# Analysis and Prediction of Rural Residents' Living Consumption Growth in Sichuan Province Based on Markov Prediction and ARMA Model

**LU Xiao-li** *

School of Geography and Resources Science, Neijiang Normal University, Neijiang 641100, China

**Abstract** I select 32 samples concerning per capita living consumption of rural residents in Sichuan Province during the period 1978 -2009. First, using Markov prediction method, the growth rate of living consumption level in the future is predicted to largely range from 10% to 20%. Then, in order to improve the prediction accuracy, time variable t is added into the traditional ARMA model for modeling and prediction. The prediction results show that the average relative error rate is 1.56%, and the absolute value of relative error during the period 2006 -2009 is less than 0.5%. Finally, I compare the prediction results during the period 2010 -2012 by Markov prediction method and ARMA model, respectively, indicating that the two are consistent in terms of growth rate of living consumption, and the prediction results are reliable. The results show that under the similar policies, rural residents' consumer demand in Sichuan Province will continue to grow in the short term, so it is necessary to further expand the consumer market.

**Key words** Rural residents, Living consumption, Markov prediction, ARMA model

Rural residents' living consumption is an important component of consumer demand, drawing great attention from the scholars. The content of existing studies is focused on regional differences in rural residents' living consumption, consumer behavior, consumption structure, and factors influencing consumption. In terms of the research methods, I conduct analysis mainly using the regression analysis method, Extend Linear Expenditure System, least squares regression, factor analysis, habitual-preference life-cycle model [1-8]. From the existing researches, most of them are based on short-term consumption data, and the research methods are focused on regression analysis, with improving the level of consumption in rural areas and promoting economic development as purpose.

In fact, correctly predicting consumer trends and needs of rural residents, and timely establishing reasonable consumption structure, is also conducive to improving the level of consumption. Sichuan Province is a major agricultural province, with a large number of rural residents. Accurate prediction of rural residents' living consumption is conducive to fully giving play to the role of rural consumption in promoting economic development. However, when predicting the level of consumption, one-time specific prediction results will be inevitably affected by social, economic and policy factors, generating great deviation. If we can predict the probability of various results at the same time, and conduct comparative analysis between the probability and the results, it will further improve the credibility of prediction and help to grasp the future level of consumption. In view of this, I use the Markov prediction to predict the event probability, and use ARMA model with high accuracy of short-term prediction for verification. These two kinds of prediction methods both need a relatively large number of sample data and take into account the volatility of the data, avoiding the deviation arising from poor stationarity or data lack in prediction, improving the reliability of the prediction.

## 1 Data source and research methods

**1.1 Data source** I select the data on per capita living consumption of rural residents in Sichuan Province during the period 1978 -2009, including food, clothing, housing, household equipment, appliances and services, health care, transportation and communication, culture, education and recreation goods and appliances, other goods and services (Data are from *Statistical Yearbook* from 1979 to 2010).

And I determine the state in accordance with the annual growth rate($r_i$). It can be divided into the following states: substantial growth($E_1$, $r_i \geq 20\%$); moderate growth($E_2$, $10\% \leq r_i < 20\%$); slow growth ($E_3$, $0\% \leq r_i < 10\%$); negative growth ($E_4$, $r_i < 0\%$). Taking 1978 as the base year, the annual growth state of rural residents' per capita living consumption can be shown in Table 1.

**1.2 Research methods**

**1.2.1** Markov prediction. Markov prediction is a way to predict the probability of something to occur, a way to predict future variation in things based on the current situation of things. In the development process of event, Markov process means that the transfer of each state is only related to the state in the previous period, nothing to do with the past state. Markov prediction is to predict the probability of various states of event to occur in the Markov process [9].

Received: July 1, 2012 Accepted: August 14, 2012
* Corresponding author. E-mail: xiaolilu@126.com

**Table 1 State transition data of rural residents' living consumption in Sichuan Province**

| Year | Order number | $r_i$//% | State | Year | Order number | $r_i$//% | State |
|---|---|---|---|---|---|---|---|
| 1978 | 1 | – | – | 1994 | 17 | 39.67 | $E_1$ |
| 1979 | 2 | 18.12 | $E_2$ | 1995 | 18 | 17.35 | $E_2$ |
| 1980 | 3 | 12.24 | $E_2$ | 1996 | 19 | 27.21 | $E_1$ |
| 1981 | 4 | 15.40 | $E_2$ | 1997 | 20 | 6.71 | $E_3$ |
| 1982 | 5 | 13.13 | $E_2$ | 1998 | 21 | 0.02 | $E_3$ |
| 1983 | 6 | 10.99 | $E_2$ | 1999 | 22 | -1.02 | $E_4$ |
| 1984 | 7 | 8.96 | $E_3$ | 2000 | 23 | 4.45 | $E_3$ |
| 1985 | 8 | 9.70 | $E_3$ | 2001 | 24 | 0.54 | $E_3$ |
| 1986 | 9 | 12.55 | $E_2$ | 2002 | 25 | 6.27 | $E_3$ |
| 1987 | 10 | 12.03 | $E_2$ | 2003 | 26 | 9.78 | $E_3$ |
| 1988 | 11 | 22.44 | $E_1$ | 2004 | 27 | 15.10 | $E_2$ |
| 1989 | 12 | 11.05 | $E_2$ | 2005 | 28 | 13.09 | $E_2$ |
| 1990 | 13 | 7.51 | $E_3$ | 2006 | 29 | 5.31 | $E_3$ |
| 1991 | 14 | 8.49 | $E_3$ | 2007 | 30 | 14.71 | $E_2$ |
| 1992 | 15 | 3.09 | $E_3$ | 2008 | 31 | 13.86 | $E_2$ |
| 1993 | 16 | 13.69 | $E_2$ | 2009 | 32 | 32.40 | $E_1$ |

The prediction steps are as follows: (i) Determining the state and deriving the transition probability $p_{ij}$ of state $E_i$ to $E_j$ state. (ii) Combining $n$ possible states of an event in the process of development, to constitute the state transition probability matrix $P_{n \times n}$. (iii) According to the initial state $\pi(0)$, calculating the probability in the period $k$ in the state $E_j$, after $k$ state transitions. The formula is as follows:

$$\pi_j(k) = \sum_{i=1}^{n} \pi_i(k-1) p_{ij},\ j = 1, 2, \cdots, n \quad (1)$$

And then we get the forecast of probability of various states of this event in the period $k$.

**1.2.2** ARMA model. ARMA model (Auto-regressive Moving Average Model) is a well-known time series prediction method advanced by Box and Jenkins in the early 1970s [10]. ARMA model includes three basic types: autoregressive model AR ($p$), moving average model and autoregressive moving average model ARMA($p$, $q$).

Autoregressive moving average model can be expressed as follows:

$$x_t = \varphi_1 x_{t-1} + \varphi_2 x_{t-2} + \cdots + \varphi_p x_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q}$$

where $p$ and $q$ are the order of autoregressive model and moving average model, respectively; $x_t$ is the value of time series $\{X_t\}$ at time $t$; $\varphi_i(i=1, 2, \cdots, p)$ $\varepsilon_t$ is the autoregressive coefficient; $\varepsilon_t$ is the error or deviation of time series $\{X_t\}$ at time $t$; $\theta_j(j=1, 2, \cdots, q)$ *is the moving average coefficient.*

## 2 Results and analysis

**2.1 Markov prediction and analysis** According to the data in Table 1, the transition probability $p_{ij}(i, j=1, 2, 3, 4)$ from $E_i$ to $E_j$ is calculated, respectively; the state transition probability matrix P is derived as follows:

$$P = \begin{pmatrix} 0 & 0.6667 & 0.3333 & 0 \\ 0.2587 & 0.5000 & 0.2143 & 0 \\ 0 & 0.3333 & 0.58333 & 0.0833 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad (2)$$

The state of rural residents' living consumption in 2009 is denoted as $\pi(0) = (1, 0, 0, 0)$. Using formula (1), we can calculate the probability of various states to possibly occur in the period 2010 – 2012 (Table 2).

**Table 2 Prediction results of living consumption growth state in the period 2010 – 2012**

| Year | State | Probability |
|---|---|---|
| 2010 | $E_1$ | 0 |
| | $E_2$ | 0.666 7 |
| | $E_3$ | 0.333 3 |
| | $E_4$ | 0 |
| 2011 | $E_1$ | 0.190 5 |
| | $E_2$ | 0.444 4 |
| | $E_3$ | 0.337 3 |
| | $E_4$ | 0.027 8 |
| 2012 | $E_1$ | 0.127 0 |
| | $E_2$ | 0.461 6 |
| | $E_3$ | 0.383 3 |
| | $E_4$ | 0.028 1 |

From Table 2, we know that the probability of state $E_2$ (ordinary growth) is greater than that of other states, indicating that in the next few years, the growth rate of rural residents' living consumption may be with a high probability.

**2.2 ARMA model prediction and analysis**

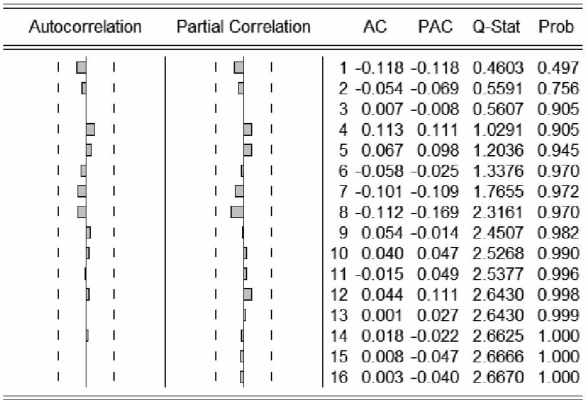
**2.2.1** Stationary analysis of data. The data on rural residents' per capita living consumption during the period 1978 – 2009 in Sichuan Province constitute time series $\{X_t\}$. Using , I draw the trend chart of $\{X_t\}$, and conduct ADF test, to judge the stationarity of time series. The test results are shown in Table 3.

**Table 3 ADF test results of sequence $\{X_t\}$ and $\{DDX_t\}$**

| Sequence | $t$-statistic | Critical value | | | Results |
|---|---|---|---|---|---|
| | | 1% | 5% | 10% | |
| $\{X_t\}$ | 2.131 907 | -3.670 170 | -2.963 972 | -2.621 007 | Non-stationary |
| $\{DDX_t\}$ | -5.705 253 | -3.679 322 | -2.967 767 | -2.622 989 | Stationary |

Table 3 shows that $t$-statistic of series $\{X_t\}$ is 2.131 907, greater than the critical value significant at level of 10%, which cannot reject the null hypothesis, and there is unit root, indicating that series $\{X_t\}$ is non-stationary series. After two difference operations, the time series is denoted as $\{DDX_t\}$. After ADF test, the $t$-statistic of $\{DDX_t\}$ is -5.705 253, smaller than the critical value significant at level of 1%, indicating that the null hypothesis can be at least rejected at 99% confidence level, and there is no unit root. Thus the series is stationary series.

**2.2.2** Model identification and establishment. Using software, I conduct auto-correlation and partial correlation analysis of the second-order differential series $\{DDX_t\}$ of $\{X_t\}$, and the results can be seen in Fig. 1. By Fig. 1, we can judge that AC and PAC are both at the random interval, in line with the characteristics of ARMA model.



| Autocorrelation | Partial Correlation | | AC | PAC | Q-Stat | Prob |
|---|---|---|---|---|---|---|
| | | 1 | -0.118 | -0.118 | 0.4603 | 0.497 |
| | | 2 | -0.054 | -0.069 | 0.5591 | 0.756 |
| | | 3 | 0.007 | -0.008 | 0.5607 | 0.905 |
| | | 4 | 0.113 | 0.111 | 1.0291 | 0.905 |
| | | 5 | 0.067 | 0.098 | 1.2036 | 0.945 |
| | | 6 | -0.058 | -0.025 | 1.3376 | 0.970 |
| | | 7 | -0.101 | -0.109 | 1.7655 | 0.972 |
| | | 8 | -0.112 | -0.169 | 2.3161 | 0.970 |
| | | 9 | 0.054 | -0.014 | 2.4507 | 0.982 |
| | | 10 | 0.040 | 0.047 | 2.5268 | 0.990 |
| | | 11 | -0.015 | 0.049 | 2.5377 | 0.996 |
| | | 12 | 0.044 | 0.111 | 2.6430 | 0.998 |
| | | 13 | 0.001 | 0.027 | 2.6430 | 0.999 |
| | | 14 | 0.018 | -0.022 | 2.6625 | 1.000 |
| | | 15 | 0.008 | -0.047 | 2.6666 | 1.000 |
| | | 16 | 0.003 | -0.040 | 2.6670 | 1.000 |

**Fig. 1 Autocorrelation and partial correlation analysis of sequence**

To further improve the prediction accuracy, the time variable t is first introduced in the model. Let $y_t = DDX_t$, we get the following model:

$$y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \cdots + \varphi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} + at \quad (3)$$

where $t = 1, 2, \cdots, 32$; $a$ is monomial coefficient.

Then using the SC Criterion [10], I add the time variable $t$ in the ARMA model of $\{y_t\}$, to determine the order. When $p = 3$, $q = 3$, $R^2$ reaches the maximum ($R^2 = 0.920\ 773$), and a set of values of AIC and SC reach the minimum (AIC = 9.351 274, SC = 9.591 244), indicating that $\{y_t\}$ is suitable for ARMA(3, 3) model. The model parameter results are shown in Table 4.

**Table 4 Model parameter estimation and test results**

| Variable | Coefficient | Standard deviation | *t*-statistic | Prob. |
|---|---|---|---|---|
| t | 0.118 142 | 0.454 146 | 0.260 141 | 0.797 2 |
| AR(2) | 0.353 207 | 0.141 974 | 2.487 825 | 0.020 9 |
| AR(3) | −0.270 966 | 0.153 503 | −1.765 222 | 0.091 4 |
| MA(2) | −0.261 986 | 0.681 729 | −0.384 296 | 0.704 4 |
| MA(3) | 3.868 338 | 0.709 464 | 5.452 483 | 0.000 0 |

In Table 4, $R^2 = 0.920\ 773$, the adjusted $R^2 = 0.906\ 369$, and D.W. = 2.055 223.

According to Table 5, the model of $\{y_t\}$ can be derived as follows:

$$y_t = 0.353\ 207 y_{t-2} - 0.270\ 966 y_{t-3} + \varepsilon_t + 0.261\ 986 \varepsilon_{t-2} - 3.868\ 338 \varepsilon_{t-3} + 0.118\ 142 t \quad (4)$$

**2.2.3** Model test and prediction. I conduct analysis of the stationarity of residual $\{\varepsilon_t\}$ of model $\{y_t\}$, and the results can be seen in Fig. 2. Then I conduct autocorrelation analysis of residual $\{\varepsilon_t\}$, and the results can be seen in Fig. 3.

From Fig. 2, we see that the residual $\{\varepsilon_t\}$ of model $\{y_t\}$ is stationary, with good fitting effect. From the autocorrelation analysis figure (Fig. 3), we see that the autocorrelation coefficients of residual series are all in the confidence interval, not significantly different from 0. It is judged as white noise sequence, and thus the model can be used for prediction.

Using Static Forecast method, I conduct test and prediction on $\{X_t\}$, and the prediction results are shown in Fig. 4. It can be seen from Fig. 4 that MAPE is 1.56, the Theil coefficient is 0.006, and CP approximates to 1 in the 3 ratios, indicating that the fitting accuracy of this model is high, and the prediction value is close to the true value.

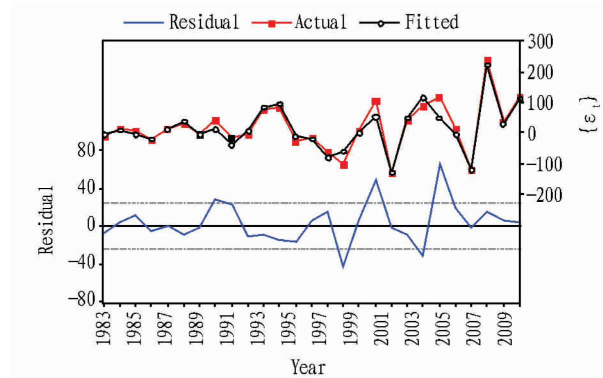
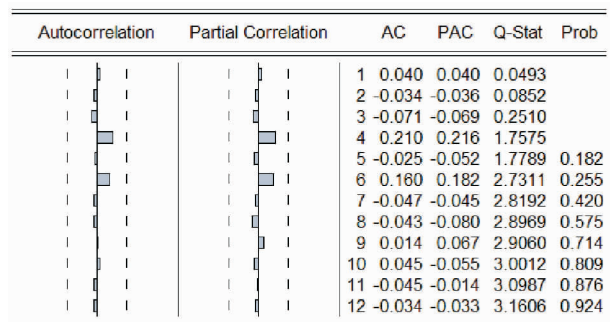



**Fig. 2 Stationarity analysis of residual sequence of model $\{y_t\}$**



| Autocorrelation | Partial Correlation | | AC | PAC | Q-Stat | Prob |
|---|---|---|---|---|---|---|
| | | 1 | 0.040 | 0.040 | 0.0493 | |
| | | 2 | -0.034 | -0.036 | 0.0852 | |
| | | 3 | -0.071 | -0.069 | 0.2510 | |
| | | 4 | 0.210 | 0.216 | 1.7575 | |
| | | 5 | -0.025 | -0.052 | 1.7789 | 0.182 |
| | | 6 | 0.160 | 0.182 | 2.7311 | 0.255 |
| | | 7 | -0.047 | -0.045 | 2.8192 | 0.420 |
| | | 8 | -0.043 | -0.080 | 2.8969 | 0.575 |
| | | 9 | 0.014 | 0.067 | 2.9060 | 0.714 |
| | | 10 | 0.045 | -0.055 | 3.0012 | 0.809 |
| | | 11 | -0.045 | -0.014 | 3.0987 | 0.876 |
| | | 12 | -0.034 | -0.033 | 3.1606 | 0.924 |

**Fig. 3 Correlation analysis of residual $\{\varepsilon_t\}$**

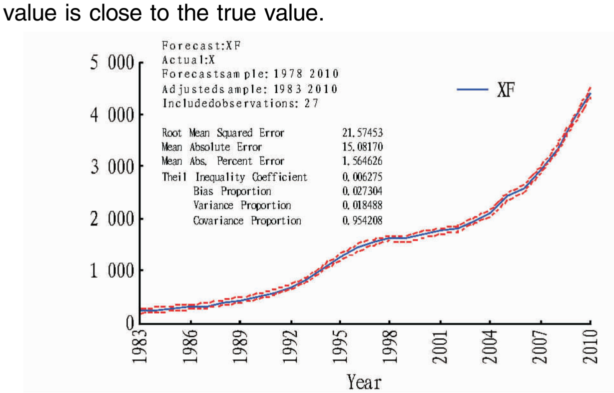



**Fig. 4 Test and prediction results of sequence $\{X_t\}$**

I use the model to fit and predict the annual rural residents' per capita consumption, and some results can be seen in Table 5.

From Table 5, we know that the relative error between the actual value and predicted value during the period 2006 – 2009 is less than 0.5%, once again indicating that the predicted value is close to the actual value of the model.

Through prediction results, I calculate the growth rate of rural residents' consumption in 2010, 2011 and 2012 at 13.74%, 13.42%, 11.56%, respectively, in line with ordinary growth. This result is consistent with that the level of consumption will experience ordinary growth trend with a great probability during the period 2010 – 2012, predicted using Markov prediction method in Table 2.

**Table 5 Prediction results of rural residents' living consumption in Sichuan Province in the period 2006 – 2012** Unit: yuan

| Year | Prediction value // yuan | Actual value // yuan | Relative error // % |
|---|---|---|---|
| 2006 | 2 572.81 | 2 572.00 | 0.03 |
| 2007 | 2 934.84 | 2 949.21 | 0.49 |
| 2008 | 3 355.57 | 3 362.00 | 0.19 |
| 2009 | 3 887.85 | 3 891.00 | 0.08 |
| 2010 | 4 425.75 | – | – |
| 2011 | 5 019.63 | – | – |
| 2012 | 5 600.02 | – | – |

## 3 Conclusions and recommendations

I select 32 samples concerning per capita living consumption of rural residents in Sichuan Province during the period 1978 – 2009. First, using Markov prediction method, the growth rate of living consumption level in the future is predicted. Then, in order to improve the prediction accuracy, ARMA model is used to verify the growth status.

Markov prediction results show that during the period 2010 – 2012, the probability of state $E_2$ (ordinary growth) of consumption growth is greater than that of other states, namely growing at the rate of 10% – 20%. ARMA model conducts short-term prediction of the living consumption value and growth rate during the period 2010 – 2012, and the prediction results are consistent with the Markov prediction results.

Thus we draw the following conclusions. Rural residents' living consumption in Sichuan Province will maintain growth rate of 10% – 20%, and the rural consumer market has brilliant prospects. However, the growth status of rural residents' consumption level is also affected by the income of rural residents, the rural market development status, consumption environment and other factors, thus we should take into account the following three aspects to ensure the stable growth of the level of consumption.

First, the steady growth of rural residents' income provides a guarantee for consumption growth.

Increase in farmers' income is inseparable from favorable internal and external environment. In terms of the external environment, the government needs to continue to focus on rural development, improve preferential policies to support agriculture and benefit farmers, and improve the level of social security in rural areas. At the same time, it should increase science and technology input into rural areas to enrich the farmers, further broaden the financing channels in rural areas, and establish the financial security system combining policy finance, cooperative finance and private fund raising. In terms of the internal environment, the farmers should actively implement the agricultural policy, constantly change and update awareness, improve literacy, further improve rural investment environment, give full play to the farming and breeding characteristics in Sichuan Province, actively develop processing enterprises of agricultural and sideline products, and broaden the income sources.

Second, it is necessary to improve the rural consumption market mechanism, and further expand rural consumer market, in order to meet rural residents' living consumption needs, improve the current consumption structure of the consumer market in rural areas and ease the regional imbalance of consumption.

It is necessary to further expand the rural production materials market; establish the commodity supply structure that is suitable for rural living consumption according to the distribution of rural residents, commodity supply adjustment, and the rural residents' consumption structure; increase the practical commodities to meet the rural residents' living needs, and solve the farmers' difficulty to buy.

Third, it is also necessary to change rural residents' consumer awareness, from passive and conservative consumption to active and open consumption, truly improving the standard of living.

For a long time, most of the consumption of rural residents has always been mainly used for the necessary means of production, housing construction, and children's education, but the consumer spending on other aspects is little. It is necessary to further optimize the consumption environment in rural areas, provide cheap and relatively high-grade goods, eliminate farmers' wariness in consumption of high-end consumer goods, and help the rural residents to increase consumer demand.

## References

[1] QI FQ, WANG ZW. Empirical analysis of rural residents consumption habits in Beijing City[J]. Chinese Rural Economy, 2007(7): 53 – 59. (in Chinese).

[2] FENG XY, LIU ZZ. Effects of consumption environment on rural residents consumption[J]. Problems of Agricultural Economy, 2008(6): 80 – 84. (in Chinese).

[3] YU HY, LIU JX, ZHANG HL. An analysis of peasants consumption in northeast China[J]. Economic Review, 2008(5): 52 – 54. (in Chinese).

[4] WU YZ. An analysis of the factors influencing the changes in the consumption structure of rural residents—empirical evidence from neo-classical economics and PLS method[J]. Economic Survey, 2008(6): 5 – 8. (in Chinese).

[5] XU HQ, LI JQ. Effects of different income sources on rural residents consumption and countermeasures[J]. Economic Review, 2009(3): 44 – 46. (in Chinese).

[6] LI WQ, YANG B. Analysis on variation factors of rural residents consumption structure in Chengdu City[J]. Consumer Economics, 2010(12): 21 – 24. (in Chinese).

[7] LIU DY. Regional characteristics of China rural residents consumption tendency[J]. Statistics and Decision, 2011(9): 103 – 105. (in Chinese).

[8] MA XH. The differential analysis of rural residents consumption status in four major economic zones[J]. Research on Development, 2011(2): 102 – 104. (in Chinese).

[9] XU JH. Surveying geography[M]. Beijing: Higher Education Press, 2006: 107 – 112. (in Chinese).

[10] PAN HY. Time sequence analysis[M]. Beijing: University of International Business and Economics Press, 2006: 62 – 112. (in Chinese).