RESEARCH ARTICLE | FEBRUARY 28 2019

A new forecasting method for cereal consumption: Combined method of LASSO and semi-parametric regression FREE

Yunqian Ma; Haiquan Xu; Yanzhi Guo; Junmao Sun; Xiuli Wang

AIP Conf. Proc. 2079, 020041 (2019)
https://doi.org/10.1063/1.5092419

A New Forecasting Method for Cereal Consumption: Combined Method of LASSO and Semi-parametric Regression

Yunqian Ma[1], Haiquan Xu[1], Yanzhi Guo[1], Junmao Sun[1], Xiuli Wang[1, a)]

[1]*Institute of Food and Nutrition Development, Ministry of Agriculture and Rural Affairs, Beijing 100081, China*

[a)]Corresponding author: wangxiuli01@caas.cn

Abstract: There is a diversified range of food available while cereal is the main food of Chinese residents. The most important food pattern of Chinese residents is that cereal is the staple food. While many people have misunderstandings about the nutrition of cereal food, the consumption of cereal is decreasing year by year. In order to ensure the dietary balance of Chinese residents, the influence factors of cereal consumption are valuable to research. Because of the high number of impact factors, we first use the LASSO method to select the main influence factors of cereal consumption, and then we construct a partially linear semi-parametric model for predicting the cereal consumption of Chinese residents. The results show that the factors affecting per capita consumption of rice, wheat and maize are different from one another and the three cereals have both common impact factors and differentiated ones; Per Capita Disposable Income (PCDI) is the common factor with a linear positive relationship to the consumption of the three cereals; the model constructed in this paper is well fitted and can accurately forecast the consumption of cereals; the average per capita consumption of rice, wheat and maize is predicted to be 78.56 kg/year, 62.73kg/year and 6.64 kg/year respectively by 2025, which is excessive and is caused by irrational dietary structure, food wastage and processing losses.

INTRODUCTION

"Cereal is the basis of a balanced diet. Enough cereals should be taken over three meals a day." According to the *Dietary Guidelines for Chinese Residents 2016*, the daily recommended intake of cereal and potato is 250~400g, in which whole grain and mixed beans is 50~150g and potato is 50~100g [1]. Cereal is the basis of a balanced diet and balanced nutrition, yet with the development of economy China has undergone a remarkable change in food structure and nutrition intake and the problem of dietary inequality is still serious. According to the monitoring data of nutrition and health of Chinese residents from 1982 to 2012, Chinese residents' cereal intake is decreasing year by year, and more than 80 percent of Chinese adults are deficient in whole grain [2]. In order to ensure the balanced nutrition of Chinese residents, research on cereal consumption and its influencing factors is increasingly important. It is necessary to note that the scope of this study is the definition of cereal, which is rice, wheat and maize, the definition of whole grain and coarse grain can be seen He et al.[3].

There is a great deal of research on food consumption structure in current mainstream literature, for example, Nzuma and Sarker used an Error Corrected Almost Ideal Demand System to study the consumption of four major cereals in Kenya[4]; Zheng and Henneberry studied food grain consumption in the urban Jiangsu province of China based on the QUAIDS and the AIDS models, and concluded that certain demographic variables have a significant impact on food grain dem[5]; and Jiang researched the impact of urban economic development on changes in food demand and associated land requirements for food production based on the Almost Ideal Demand System [6]; Hovhannisyan and Gould examined food demand structure and its dynamics for 11 commodities in urban China [7]; and Valin et al. studied the 2050 world food demand based on computable general equilibrium, partially equilibrium and eight other global economic models [8].

2018 International Conference on Biotechnology and Bioengineering (8th ICBB)
AIP Conf. Proc. 2079, 020041-1–020041-10; https://doi.org/10.1063/1.5092419
Published by AIP Publishing. 978-0-7354-1807-3/$30.00

Nevertheless, the study of grain classification (rice, wheat, maize) consumption and its influencing factors is obviously insufficient. The current methods (such as LES/ELES, AIDS, and PE) are mainly linear in essence, but there is a certain nonlinear relationship between dependent variables and independent variables. Adopting linear models entirely is too strict; the relationship between them can't be captured accurately [9]. On the other hand, it is difficult to provide economic interpretation by using an entirely nonlinear model; therefore, a partially linear semi-parametric model will be used to capture the relationship between cereal consumption and its impact factors, and then to predict future cereal consumption. Cereal consumption is affected by numerous factors, in which the dominant factors for the variation of cereal consumption are worth further study. Too many factors will affect the estimation and prediction accuracy of the model and also cause the problem of multicollinearity. As a result, it is necessary to screen variables before building the model. The general methods of variable selection are principal component analysis (PCA) and grey correlation degree. The former is the linear transformation of the original variables, which cannot carry out the screening of variables and is also not conducive to the explanation of variables. At the same time, the grey correlation degree method possesses definite subjectivity in variable selection. In short, the above two methods cannot achieve an effective variable filter. But LASSO is an effective variable selection method which can eliminate the multicollinearity problem and makes some coefficients to be zero in order to reduce indicators by constructing a penalty function. Therefore, we will combine the partially linear semi-parametric model and the LASSO method to analyze the main factors of cereal consumption and then predict the future cereal consumption of Chinese residents.

METHOD

LASSO method

The Least Absolute Shrinkage and Selection Operator (LASSO) is proposed by Tibshirani [10]. The idea of this method is making some coefficients to be zero to achieve the purpose of reducing indicators by constructing a penalty function.

Suppose the linear regression model is $Y = X\beta + \varepsilon$, Y is the vector of dependent variable and X is the matrix of independent variable. $\beta = (\beta_1, \beta_2, \cdots, \beta_p)$, ε is coefficient vector and error vector, respectively. x_{ij} is the standardized independent variable, y_i is the centralized dependent variable, which $i = 1,2,\cdots,n$, $j = 1,2,\cdots,p$. The traditional ordinary least squares method is:

$$\min_{\beta} \sum_{i=1}^{n} (y_i - \sum_{j=1}^{p} x_{ij}\beta_j)^2 \tag{1}$$

While the LASSO is a L_1 regularization with a L_1 norm punishment, that is

$$\min_{\beta} \sum_{i=1}^{n} (y_i - \sum_{j=1}^{p} x_{ij}\beta_j)^2 + \lambda \sum_{j=1}^{p} \left|\beta_j\right| \tag{2}$$

L_1 norm punishment is equal to

$$\min_{\beta} \sum_{i=1}^{n} (y_i - \sum_{j=1}^{p} x_{ij}\beta_j)^2 \tag{3}$$

$$s.t. \sum_{j=1}^{p} \left|\beta_j\right| \le s. \tag{4}$$

The idea of the LASSO method is that the sum of absolute coefficients cannot be too large. Under this premise, applying the ordinary least squares method, the sum of squares of residuals is the least. L_1 norm punishment (3) is no analytic solution, but can be solved by convex quadratic programming. The Least Angle Regression proposed by Efron can effectively solve this problem[11]. The most obvious advantage of L_1 norm punishment is achieving shrinkage while implementing the selection of variables, which is making some coefficients β_j to be zero during the process of optimum solution.

Semi-parametric regression model

Suppose a simple semi-parametric regression model as follows:

$$y_i = \beta_0 + \beta_1 x_1 + f(x_2) + \varepsilon_i \tag{5}$$

where y_i is the dependent variable, x_1, x_2 are the independent variables, and ε_i is the error term. The model has a nonparametric component $f(x_2)$ and a parametric component $\beta_0 + \beta_1 x_1$.

We can fit (5) using a penalized linear spline through the mixed model:

$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \sum_{k=1}^{K} u_k (x_2 - k_K)_+ + \varepsilon_i \tag{6}$$

where $u_k \overset{ind.}{\sim} N(0, \sigma_u^2)$ and $\varepsilon_i \overset{ind.}{\sim} N(0, \sigma_\varepsilon^2)$.

Note that this is a special case of the Gaussian linear mixed model:

$$Y = X\beta + ZU + \varepsilon \tag{7}$$

with Y containing the y_i values.

$$X = \begin{bmatrix} 1 & x_{11} & x_{21} \\ \vdots & \vdots & \vdots \\ 1 & x_{1n} & x_{2n} \end{bmatrix},\ \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix},\ Z = \begin{bmatrix} (x_{21} - k)_+ & \cdots & (x_{21} - k_K)_+ \\ \vdots & \ddots & \vdots \\ (x_{2n} - k)_+ & \cdots & (x_{2n} - k_K)_+ \end{bmatrix}, U = \begin{bmatrix} u_1 \\ \vdots \\ u_K \end{bmatrix}.$$

where $Cov(U) = \sigma_u^2 I$ and $Cov(\varepsilon) = \sigma_\varepsilon^2 I$.

More model forms and details about semi-parametric regression models can be seen in Ruppert et al. [12].

EMPIRICAL ANALYSIS

Data Description

The sample data in this paper ranges from 1991 to 2013. The primary data of cereal consumption comes from the food balance sheet of the Food and Agriculture Organization (FAO). Cereal producer price downloads are from the National Agricultural Product Cost Benefit Data Compilation. Per Capita Disposable Income (PCDI), Per Capita GDP (GDP), Consumer Price Inflation (CPI), Food Expense (FE), Engel Coefficient (EC), Aging Rate (AR), and Urbanization Rate (UR) are from the *Annals of Statistics* by the National Bureau of Statistics of China (NBSC). The Aging Rate (AR) is equal to the number of people over 60 divided by the total population at the year-end, and Urbanization Rate (UR) is the urban population divided by the total population at the year-end. In order to represent the concept of elasticity, we take the natural logarithm processing of sample data.

The trend plot of per capita cereal consumption from 1991 to 2013 is shown in Fig. 1 and its descriptive statistics are listed in Table 1. The results show that the mean of per capita consumption of rice, wheat and maize is 76.48

kg/year, 70.82kg/year and 6.58 kg/year respectively, and the total is 153.88kg/year. Between 1991 and 1997, the consumption of wheat is higher than rice consumption, while since 1998 the per capita rice consumption has exceeded the consumption of wheat. There has been a trend of increasing volatility, but the overall trend is relatively stable. Wheat consumption declines from 1997 to 2009 and smoothly overall after 2010. The reasons for the change in cereal consumption trends and the factors impacting cereal consumption are the research contents of this paper.

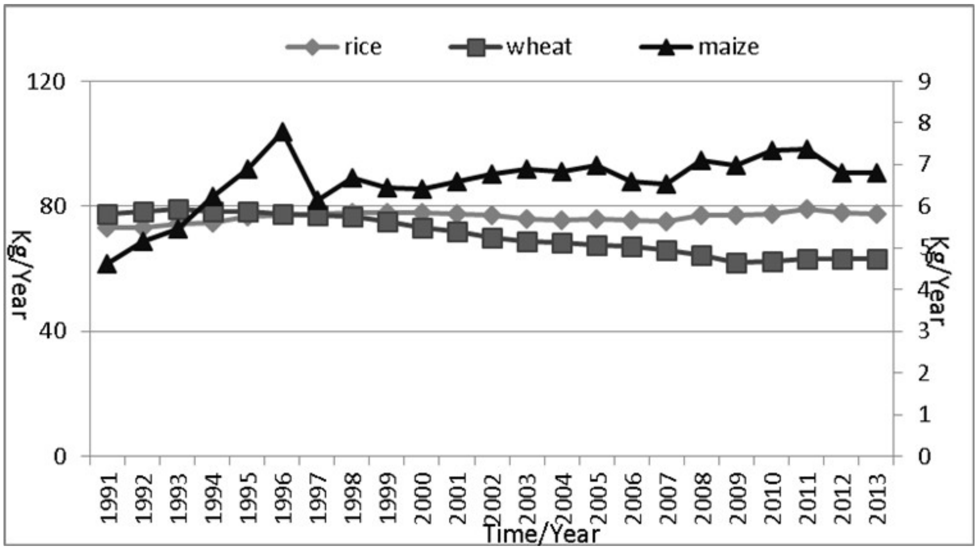

FIGURE 1. Trend plot of Chinese residents' per capita cereal consumption from 1991 to 2013.

TABLE 1. Descriptive statistics of per capita cereal consumption

	Mean	**Median**	**Maximum**	**Minimum**	**Std. Dev.**	**Skewness**	**Kurtosis**
Rice	76.4822	76.9900	79.2400	73.0600	1.5786	-0.6897	2.8247
Wheat	70.8152	69.9900	79.1300	62.0900	6.2287	-0.0352	1.4351
Maize	6.5839	6.7600	7.8000	4.6300	0.7112	-1.1484	4.3642

The changing trend of grain consumption is affected by many factors, such as the level of residents' income, economic development status, demographic structure, and cereal price. Based on the research of Zhou [16], Chen and Fu [2] and Abdelmonem and Adnen [1], in this paper, we choose the following eight influence factors as independent variables: Cereal (Rice, Wheat, and Maize) Producer Price (RPP, WPP, MPP), Per Capita Disposable Income (PCDI), Per Capita GDP (GDP), Consumer Price Inflation (CPI), Food Expense (FE), Engel Coefficient (EC), Aging Rate (AR), and Urbanization Rate (UR). Per Capita Cereal (Rice, Wheat, and Maize) Consumption is the dependent variable.

Variable Selection

The LASSO method is an estimation method with which to implement effective variable selection and eliminate multiple co-linear problems. For effectively screening out the causes of cereal consumption, in this paper, the LASSO regression method is adopted to select the variables of rice, wheat and maize consumptions respectively accomplished by LARS package in R software.

The LASSO variable selection process is introduced in detail with maize sample data. The results show that all the solutions of LASSO can be obtained by six steps, adjusted $R^2 = 0.833$, which means the model is well fitted. According to the principle of Least Angle Regression, the fitted equation is selected with least Cp statistics. As shown in Table 2 and Fig. 2, the least Cp statistic is 3.7882 in the sixth step which corresponds to the optimal model. There are four variables in the sixth step, the coefficients of the following four variables: Consumer Price Inflation (CPI), Food Expense (FE), Engel Coefficient (EC), and Urbanization Rate (UR), are zero, which indicates they are screened out. The remaining variables are Maize Producer Price (MPP), Per Capita Disposable Income (PCDI), Per Capita GDP (GDP), and Aging Rate (AR).

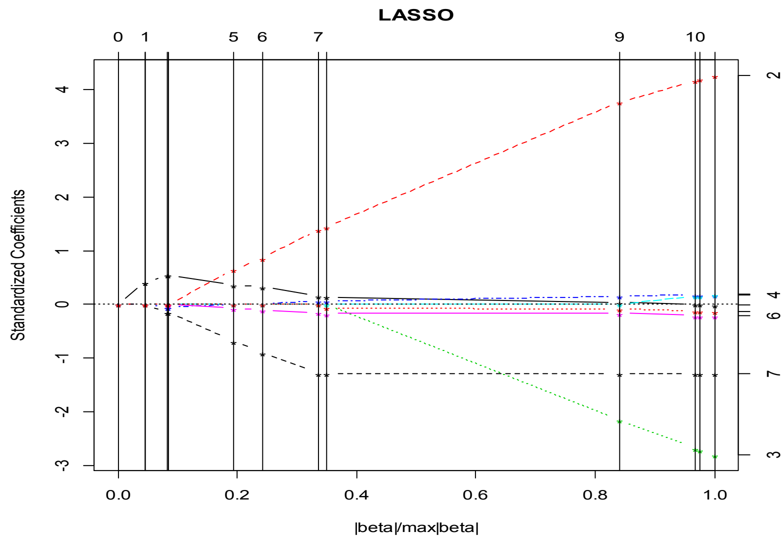

FIGURE 2. LASSO regression plot

TABLE 2. Residual sum of squares and Cp statistics

	Df	Rss	Cp
0	1	0.3015	63.0244
1	2	0.1092	11.4209
2	3	0.1090	13.3734
3	4	0.0927	10.8294
4	5	0.0923	12.7351
5	6	0.0667	7.5800
6	5	0.0602	3.7882
7	6	0.0548	4.2821
8	7	0.0546	6.2234
9	8	0.0505	7.0644
10	9	0.0502	9.0025
11	8	0.0502	7.0014
12	9	0.0502	9.0000

TABLE 3. Parameter estimates of LASSO regression

	X1	X2	X3	X4	X5	X6	X7	X8
1	0	0	0	0	0	0	0	0
2	0.1991	0	0	0	0	0	0	0
3	0.1997	0	0	-0.0051	0	0	0	0
4	0.2660	0	0	-0.2207	0	0	-0.2205	0
5	0.2666	0	0	-0.2197	0	-0.0057	-0.2279	0
6	0.1760	0.1639	0	0	0	-0.0976	-0.9931	0
7	0.1598	0.2168	0	0	0	-0.1519	-1.3115	0
8	0.0745	0.3510	0	0.2398	0	-0.2064	-1.8779	0
9	0.0669	0.3650	0	0.2506	0	-0.2385	-1.8768	-0.0596
10	0.0174	0.9498	-0.5084	0.5717	0	-0.2170	-1.8834	-0.0755
11	0	1.0518	-0.6314	0.6471	0.0514	-0.2818	-1.8774	-0.1169
12	0	1.0586	-0.6404	0.6513	0.0541	-0.2849	-1.8767	-0.1157
13	-0.0043	1.0770	-0.6618	0.6658	0.0642	-0.2978	-1.8759	-0.1271

Using the same method, four influence factors of rice consumption are selected: Per Capita Disposable Income (PCDI), Consumer Price Inflation (CPI), Aging Rate (AR), and Urbanization Rate (UR). Meanwhile, the selected wheat consumption affecting factors are the following six variables: Wheat Producer Price (WPP), Per Capita Disposable Income (PCDI), Consumer Price Inflation (CPI), Food Expense (FE), Aging Rate (AR), and Urbanization Rate (UR).

Empirical Analysis based on Linear Semi-parametric Model

In this section, on the basis of variable selection, we will construct a partially linear semi-parametric model to study the relationship between cereal consumption and its influence factors, and then predict the future cereal consumption of Chinese residents. We take rice consumption data as an example to introduce the process of modeling with the partially linear semi-parametric method.

As mentioned above, the selected influence factors of wheat consumption are Wheat Producer Price (WPP), Per Capita Disposable Income (PCDI), Consumer Price Inflation (CPI), Food Expense (FE), Aging Rate (AR), and Urbanization Rate (UR). In order to build the partially linear semi-parametric model, the first step is to ascertain the linear and nonlinear parts, which can be decided by the Pearson correlation coefficient. The absolute Pearson correlation coefficient of wheat consumption and its influence factors – WPP, PCDI, CPI, FE, AR, and UR – is 0.7435, 0.944, 0.4204, 0.9263, 0.9691, and 0.9789, respectively. As we can see, the lowest two correlation coefficients are Wheat Producer Price (WPP) and Consumer Price Inflation (CPI), that is to say, they have a low linear correlation with wheat consumption, which makes them nonlinear parts. From the above, the partially linear semi-parametric model of wheat consumption is initially identified as:

$$wheat = \beta_0 + \beta_1 pcdi + \beta_2 fe + \beta_3 ur + \beta_4 ar + f(wpp) + f(cpi) \tag{8}$$

The estimation results show that the coefficient of Aging Rate (AR) is eliminated as it is below the 0.05 significance level. The final model is as follows:

$$wheat = \beta_0 + \beta_1 pcdi + \beta_2 fe + \beta_3 ur + f(wpp) + f(cpi) \tag{9}$$

The model estimation and results of wheat sample are shown in Table 4. The parameters of partially linear semi-parametric model in this paper are estimated by R package of SemiPar, in which the penalty spline estimation is used for non-parametric estimation. As shown in Table 4, the partially linear semi-parametric model of wheat sample is modeled as follow:

$$wheat = 2.585 + 0.3354 pcdi - 0.2322 fe - 0.9467 ur + f(wpp) + f(cpi) \tag{10}$$

TABLE 4. Model estimation and results of wheat sample data

Summary for linear components				
	Coef	**Se**	**ratio**	**p-value**
intercept	2.5850	0.53200	4.859	0.0000
Pcdi	0.3354	0.08244	4.069	0.0001
Fe	-0.2322	0.09172	-2.531	0.0136
Ur	-0.9467	0.12590	-7.518	0.0000
Summary for nonlinear components				
	Df	**Spar**	**knots**	
f(wpp)	1.314	3.364	4	
f(cpi)	1.000	252.500	4	

The parameters of partially linear semi-parametric model are all at the significance level of 0.05, which indicates that the model is well fitted. The key factors of per capita wheat consumption are Per Capita Disposable Income (PCDI), Food Expense (FE), Urbanization Rate (UR), Wheat Producer Price (WPP), and Consumer Price Inflation (CPI). The last two factors have a lower linear correlation with the dependent variable, and are fitted as nonlinear parts. Per Capita Disposable Income (PCDI) is positively correlated with per capita wheat consumption, while Food Expense (FE) and Urbanization Rate (UR) have a negative correlation with the dependent variable. As the first factor rises 1%, the wheat consumption accordingly increases 0.3354. Nonetheless, when Food Expense (FE) and Urbanization Rate (UR) increase 1%, the wheat consumption decreases 0.2322% and 0.9467%, respectively.

TABLE 5. Model estimation and results of rice sample data

	Summary for linear components			
	Coef	**Se**	**Ratio**	**p-value**
intercept	4.1010	0.22810	17.980	0.0000
Pcdi	0.0927	0.01538	6.029	0.0001
Ur	-0.3068	0.05888	-5.210	0.0000
	Summary for non-linear components			
	Df	**Spar**	**Knots**	
f(cpi)	1.000	753.1	4	

The process of modeling rice and maize sample data is similar to that of the wheat sample, which is well-documented above; therefore, we will give no more detail here. As illustrated in Table 5 and Table 6, the parameters are all at the significance level of 0.05. The major impact factors of rice consumption are Per Capita Disposable Income (PCDI), Urbanization Rate (UR), and Consumer Price Inflation (CPI). The former two factors are fitted as linear parts and the last is nonlinear. The partially linear semi-parametric model of the rice sample is modeled as follows:

$$rice = 4.101 - 0.3068ur + 0.0927\,pcdi + f(cpi) \tag{11}$$

Per Capita Disposable Income (PCDI) is positively correlated with rice consumption, while Urbanization Rate (UR) has a negative correlation with the dependent variable. When Per Capita Disposable Income (PCDI) grows 1%, the rice consumption will grow 0.0927%. Whereas, when the Urbanization Rate (UR) increases 1%, the rice consumption decreases 0.3068.

TABLE 6. Model estimation and results of maize sample data

	Summary for linear components			
	Coef	**Se**	**ratio**	**p-value**
intercept	-7.0420	1.50400	-4.683	0.0001
Pcdi	0.4371	0.06433	6.795	0.0000
	Summary for non-linear components			
	Df	**Spar**	**knots**	
f(ar)	1.000	367.7	4	

As Table 6 shows, the main influence factors of per capita maize consumption are Per Capita Disposable Income (PCDI) and Aging Rate (AR). At the same time, the former one is modeled as a linear part of the overall model, which has a positive linear correlation with the dependent variable. And the partially linear semi-parametric model of the maize sample is fitted as follows:

$$mazie = -7.042 + 0.4371\,pcdi + f(ar) \tag{12}$$

When the Per Capita Disposable Income (PCDI) rises 1%, per capita maize consumption increases 0.0927%.

Three biomarkers, mean absolute error (MAE), mean absolute percentage error (MAPE) and mean squared error (MSE), are used for evaluating the accuracy of the models in our paper. They are calculated separately as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|e_i|,\quad MSE = \frac{1}{n}\sum_{i=1}^{n}e_i^2,\quad MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100 \tag{13}$$

In which, n is the number of observations, e_i is the standard error of the fitting model, $\hat{y}_i$ is the fitted value and y_i is the observed value.

The calculated values of the three indexes are listed in Table 7. As the results show, the MAE, MSE and MAPE of the partially linear semi-parametric models of wheat, rice and maize samples are all very small. That is to say, the models built in this paper are well fitted.

TABLE 7. Model evaluation biomarkers

	MAE	MSE	MAPE
Rice	0.0043	0.000019	0.18%
Wheat	0.0060	0.000039	0.19%
Maize	0.0189	0.000398	2.14%

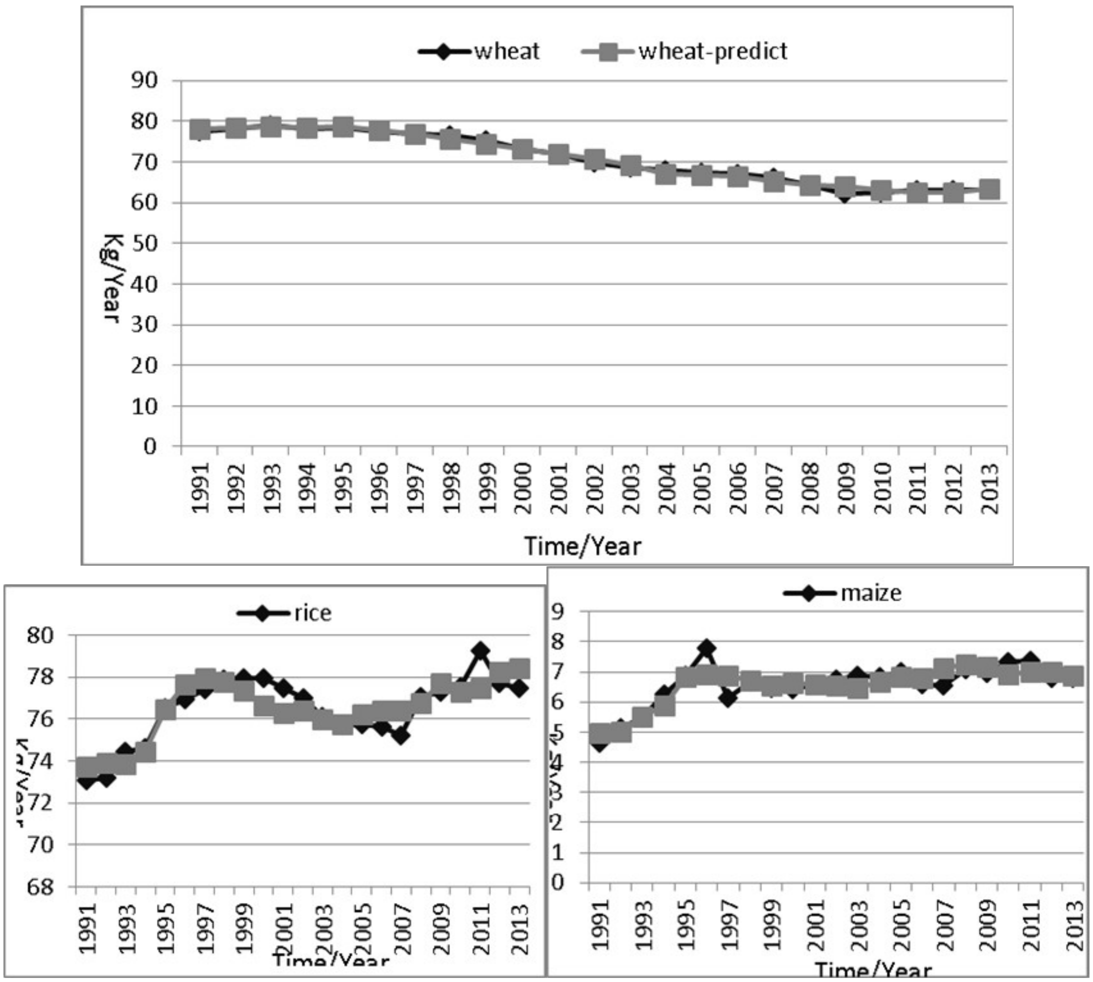

FIGURE 3. Prediction chart of partially linear semi-parametric model of wheat, rice and maize sample

Figure 3 is the prediction plot of the partially linear semi-parametric model of cereal samples. As the charts show, the models built in this paper can accurately forecast the consumption of cereal, as the prediction values approximately agree with the factual values. The most accurate is the model of the wheat sample, whose model evaluation biomarkers are also the lowest.

To study the balance of the diet of Chinese residents and whether the consumption of cereal is reasonable, we will forecast the per capita cereal consumption of the next ten years based on the models' built-in preliminaries. The predictions of independent variables are obtained by the five-year moving average method. As mentioned above, the average per capita consumption of rice, wheat and maize in the sample range is respectively 76.48 kg/year, 70.82kg/year and 6.58 kg/year, and the total is 153.88kg/year. The results in Table 8 show that the per capita consumption of cereal by 2025 will be 147.93kg/year, in which rice, wheat and maize will be 76.48 kg/year, 70.82kg/year and 6.58 kg/year, respectively. Compared with the average consumption of cereal from 1991 to 2013, the predicted value by 2025 will drop slightly to 3.87%. Meanwhile, there will be a modest increase in the per capita consumption of rice and maize by 2025, rising to 2.72% and 0.91% respectively, and a significant decline in the per capita consumption of wheat to 11.42%.

TABLE 8. Predicted values of cereal consumption for next ten years

Year	Rice	Wheat	Maize
2014	78.7262	62.8739	6.5731
2015	78.8560	62.3656	6.3548
2016	78.3440	62.7245	6.7555
2017	78.5156	62.7728	6.7079
2018	78.5703	62.8453	6.6512
2019	78.6022	62.7174	6.6070
2020	78.5775	62.6854	6.6138
2021	78.5219	62.7493	6.6668
2022	78.5575	62.7541	6.6492
2023	78.5659	62.7503	6.6376
2024	78.5650	62.7313	6.6349
2025	78.5575	62.7341	6.6404

According to the forecast, the per capita consumption of cereal will be 153.88kg/year by 2025, amounting to 422g/day. According to the *Dietary Guidelines for Chinese Residents 2016*, the daily recommended intake of cereal and potato is 250~400g, in which whole grain and mixed beans is 50~150g, and potato is 50~100g. The consumption of the main three cereals will exceed the recommended intake suggested by the dietary guidelines by 2025. It should be noted that there is a large gap between consumption and intake, which may be caused by food wastage and processing losses.

CONCLUSION

In this article, we conducted an empirical study on the per capita cereal consumption of Chinese residents and its influence factors based on the proposed combination of LASSO and the partially linear semi-parametric model. According to the research, there are several important findings. The factors affecting the per capita consumption of rice, wheat and maize are different from each other. The main influence factors of rice are Per Capita Disposable Income (PCDI), Urbanization Rate (UR) and Consumer Price Inflation (CPI). The following five factors have a significant impact on wheat consumption: Per Capita Disposable Income (PCDI), Food Expense (FE), Urbanization Rate (UR), Wheat Producer Price (WPP), and Consumer Price Inflation (CPI). Meanwhile, only two major factors have an effect on maize consumption: Per Capita Disposable Income (PCDI) and Aging Rate (AR). The results show that the three cereals have both common impact factors and differentiated ones. Per Capita Disposable Income (PCDI) is the common factor, which has a linear positive effect on the consumption of rice, wheat and maize. Urbanization Rate (UR) and Consumer Price Inflation (CPI) have an impact on both rice and wheat consumption. The former is linear negatively with rice consumption while the latter has a nonlinear relationship with wheat consumption. Wheat Producer Price (WPP) and Food Expense (FE) only affect wheat consumption, and Aging Rate (AR) is only the main factor of maize consumption, which has a nonlinear relationship with maize consumption.

The model constructed in this paper is well fitted according to the model evaluation indexes. The prediction shows that the per capita consumption of cereal is 153.88kg/year, of which the average per capita consumption of rice, wheat and maize will be 78.56 kg/year, 62.73kg/year and 6.64 kg/year respectively by 2025. According to the *Dietary Guidelines for Chinese Residents 2016* and the prediction data, the consumption of the main three cereals will be excessive by 2025, which may be caused by irrational dietary structure, food wastage and processing losses.

ACKNOWLEDGEMENTS

This work was supported by the Agricultural Science and Technology Innovation Project of Chinese Academy of Agricultural Sciences (CAAS-ASTIP-2018-IFND), the Special Scientific Research Fund of Agricultural Public

Welfare Profession of China (201503001), and Fundamental Research Funds for Central Non-profit Scientific Institution (1610422018001, 1610422018002).

REFERENCES

1. Chinese Nutrition Society, *Dietary Guidelines for Chinese Residents 2016*, (People's Medical Publishing House, Beijing, 2016), pp.3-29.
2. More than 80 percent of Chinese adults cereal intake is inadequate *(In Chinese)*, *Huanqiu Net*. Retrieved Aug. 06, 2018, from http://health.huanqiu.com/health_news/2016-05/8933544.html
3. Y. N. He, L. Y. Zhao, et al., Acta Nutrimenta Sinica, **38**, 115-118 (2016).
4. J. M. Nzuma,, R. Sarker, Agri. Econ., **41**, 43-50 (2010).
5. Z. H. Zheng, S.R. Henneberry, J. Agri. Applied Econ., **42**, 337-355 (2010).
6. L. Jiang, K. C. Seto, J. F. Bai, Chin. Agri. Econ. Review, **7**, 240-261 (2015).
7. V. Hovhannisyan, B. W. Gould, Agric. Econ., **42**, 1-17 (2011).
8. H. Valin, R. D. Sands, et al., Agric. Econ., **45**, 1-17 (2014).
9. X. Tian, X. H. Yu, Food Policy, **53**, 67-81 (2015).
10. R. Tibshirani, J. Royal Statis. Soc. B (Methodological), **58**, 267-288 (1996).
11. B. Efron, T. Hastie, I. Johnstone, R.Tibshirani, Ann. Statist., **32**, 407-499 (2004).
12. D. Ruppert,, M. P. Wand, R. J. Carroll, *Semiparametric Regression*. (Cambridge University Press, Cambridge, 2003), pp.25-47.
13. Z. J. Zhou, J. Agrote. Econ., **5**, 68-75 (2015).
14. R. Chen, X. H. Fu, J. Rural Econ., **7**, 37-41 (2012).
15. N.Abdelmonem, O. Adnen, Intern. J. Envir. Agri. Research, **2**, 142-148 (2016).