}^{n}[\int X_i(t)\varphi_1(t)dt]^2$$

$$s.t.\left|\varphi(t)\right|^2=\int[\varphi_1(t)]^2\,dt=1$$

If we find the maximum solution that satisfies these constraints, X_{i1} is called the first functional principal component, it is the projection $X_i(t)$ onto $\varphi_i(t)$. $\varphi_i(t)$ is the characteristic function of the first principal component. The kth functional principal component can be obtained using a similar method.

$X_i(t)$ can be represented as a linear combination of a set of basic functions:

$$X(t) = \sum_{k=1}^{K} \xi_k \varphi_k(t)$$

In univariate function principal component analysis, only the order quantity is used as an analysis indicator. However, the number of orders and sales may be influenced by many other factors. Therefore, using only the order quantity is not comprehensive enough. By considering the actual situation, we can observe that the number of orders and sales may be related to average customer price and average discount amount. Therefore, we further collected the average customer price and average discount amount for each region in a day as supplemental data and conducted multivariate function principal component analysis.

Suppose $X \in R^2$, there exist functions $h_{1k}(X_1), h_{2k}(X_2), k = 1,2,\cdots,M_1$, $M_1 \times M_2$ tensor basis is

$$g_{jk}(X) = h_{1j}(X_1)h_{2k}(X_2), j = 1,2,\cdots,M_1, k = 1,2,\cdots,M_2 \quad (1)$$

More generally, we can express $f \in R^2$ as an expansion of a large set of base functions. For example, we can use univariate B-splines as base functions.

To construct the basis, we form the tensor product of all pairwise univariate smoothing spline basis functions in Equation (2.1), usually reducing the number of basis functions for each coordinate accordingly. When observing a N sample with P measurements, where there exists function $x(s) = (x^1(s), x^2(s), \cdots, x^p(s))'$, the principal component of the multivariate function variable is defined as the following linear combination:

$$f = \int \beta(s)' x_i(s) ds \quad (2)$$

Where the weight function for the functional principal components f is $\beta(s) = (\beta_1(s), \beta_2(s), \cdots, \beta_p(s))$ $b(s) = (b_1(s), b_2(s), ..., b_p(s))$, $\beta_i(s)$ $b_i(s)$ representing the relative weight of the i-th function.

In this study, we collected the order volume, average order value, and average order discount amount of major cities and regions nationwide on a large domestic e-commerce platform in a day, and expanded the order volume of major cities and regions nationwide using B-spline basis functions. We can observe the data situation in a three-dimensional pattern by expanding the data based on time and region, as shown in Figure 1,

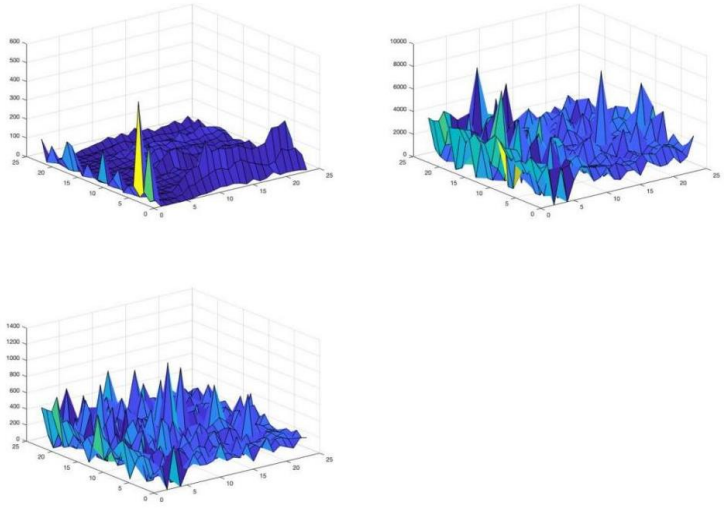


Fig. 1 3D image of raw data

The figure above mainly displays the order volume, average order value, and average order discount price of different cities and regions, and it can be found that the original data gradually presents the characteristics of a function. Therefore, the B-spline basis functions were selected to expand the original data, transforming it into functional data represented in Figure 2,

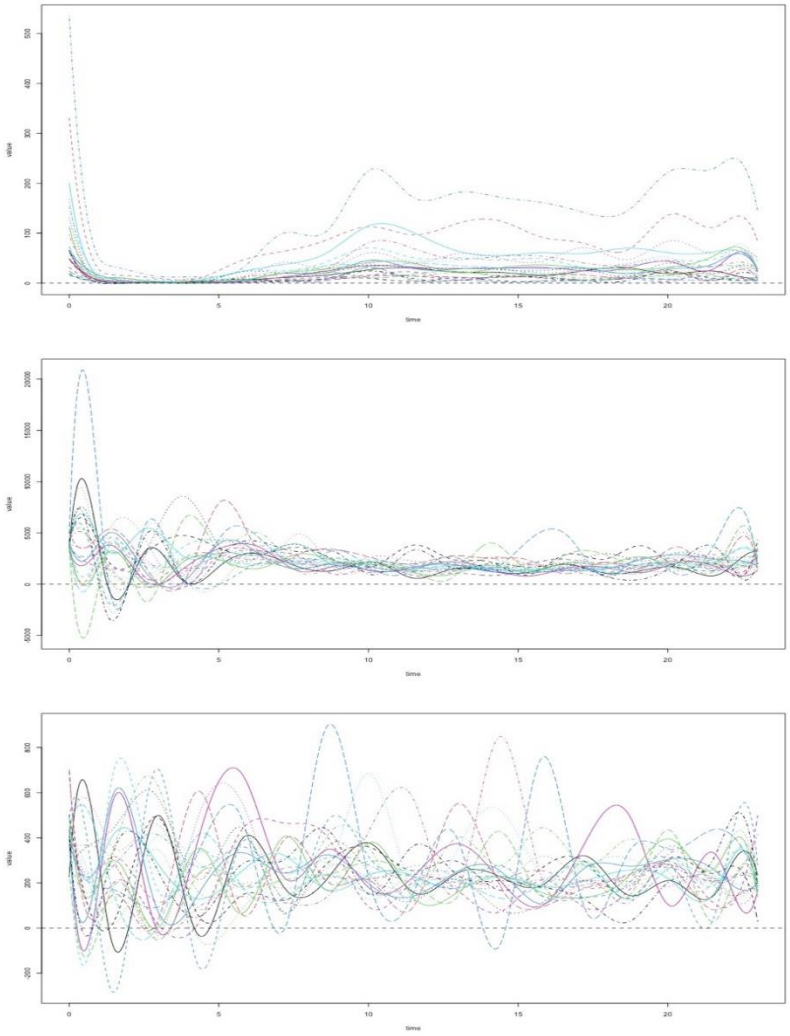


Fig. 2 Image of B-spline transform

Since the data set is three-dimensional, it is difficult to display the three-dimensional data on a plane. Therefore, the analysis of the original data will be conducted separately for each dimension.

By expanding the original data into functional data, the fitting of the order quantity, average customer price, and average order discount amount can be obtained. Due to the large difference in values between the average customer price, average discount amount, and order quantity, the three sets of data were standardized before expanding. The expansion situation at this time is shown in Figure 3,

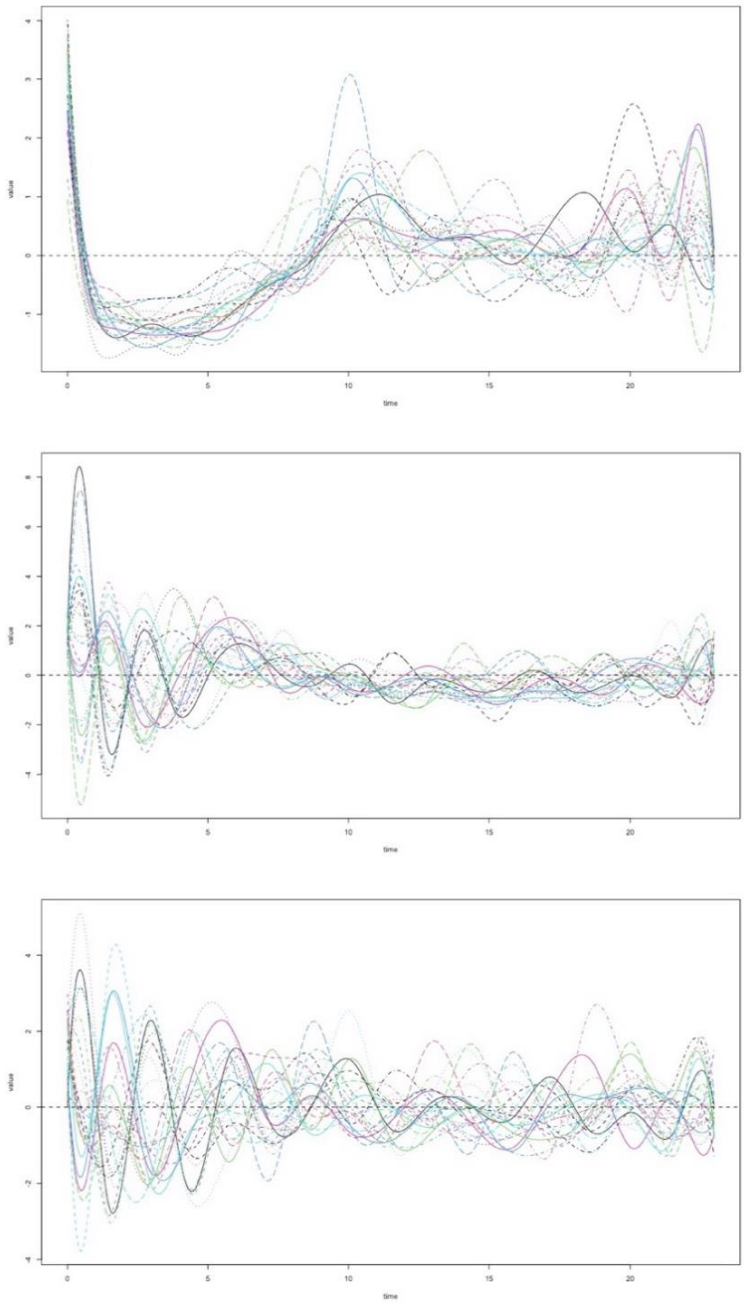


Fig. 3 Standardized B-spline transform

By expanding the original data into functional data, the mean function images of the order quantity, average customer price, and average discount amount can also be obtained, as shown in Figure 4,

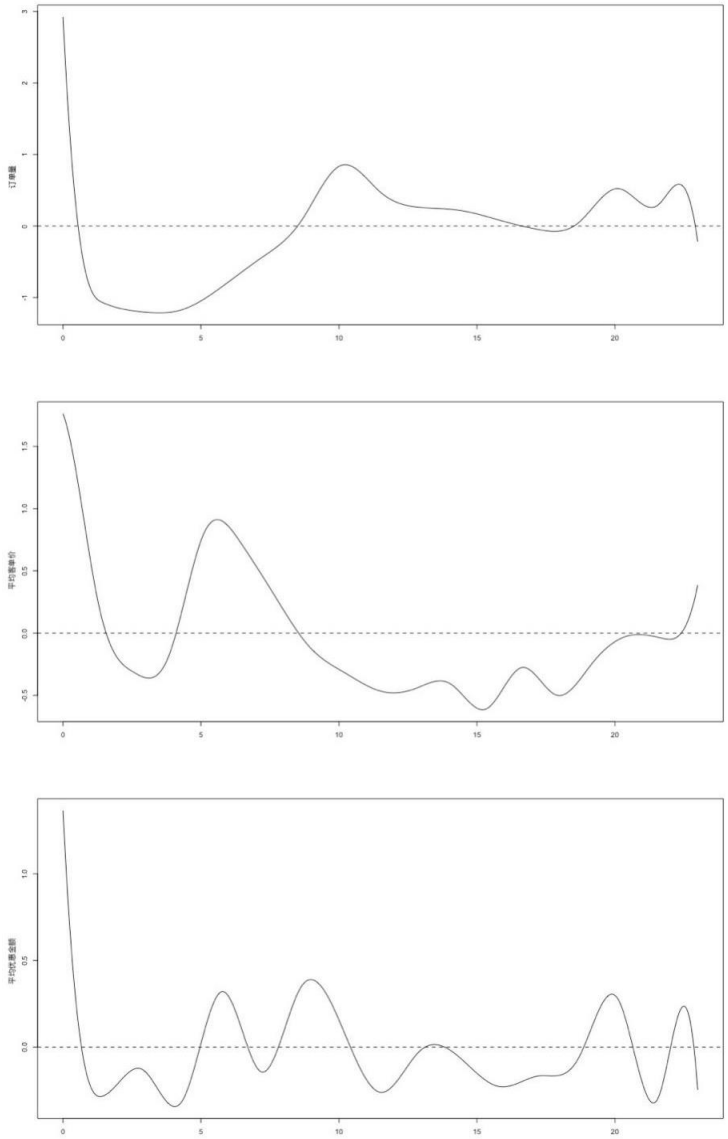


Fig. 4 Mean function image of original data

Since the data has been standardized, for discount amounts below 0, it does not represent an increase in price, but rather a relatively low mean discount. It can be observed that even after standardization, the amount of order remains similar to the non-standardized case. As for the average order value, it is highest at 0:00 and then rapidly declines, likely due to the lower order volume during this time period. However, it increases again from 6:00 to 8:00 and then declines once more. The average discount amount follows a similar pattern, with the highest at 0:00 and then rapidly decreasing. The cumulative variance contribution rates of the three sets of functional data are shown in the following table 1-3.

Table 1. Cumulative Variance Contribution Rates of Amount of Order

	first	second	third	forth
eigenvalue	1.0864	0.8976	0.6895	0.3372
variance contribution rates	25.9%	21.4%	16.4%	8%
cumulative variance contribution	25.9%	47.3%	63.7%	71.7%

Table 2. Cumulative Variance Contribution Rates of The Average Order Value

	first	second	third	forth
eigenvalue	4.6988	2.3969	1.8387	1.6216
variance contribution rates	34.1%	17.4%	13.4%	11.8%
cumulative variance contribution	34.1%	51.5%	64.9%	76.7%

Table 3. Cumulative Variance Contribution Rates of Average Discount Amount

	first	second	third	forth
eigenvalue	4.6898	3.1806	2.2292	1.6576
variance contribution rates	27%	18.3%	12.9%	9.6%
cumulative variance contribution	27%	45.3%	58.2%	67.8%

It can be seen that although the cumulative variance contribution rates for multivariate functional principal component analysis are not as high as those for univariate functional principal component analysis, 70% of the variance still represents the majority of the original information.

Furthermore, the weight coefficients of the first four principal components of order quantity, average order value, and average discount amount can be calculated respectively. they can be substituted into the weight function to generate corresponding weight function image, as shown in Figure 5,

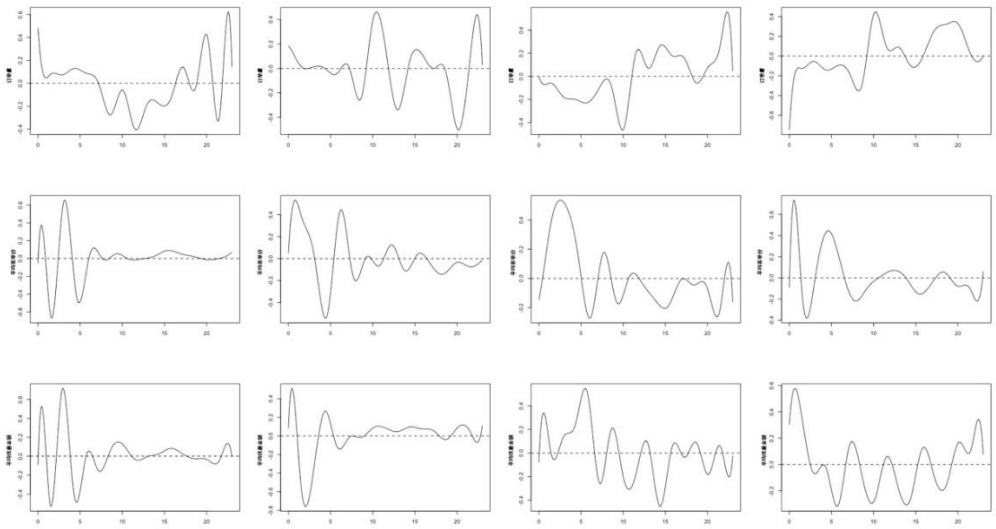


Fig. 5 Weight function image of the forth principal components

Figure 5 shows the weight function image of the top four principal components for three sets of data. Each set of four curves represents the characteristic function curves of each indicator.

For the average customer unit price, the first principal component's weight function has a variance contribution rate of 34.1%. From the figure, it can be seen that the weight function fluctuates greatly from 0:00 to 7:00, but gradually stabilizes after 7:00. The first principal component mainly represents the average customer unit price from 0:00 to 7:00. The second principal component's weight function also shows some fluctuations from 7:00 to 23:00, reflecting the changes in average customer unit price during that time period. The third principal component's weight function exhibits noticeable variations throughout the day, but in most cases, the unit price is lower than the average. The third principal component reflects the pricing situation in some areas where it is below average. The weight function of the fourth principal component is similar to that of the third principal component, but it mostly fluctuates around the average value. The fourth principal component reflects the price situation in areas around the average value.

As for the average discount amount, the weight function in the first principal component fluctuates drastically from 00:00 to 07:00, indicating that people pay more attention to discounts during this time. The weight function in the second principal component is relatively stable compared to the first principal component from 00:00 to 07:00, and after 07:00, it mainly hovers around the mean value with little fluctuation. The weight function in the third principal component fluctuates greatly throughout the day, reflecting the fluctuation of discounts in regions that rely more on discounts. The weight function of the fourth principal component is similar to that of the third principal component, but in most cases, it is below the mean value, reflecting the situation in areas where discounts are not relied upon during daytime shopping.

Also, the corresponding principal component scores for average customer spending and average order discount amount can be obtained,

Table 4. Principal Component Scores of Order Volume

	First	Second	Third	Forth
ANHUI	-0.7717	-0.1055	0.0040	0.8518
BEIJING	0.3116	-0.4353	0.3921	-0.5807
FUJIAN	-0.1943	0.5550	0.1865	0.2219
GUANGDONG	0.3828	-0.0053	0.4989	0.0214
HEBEI	-0.6012	-0.3719	0.3365	-0.1074
HENAN	0.7760	0.1657	1.2238	0.4922
HEILONGJIANG	1.2304	-2.5192	-0.4996	0.8061
HUBEI	0.6618	-0.1942	-0.5247	-0.0416
HUNAN	-0.2992	-0.3941	-1.2998	-1.4079
JILIN	0.4647	1.3228	-2.5844	0.6564
JIANGSU	-0.5567	0.5104	-0.5765	0.3488
JIANGXI	-1.5720	1.0773	0.2783	-0.6145
LIAONING	0.3133	-0.3396	0.1257	0.2484
SHANDONG	-0.9084	0.7929	-0.0287	0.8356
SHANXI	-2.4260	-2.3664	-0.0219	-0.3404
SHANXI	0.3808	1.1669	1.0285	0.0940
SHANGHAI	-0.0027	0.3551	-0.1738	-0.4668
SICHUAN	0.8435	-0.6769	0.4098	0.2867
TIANJIN	1.3077	0.1317	-1.0170	-0.6834
YUNNAN	-2.0769	0.4546	0.1675	0.4755
ZHEJIANG	0.5152	0.1888	0.7732	-0.5684

Table 5. Principal Component Scores of The Average Order Value

	First	Second	Third	Forth
ANHUI	6.0618	1.1417	0.0353	0.5923
BEIJING	-4.9480	0.0879	-0.1104	0.1076
FUJIAN	5.8386	-0.6115	0.6040	-1.3674
GUANGDONG	-0.3119	1.7316	0.8370	0.1272
HEBEI	-4.7394	-0.0124	0.2757	-0.3525
HENAN	-2.2504	1.4463	-1.4387	1.3950
HEILONGJIANG	1.6701	-1.7200	0.7989	2.5950
HUBEI	-2.4790	1.9733	2.4913	-2.3566
HUNAN	0.8123	-0.2781	-1.8157	1.0921
JILIN	4.9270	1.1928	0.1851	0.4682
JIANGSU	2.2916	1.0472	3.1605	0.6320
JIANGXI	-0.5531	0.8004	-1.6771	-1.1830
LIAONING	1.5029	-3.3738	1.6595	2.1288
SHANDONG	0.7423	-0.0252	-2.6977	0.3586
SHANXI	-5.0969	-3.9580	1.0464	-0.8548
SHANXI	-2.0992	2.3886	-0.5695	0.0288
SHANGHAI	-3.2144	0.0243	1.0631	0.8070
SICHUAN	-0.5831	2.5611	-0.0191	0.8220
TIANJIN	3.8021	-0.8756	-1.1100	-1.4684
YUNNAN	4.4472	-1.5618	-0.0342	-2.5967
ZHEJIANG	-0.5351	-2.4764	-2.0310	-0.0071

Table 6. Principal Component Scores for Average Order Discount Amount

	First	Second	Third	Forth
ANHUI	4.2551	1.4053	0.6642	-0.1904
BEIJING	-1.9717	0.7255	-0.8073	1.8295
FUJIAN	2.4133	0.6205	-1.6034	0.3566
GUANGDONG	1.0338	-1.9502	1.7221	-0.6495
HEBEI	-0.8637	3.9603	1.8477	1.7873
HENAN	-3.0327	-0.8869	1.9500	-1.9115
HEILONGJIANG	2.1978	-0.0202	-0.7711	-0.5794
HUBEI	2.2814	-0.6442	0.6435	3.0586
HUNAN	2.0887	-2.6296	0.5558	-1.2008
JILIN	-0.1841	1.7068	-2.7076	-0.2942
JIANGSU	-1.7687	1.0789	-2.1748	-1.6413
JIANGXI	-2.4459	-2.0278	0.0451	-0.0591
LIAONING	-3.0233	-3.1384	-1.3172	0.5752
SHANDONG	-2.3293	1.8248	2.4651	-0.9706
SHANXI	-0.2246	-1.4302	-0.8310	0.5316
SHANXI	4.1239	0.3660	-0.0947	-0.6497
SHANGHAI	1.2891	0.8701	-1.3381	-2.4706

From the table, it can be observed that in terms of order volume, Heilongjiang, Jiangxi, Shanxi, and Yunnan are influenced the most by the first principal component; Heilongjiang and Shanxi are influenced the most by the second principal component; Henan, Jilin, and Tianjin are influenced the most by the third principal component; and Hunan is influenced the most by the fourth principal component. Additionally, some provinces are influenced by multiple principal components. For example, Heilongjiang and Shanxi are significantly influenced by both the first and second principal components.

Regarding the average customer transaction amount, Anhui, Beijing, Fujian, Hebei, Jilin, Shanxi, Shanghai, Tianjin, and Yunnan are influenced the most by the first principal component; Liaoning, Shanxi, Shaanxi, Sichuan, Western regions, and Zhejiang are influenced the most by the second principal component; Henan, Hubei, Jiangsu, Shandong, and Zhejiang are influenced the most by the third principal component; and Heilongjiang, Hubei, Liaoning, and Yunnan are influenced the most by the fourth principal component.

Furthermore, some provinces are simultaneously affected by multiple principal components. For example, Shaanxi is greatly influenced by the first and second principal components simultaneously, while Hebei is greatly influenced by the second and third principal components simultaneously. Moreover, it can be observed that some cities are simultaneously affected by multiple components of multiple indicators. For instance, Fujian is simultaneously influenced by the first component of average order value and the first and third components of average discount amount.

The first principal component of Yunnan is significantly influenced by the order volume, average customer unit price, and average discount amount.

The mean disturbance function of the principal component is shown in Figure 6.

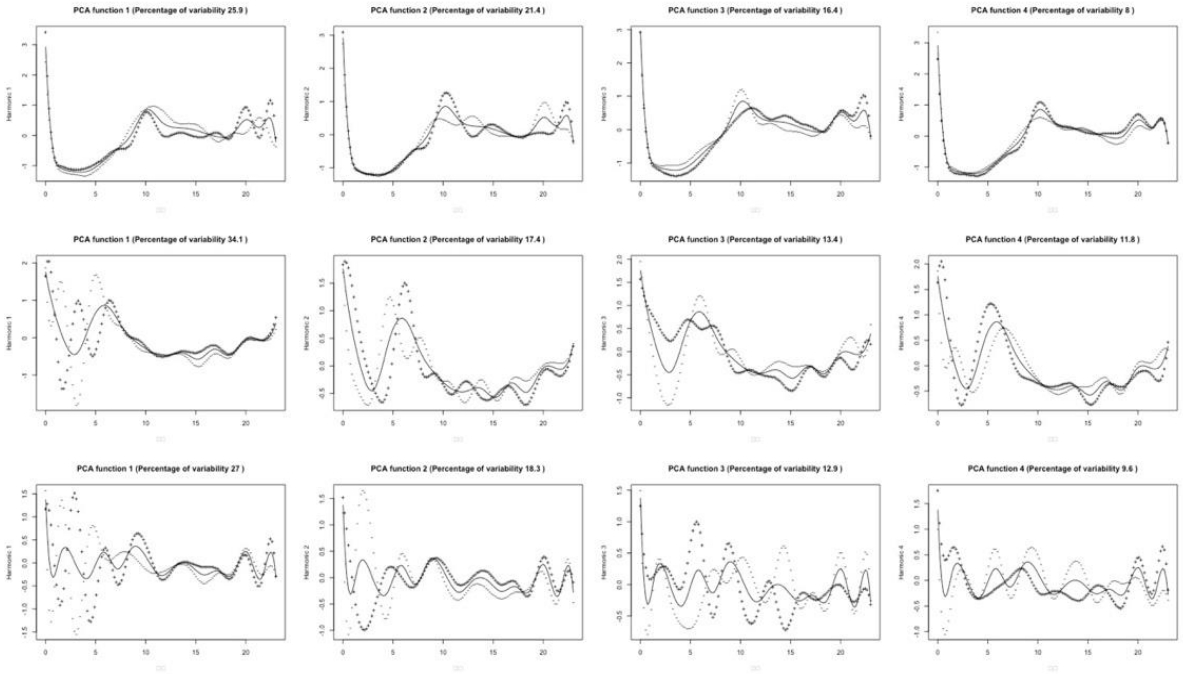


Fig.6 Mean disturbance function of the principal component

The variance contribution rate of the first principal component of the order volume is 25.9%. It can be seen from the figure that there is a significant fluctuation in the order volume from 10:00 to 16:00, indicating that the provinces influenced by the first principal component of the order volume have significant fluctuations in their order volume during this time period. As for the second principal component, its variance contribution rate is 21.4%. There is a significant fluctuation from 10:00 to 11:00, 13:00 to 15:00, and 22:00 to 23:00, which is due to the fact that the e-commerce platform promises that orders placed before 11:00 can be delivered on the same day and orders placed before 23:00 can be delivered the next day. This indicates that the provinces influenced by the second principal component of the order volume have higher requirements for delivery time. As for the third principal component, its variance contribution rate is 16.4%. It mainly reflects the fluctuation of orders from 12:00 to 16:00 and 17:00 to 18:00, which may be similar to the second principal component, but mainly reflects the user's ordering situation during the lunch break. This indicates that the users influenced by the third principal component have requirements for delivery time. As for the fourth principal component, its variance contribution rate is 8%. The main component reflects the fluctuation of order volume from 7:00 to 9:00 and 18:00 to 20:00, combined with the actual findings indicating that this period is mainly the commuting time for going to and from work, suggesting that users in cities with a significant impact on the fourth main component of order volume tend to place orders during commuting hours.

Regarding the average order value, the contribution rate of the first principal component variance is 34.1%. From the graph, it can be observed that there is a significant fluctuation in the average order value from 0:00 to 7:00, indicating that provinces with a significant influence from the first principal component of average order value will pay more attention to price issues when placing orders at night. For the second principal component, the contribution rate of the variance is 17.4%. It can be seen that there is significant fluctuation during 12:00 to 15:00 and 18:00 to 20:00, which is because these time periods are usually lunch break hours for most companies. This suggests that users from provinces with a significant influence from the second principal component of average order value will consider price issues more when placing orders during lunch breaks. For the third principal component of average order value, the contribution rate of the variance is 13.4%. There is significant fluctuation from 15:00 to 19:00 and 20:00 to 23:00, indicating that provinces with a significant influence from the third principal component of average order value will emphasize price issues when placing orders during these time periods. For the fourth principal component of average order value, its contribution

rate to variance is 11.8%. There is significant price fluctuation between 7:00-10:00 and 18:00-20:00, which is mainly due to the commuting hours. This suggests that customers who place orders during this time period pay more attention to the price.

As for the first principal component of average discount amount, its contribution rate to variance is 27%. From the graph, we can observe significant fluctuation in discount amount between 0:00-7:00, indicating that these customers are more concerned about discounts when placing orders at night. The second principal component of average discount amount has a variance contribution rate of 18.3%, and shows significant fluctuation between 10:00-18:00. This implies that provinces heavily influenced by the second principal component of average discount amount tend to prioritize discounts when placing orders during the day. For the third principal component of average order value, its contribution rate to variance is 12.9%. In this component, there is significant fluctuation throughout the day, but not as strong as the first principal component. It can be inferred that provinces influenced by the third principal component of average discount amount do not show a clear time preference, but are more sensitive to the extent of discounts. For the fourth principal component of the average discount amount, the variance contribution rate is 9.6%. There is a significant fluctuation between 5:00-11:00 and 16:00-19:00. It is observed that these time periods mainly correspond to the commuting hours for going to and returning from work, indicating that users who place orders during this time period prioritize discounts.

In terms of order volume, Heilongjiang, Jiangxi, Shanxi, and Yunnan are greatly influenced by the first principal component. This suggests that these provinces tend to place orders between 10:00-16:00. Heilongjiang and Shanxi are greatly influenced by the second principal component, indicating that these provinces have higher expectations for delivery time. Hunan is greatly influenced by the fourth principal component, indicating that this province tends to place orders during commuting hours. For provinces that are influenced by multiple principal components, such as Heilongjiang and Shanxi being significantly influenced by the first and second principal components, it would be recommended to focus on promoting products between 10:00-16:00 and 22:00 for these users.

In terms of average customer spend, Anhui, Beijing, Fujian, Hebei, Jilin, Shanxi, Shanghai, Tianjin, and Yunnan are greatly influenced by the first principal component. It can be observed that these provinces consider price more when placing orders during the night. For provinces like Beijing, Guangdong, and Jiangsu, which are greatly influenced by the first principal component of order volume, it is advisable to offer products in their price range during the night. On the other hand, provinces like Liaoning, Shanxi, Shaanxi, Sichuan, western provinces, and Zhejiang are greatly influenced by the second principal component, showing significant fluctuations from 12:00 to 15:00. Based on analysis, it can be inferred that this time period corresponds to the lunch break, indicating that users from these provinces are more price-conscious when placing orders during the lunch break. Considering the provinces that are greatly influenced by the third principal component in order volume, such as Henan, Jilin, and Tianjin, it is recommended to promote products with suitable prices during the lunch break. Similarly, provinces like Henan, Hubei, Jiangsu, Shandong, and Zhejiang, which are greatly influenced by the third principal component in average customer spending, should also have appropriately priced products promoted during the corresponding time period. Heilongjiang, Hubei, Liaoning, and Yunnan are greatly influenced by the fourth principal component. For users in these provinces, it is recommended to focus on pushing products with suitable prices during commuting hours. At the same time, considering the fourth principal component of order quantity, users in heavily influenced provinces such as Hunan are more inclined to place orders during commuting hours and pay more attention to pricing issues when placing orders.

As for the average discount amount, Anhui, Henan, Shanghai, and Yunnan are greatly influenced by the first principal component. For users in these provinces, when recommending products at night, it is important to focus on recommending products with larger discounts, taking into account the first principal component of order quantity. For provinces heavily influenced by these two factors, such as Yunnan, it is advisable to recommend products with larger discounts at night. Hebei and Shanghai are greatly influenced by the second principal component, so it is recommended to focus on

recommending products with larger discounts to users in these provinces during the day. Hebei, Henan, Jilin, Liaoning, Shandong, Shanxi, and Sichuan are greatly influenced by the third principal component. For users in these provinces, it is recommended to remember to push products with larger discounts.

Hubei and Zhejiang are heavily influenced by the fourth principal component. For users in these provinces, it is important to prioritize the promotion of products with high discounts during short commute times. Additionally, taking into account the fourth principal component of order quantity and average order value, users in heavily impacted provinces such as Hunan tend to be inclined towards purchasing affordable products with significant discounts during commute hours.

This study primarily conducts a multidimensional principal component analysis on the daily order quantity, average order value, and average discount amount data of various regions. The analysis reveals that users in Shanxi and Yunnan prefer placing orders at night and consider price issues more significantly during their purchase. Yunnan users also value discounts, so when making recommendations to Yunnan users, the discount level should be taken into consideration. Users in Henan, Hubei, Jiangsu, Shandong, and Zhejiang are more likely to place orders during the lunch break period, with users in Henan paying more attention to price and discount levels.

Users in Shandong are more inclined to place orders during the nap time period and are more concerned about the intensity of discounts; users in Hunan, on the other hand, are more inclined to place orders during the commuting time period and are more concerned about discounts. Although users in Heilongjiang, Hubei, Liaoning, and Yunnan are not more inclined to place orders during the commuting time period, they are more concerned about the price when placing orders during this time period. Users in Hubei and Zhejiang are more concerned about the intensity of discounts

References

[1] J.O Ramsay, Giles Hooker, Spencer Graves, Functional Data Analysis with R and MATLAB, 2021.

[2] F.A. Ocaña, A.M. Aguilera, M.J. Valderrama, Functional Principal Components Analysis by Choice of Norm, Journal of Multivariate Analysis, 1999,17(2): 262-276.

[3] Philip T. Reiss, Meng Xu, Tensor product splines and functional principal components, Journal of Statistical Planning and Inference, 2020,16(1): 1-12.

[4] Jingxiang Liu, Junghui Chen, Dan Wang, Wavelet functional principal component analysis for batch process monitoring [J], Chemometrics and Intelligent Laboratory Systems, 2020,196(3): 103-124.

[5] Jeng-Min Chiou, Ya-Fang Yang, Yu-Ting Chen, Multivariate functional linear regression and prediction, Journal of Multivariate Analysis, 2021,146(3): 13-27.

[6] J.R. Berrendero, A. Justel, M. Svarc, Principal components for multivariate functional data, Computational Statistics & Data Analysis, 2011,55(9): 2619-2634.

[7] Trevor Hastie, Robert Tibshirani, Jerome Friedman, The Elements of Statistical Learning: Data Mining, and Prediction 2nd ed, 2021, Page 139-190.

[8] Jingxiang Liu, Dan Wang, Junghui Chen, Global-local based wavelet functional principal component analysis for fault detection and diagnosis in batch processes, Chemometrics and Intelligent Laboratory Systems, 2021,12: 261-279.