

## Design for Forecasting System of Tobacco Sales Based on Time Series

Zhenyu Yang<sup>1</sup>

<sup>1</sup>China Tobacco Guangxi Industrial Co., Ltd. Nanning Guangxi, 530000, China  
[dreamgird@qq.com](mailto:dreamgird@qq.com)

### Abstract

*The sales forecast in the tobacco sales management plays a great role in the production and operation activities of an enterprise as a bridge between the production activity and the economic benefit. Forecasting system based on time series is put forward in accordance with the characteristics of tobacco sales by knowing about market supply and demand as well as development tendency at home and abroad through the sales forecast, with the help of studies on the sales forecasting technology. Time series is a sequence of random variables in chronological order and the time series analysis technique includes the moving average forecasting, exponential smoothing and regression analysis. According to the characteristics of tobacco sales data, the moving average forecasting model, exponential smoothing model and regression analysis model are established. Through programming, a forecasting system of tobacco sales is set up, which is used in combination with examples to simulate the forecasting model, and then model tests are conducted on the simulation results to get the appropriate model, which is used for sales forecast. The establishment of the system has an important influence on the management of tobacco companies. According to the forecast of tobacco sales, they will know the sales of the next period, so that they can take effective sales management strategies to improve the accuracy and efficiency of the decisions on the company's production and operation.*

**Keywords:** Tobacco sales; Forecasting technique; Time series forecast; Exponential smoothing; Regression analysis

### 1. Introduction

As it is about to break up the tobacco monopoly, it is required to carry out market segmentation and accurate market orientation from the point of view of sales by studying historical data of tobacco sales and forecasting the future sales. Measures should be taken to find the change rules of the sales data and make more effective and accurate forecast of tobacco sales, to meet the needs of market flexibly through sales decisions. These are very important to realize the transition from the tobacco monopoly to marketization, so that the tobacco industry can adapt to market competition.

Forecast is a data mining related technique, which is always based on a variety of time series or regression methods to find out useful trends or patterns from historical data and then use these trends or patterns to obtain the forecast values of the next period or cycle. A preliminary market analysis has been carried out to the tobacco industry to study some basic environmental factors influencing tobacco sales, including tobacco planting area, price, import and export volume, *etc.*, The results show that there are a variety of the factors that affect the sales of tobacco and these factors interact with each other. Tobacco sales are characterized by double trend changes of time series, namely the changes in the overall trend and the seasonal fluctuations. Data mining is a process of extracting the hidden and

unknown but potentially useful information and knowledge from large amounts of incomplete, noisy, fuzzy and random real application data. We must use it to mine information data deeply.

## 2. Methods of Time Series Forecast Modeling

### 2.1. Average Value Forecasting

As one of the advantages in forecasting, average value can eliminate any fluctuations, thus eliminating the noise (something that time series techniques cannot forecast), but at the same time, the average value also weakens the trend and seasonality.

$$F_{t+1} = \left( \sum_{i=1}^N S_t \right) / N \quad (1)$$

Where,  $F_{t+1}$  is the forecast value of the No.  $t+1$  period,  $N$  is the number of the selected period, and  $\sum_{i=1}^N S_t$  is the sum of the selected actual sales. Conclusion: the average value method can only be used to predict the sales data that contains only horizontality and noise, so the following method that is more advanced is used here instead of the average value to forecast for modeling.

### 2.2. Moving Average Modeling

(1) MA (Moving Average) is an improved version based on the average value forecasting, which is mainly about selection of the number  $N$  of the period. Its formula is expressed as:

$$F_{t+1} = (S_t + S_{t-1} + S_{t-2} + \dots + S_{t-N+1}) / N \quad (2)$$

Where,  $F_{t+1}$  is the forecast value of the No.  $t+1$  period and  $S_t$  is the actual value of the No.  $T$  period, characterized by that the moving average is closer to the average value method with a larger value of  $N$ , and the change forecast based on the moving average is stronger and the change is reflected faster with a smaller value of  $N$ .

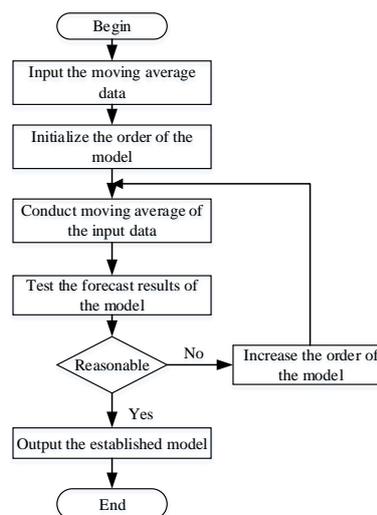


Figure 1. Flow Chart of Moving Average Modeling

Moving average modeling is an improvement made on the basis of the average value forecasting. As shown in Figure 1, the goal of conducting moving average of the input data is mainly to select the value of N, “moving period N”. After the value of N is selected, moving average is conducted according to the formula (2) and  $F_{t+1}$  is the result output by the model. The so-called model order  $M_j$  refers to the number of modeling times as needed, for example, for the first time of model building, the model order  $M_j=1$ ; if the results forecasted by the model are tested unreasonable, the model order  $M_j=2$  for the second time of model building. Model test is carried out in accordance with the characteristics of the sales forecast data and on the basis of the required model that has the smallest average error.

(2) Realization of moving average modeling and analysis of the forecast results

After entering the moving average forecasting technique module in the tobacco forecasting system, a new file named "Moving Average Forecasting Technique" will be opened by calling the component ExcleApplication and importing it into Excle in the Delphi system after pressing the key “Import New Data”.

At this time, according to the moving average modeling process, the moving average data of tobacco sales S reached from January 2013 to May 2014 is input for simulation while building the model, as shown in Figure 2, in order to facilitate the following explanation. For moving average of the input data, the “moving period  $N=2$ ” and the model order  $M_j=1$  are used in the first model. According to the moving average formula  $F_{t+1}=(S_t+S_{t-1}+S_{t-2}+ \dots +S_{t-N+1})/N$ , we get the results forecast based on the moving average of each month. Now we start to test if the forecast results of the model are reasonable. According to the characteristics of the sales forecast data, we use the average absolute error  $MAE=(\sum |E|)/N$  for testing the forecast results of the model and get that the average absolute error of the first model is 22.57. According to the test, the forecast results of the first model are not reasonable. Then we start the second moving average modeling, where “moving period  $N=4$ ” and model order  $M_j=2$  are used. After many times of modeling by repeating the same steps above, the forecast result of the fourth model is tested more reasonable with the smallest average absolute error.

So we take the fourth model where “moving period  $N=8$ ” for sales forecasting. According to the moving average formula (2), the forecast tobacco sales  $F_{t+1}$  of June 2014 is 1,145,000 tons, as shown in Figure 2 Forecast Results of Moving Average.

Month	Tobacco sales	Ft+1 (N=2)	Error	Ft+1 (N=4)	Error	Ft+1 (N=6)	Error	Ft+1 (N=8)	Error
2013.1	100								
2013.2	119								
2013.3	166	109.5	56.60						
2013.4	98	142.5	45.50						
2013.5	110	132	22.00	120.75	10.75				
2013.6	93	104	11.00	123.25	30.25				
2013.7	104	101.5	2.50	116.75	12.75	114.33	10.33		
2013.8	96	98.5	2.50	101.25	5.25	115.00	19.00		
2013.9	99	100	1.00	100.75	1.75	111.17	12.17	110.75	11.75
2013.10	103	97.5	5.50	98.00	5.00	100.00	3.00	110.63	7.63
2013.11	94	101	7.00	100.50	6.50	100.83	6.83	108.63	14.63
2013.12	102	98.5	3.50	98.00	4.00	98.17	3.83	99.63	2.38
2014.1	98	98	0.00	99.50	1.50	99.67	1.67	100.13	2.13
2014.2	120	100	20.00	99.25	20.75	98.67	21.33	98.63	21.38
2014.3	187	109	78.00	103.50	83.50	102.67	84.33	102.00	85
2014.4	99	153.5	54.50	126.75	27.75	117.33	18.33	112.38	13.38
2014.5	113	143	30.00	126.00	13.00	116.67	3.67	112.75	0.25
2014.6		106		129.75		119.83		114.50	
Average error			22.57		17.13		16.77		10.57

Figure 2. Forecast Results of Moving Average

## 2.3. ES Modeling

**2.3.1. Exponential Smoothing:** ES (Exponential Smoothing) is the basis for almost all time series techniques that are in use, which was originally known as “exponential weighted moving average”. Based on the following formula, it is easy to know that a larger value of  $a$  means the forecasting places much more stress on the role of the sales of this period in the forecasting process while the weight of the sales of other periods becomes smaller, and vice versa.

$$F_{t+1}=aS_t+(1-a)F_t \quad (3)$$

Where,  $F_{t+1}$  is the forecast value of the No.  $t$  period and  $S_t$  is actual;  $a$  is a number between 0 and 1, that is, the value forecast for the next period based on exponential smoothing is the function of the actual and forecast values of this period [9].

In determining the value of  $a$ , when using the exponential smoothing based forecasting method, it is very critical to determine an appropriate value of  $a$  for it directly affects the accuracy of the forecast value. The size of the value of  $a$  reflects its weight ratio with the forecast value that the greater the value of  $a$  is, the more the actual sales  $S_t$  of the No.  $t$  period contributes to the newly forecast value  $F_{t+1}$ , for example, when  $a = 1$ , the actual value of the No.  $t$  period is the forecast value of the No.  $t+1$  period; and the smaller the value of  $a$  is, the less the actual value of the No.  $t$  period affects the newly forecast value, for example, when  $a = 0$ , the forecast value of the No.  $t+1$  period is the actual value of the No.  $t$  period, at this time, there is no consideration of the latest information of the No.  $t$  period but the historic influence. Therefore, when we are more dependent on the recent information to make forecast, we can use a larger value of  $a$ , but if the historic influence is greater, a smaller value of  $a$  can be used. In practical application, the value of  $a$  can be selected based on experience and the specific conditions. The commonly used method is to make forecast with some different values of  $a$  and then compare their forecast errors to select the one that has the smallest error.

Conclusion of the influence of the value of  $a$  on the model:

(1) If the level change is the greater,  $a$  should get a large value, so that the forecast can be adjusted as soon as possible.

(2) If the data is more random,  $a$  should get a smaller value correspondingly, so that the exponential smoothing can eliminate more noise.

Adaptive exponential smoothing is based on automatic selection of the value of  $a$ , used for adjusting the value of a forecast for the next period according to the percentage error of the current forecast. Therefore, we can still use the exponential smoothing formula  $F_{t+1}=aS_t+(1-a)F_t$ , however, when the actual sales value of each period is obtained, the following formula should be used to adjust  $a$ :

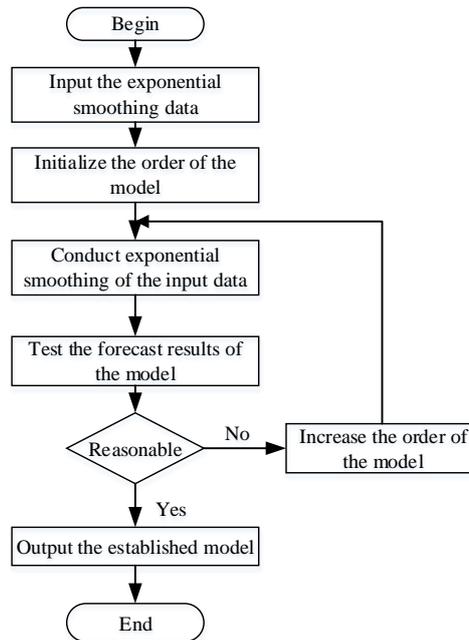
$$a_{t+1}=|(F_{t+1}-S_{t+1})/S_{t+1}|=|PE_{t+1}| \quad (4)$$

Because the value given by the above formula may be beyond the range of 0 to 1, it is required to make adjustment according to the following rules:

If  $|PE_{t+1}|$  is greater than or equal to 1, take  $a_{t+1}=0.9999$ ;

If  $|PE_{t+1}|$  is less than or equal to 0, take  $a_{t+1}=0.0001$ ;

### 2.3.2. Realization of ES Modeling and Analysis of the Forecast Results:



**Figure 3. Flow Chart of ES Modeling**

According to the same steps of the moving average modeling, now we directly enter ES modeling according to Figure 3, Flow Chart of ES Modeling, by inputting the data of tobacco sales  $S$  reached from January 2013 to May 2014 for simulation and conducting exponential smoothing to the input data to build the first model, where the “weighted coefficient”  $a = 0.2$  and the model order  $M_j = 1$ . According to the exponential smoothing formula  $F_{t+1} = aS_t + (1-a)F_t$ , the results forecast based on exponential smoothing for each month are obtained, as shown in Figure 4 Forecast Results of Exponential Smoothing.

Now we start to test if the forecast results of the model are reasonable. According to the characteristics of the sales forecast data, we use the average absolute error  $MAE = (\sum |E|) / N$  for testing the forecast results of the model and get that the average absolute error of the first model is 17.80. If we cannot tell if the forecast results that have been test are reasonable, we can now start the second ES modeling, where the “weighted coefficient”  $a = 0.3$  and the model order  $M_j = 2$  are used. After many times of modeling by repeating the same steps above, the forecast result of the first model ( $a = 0.2$ ) is tested more reasonable with the smallest average absolute error, as shown in Figure 4.

So we take the first model where the “weighted coefficient”  $a = 0.2$  for sales forecasting. According to the exponential smoothing formula (4), the forecast tobacco sales  $F_{t+1}$  of June 2014 is 1,160,600 tons, as shown in Figure 5 Forecast Results of Exponential Smoothing.

Month	Tobacco sales	Exponenti	Error	ExponentiError	ExponentiError	ExponentiError	ExponentiError	ExponentiError	ExponentiError	
2013.1	100									
2013.2	119	100	19.00	100.00	19.00	100.00	19.00	100.00	19.00	
2013.3	166	103.8	62.20	105.70	60.30	107.60	58.4	111.4	54.6	
2013.4	98	116.24	18.24	123.79	25.79	130.96	32.96	144.16	46.16	
2013.5	110	112.59	2.59	116.05	6.05	117.78	7.78	115.45	6.46	
2013.6	93	112.09	19.07	114.24	21.24	114.67	21.67	112.59	19.59	
2013.7	104	108.26	4.26	107.87	3.87	106.00	2.00	100.83	3.17	
2013.8	96	107.41	11.41	106.71	10.71	105.20	9.20	102.73	6.73	
2013.9	99	105.13	6.13	103.49	4.49	100.52	2.52	98.69	0.31	
2013.10	103	103.9	0.90	102.15	0.85	101.51	2.49	98.22	4.12	
2013.11	94	103.72	9.72	102.40	8.40	101.51	7.51	101.35	7.35	
2013.12	102	101.78	0.22	99.88	2.12	98.50	3.50	96.94	5.06	
2014.1	98	101.82	3.82	100.52	2.52	99.90	1.90	99.98	0.20	
2014.2	120	101.06	18.94	99.76	20.52	99.14	20.86	98.79	21.21	
2014.3	187	104.85	82.15	105.83	81.17	107.48	79.52	111.52	75.48	
2014.4	99	121.28	22.28	130.18	31.18	138.29	40.29	156.81	57.81	
2014.5	113	116.82	3.82	120.83	7.82	123.17	10.17	122.12	9.12	
2014.6		116.06								
Average error			17.80		19.11		19.98		21.13	23.31

**Figure 4. Forecast Results of Exponential Smoothing**

ES modeling is characterized by that the forecast value for the second month of each model is equal to the value of the tobacco sales  $S$  of the first month, taking the first model ( $a=0.2$ ) above for example, the forecast value of exponential smoothing for February 2013 is equal to the value of the tobacco sales in January 2013.

### 3. Regression Analysis

#### 3.1. Regression Analysis

Regression analysis is a mathematical method to study the dependence relationship between variables. Generally, the relationship between variables can be broadly divided into two types: the first is the function relationship where the relationship between variables is completely determined, that is, a variable can be determined by one or a number of other variables according to a certain rule; the second is the related-to relationship where there is a nondeterministic dependence relationship between variables, that is, variables are closely related to each other in quantity but we cannot accurately get the value of one variable according to one or a number of other variables except that the regularities of quantity changes among such variables can be distinguished on the basis of a large number of statistical data.

Regression analysis is to study this related-to relationship, which, although nondeterministic, is often found to have certain regularity based on a large number of observations. If the corresponding observed values of the two variables in a related-to relationship are regarded as the coordinates of the points on a Cartesian plane and these points are marked on the plane, a scatter diagram of points can be obtained. This diagram is called the scatter diagram of observed values. In general, the statistical regularity of the variable relation can be seen from the scatter diagram. Although the related-to relationship between variables is not a function, we can still use the corresponding function to express their regularity and such function is called regression function. Studying the relationship between variables, determining the regression function, forecasting and controlling change range of variables accordingly and so on constitute the contents of the regression analysis. In regression analysis, the first is to determine a variable as the dependent variable in many variables, with the rest serving as independent variables. When the forecast is carried out, the variable of the to-be-forecast value is a dependent variable. If only involving two variables, the regression analysis is called monadic regressive analysis; if involving more than two variables, it is called multivariate regression analysis.

We select a factor influencing most closely from many influencing factors of the forecast quantity (Y) as the independent variable (X) to set up a unary linear regression equation:

$$Y_t = \beta_1 + \beta_2 X_t + u_t \quad (5)$$

Where,  $\beta_1$  and  $\beta_2$  are parameter to be estimated,  $u_t$  is called a random error (the influence of other factors on  $Y_t$  except the linear influence of  $X_t$  on  $Y_t$ ). If the random errors generally obey the normal distribution  $N(0, \sigma^2)$  and keep independent of each other, we can fit the regression curves based on the least squares method in the observation samples of  $S_t$  and  $F_t$ . Such regression curve satisfies the least sum of squared residuals. Where,  $\hat{Y}_t$  is the estimated value of  $Y_t$ ,  $L_t$  is the residual,  $\sigma$  is the standard difference of all  $L_t$ s,  $\hat{\beta}_1, \hat{\beta}_2$  are the estimated values of  $\beta_1$  and  $\beta_2$  respectively. Based on this method, we can get:

$$\sum L_t^2 = \sum (Y_t - \hat{Y}_t)^2 = \sum (Y_t - \hat{\beta}_1 - \hat{\beta}_2 X_t)^2 \quad (6)$$

$$\hat{\beta}_2 = \frac{\sum (X_t - \bar{X})(Y_t - \bar{Y})}{\sum (X_t - \bar{X})^2} \quad (7)$$

$$\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2 \bar{X} \quad (8)$$

$$\hat{Y}_t = \hat{\beta}_1 + \hat{\beta}_2 X_t$$

Where,  $\bar{Y}$  and  $\bar{X}$  are sample averages of Y and X respectively. The regression model of the sample has been estimated so far, and the unary linear regression equation is  $\hat{Y}_t = \hat{\beta}_1 + \hat{\beta}_2 X_t$ .

Obtaining the regression curve by fitting based on the least square method does not mean that forecast can be made, for there is need to carry out necessary statistical tests.

Goodness-of-fit tests by determination coefficient  $r^2$

The goodness-of-fit refers to the excellence degree or grade of the observed data fitting by the sample regression line. The determination coefficient  $r^2$  is a comprehensive scale to indicate the fitting degree or grade of the data by the sample regression line, which measures the percentage of the deviation explained by the regression model in the total deviation of F based on the calculation formula as follows:

$$r^2 = 1 - \frac{\sum L_t^2}{\sum (Y_t - \bar{Y})^2} = \hat{\beta}_2^2 \frac{\sum (Y_t - \bar{Y})^2}{\sum (Y_t - \bar{Y})^2} \quad (9)$$

The value of  $r^2$  ranges from 0 to 1. A larger value of  $r^2$  indicates better fitting of the regression line, that is, the effect of the linear regression is better. The so-called "fitting" is to find the straight line through the data, so as to minimize the mean square error, namely to minimize the sum of squared distances from each point to the straight line.

"Standardized residual" means turning the residual  $L_t = Y_t - \hat{Y}_t$  into a standard value based on the formula as follows

$$K_t = (Y_t - \hat{Y}_t) / \sigma_t \quad (10)$$

Where,  $K_t$  is a standardized residual and  $\sigma$  is the standard difference of all  $L_t$ s.

### 3.2. Realization of Regression Analysis Modeling and Analysis of the Forecast Results

According to the same steps of the moving average modeling, here we directly enter regression analysis modeling according to Figure 5 Flow Chart of Regression Analysis Modeling.

First, we select a factor influencing most closely from many influencing factors of the forecast quantity (Y) as the independent variable (X) to set up a unary linear regression equation  $Y_t = \beta_1 + \beta_2 X_t$ . The tobacco sales S is taken as the dependent variable Y and cigarette consumption is taken as the independent variable X, the data of tobacco sales S reached from January 2013 to May 2014 as well as cigarette consumption is input for simulation, as shown in Figure 6. Regression analysis is conducted to the input data to establish the first regression model, where the model order  $M_j = 1$  and the estimating equation is  $\hat{Y}_t = \hat{\beta}_1 + \hat{\beta}_2 X_t$ .

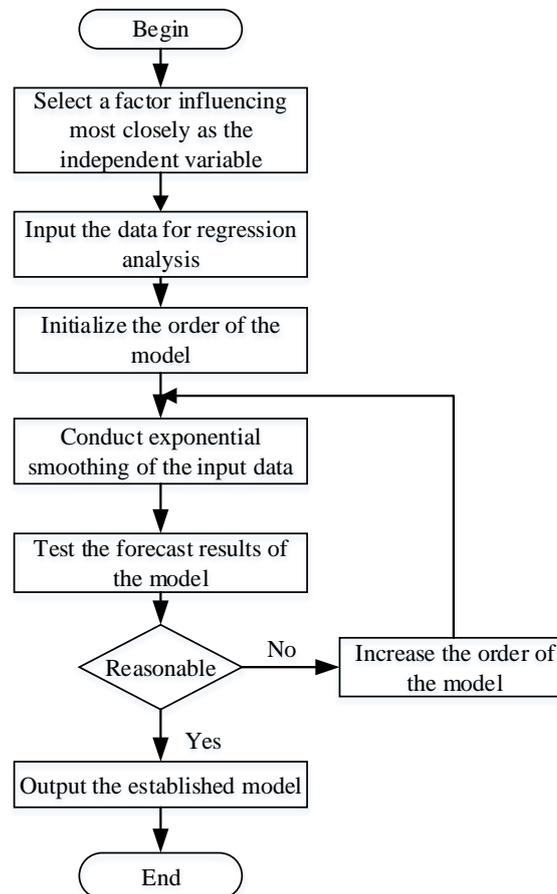


Figure 5. Flow Chart of Regression Analysis Modeling

Month	Tobacco sales	Cigarette consumption
2013.1	100	79
2013.2	119	83.24
2013.3	166	128.33
2013.4	98	80.67
2013.5	110	81.1
2013.6	93	76
2013.7	104	78.5
2013.8	96	76.42
2013.9	99	77.82
2013.10	103	80.57
2013.11	94	75.63
2013.12	102	79.47
2014.1	98	78.15
2014.2	120	83.86
2014.3	187	139.53
2014.4	99	76.5
2014.5	113	80.56
2014.6		92.6

**Figure 6. Input Data for Regression Analysis**

According to the regression analysis formulas (9) and (10), the obtained estimated coefficients  $\hat{\beta}_1$  and  $\hat{\beta}_2$  are -5.633 and 1.372 respectively.

Regression statistics	
Multiple	0.978
Determina	0.956
Adjustmer	0.953
Standard	5.596
Observed	17

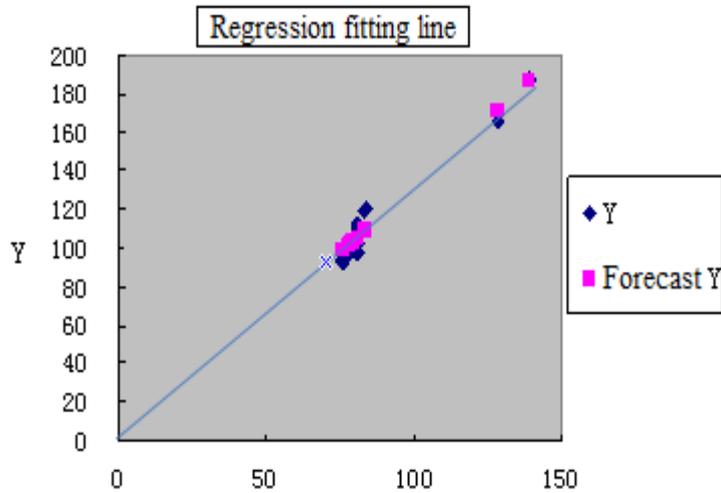
**Figure 7. Regression Analysis Coefficient**

	Coefficient	Standard error
Constant	-5.633	6.633
Variator	1.372	0.076

**Figure 8. Regression Statistics**

As shown in Figure 7, the regression model of the sample has been estimated so far, and the unary linear regression equation is  $\hat{Y}_i = -5.633 + 1.372X_i$ .

According to goodness-of-fit tested by the “determination coefficient  $r^2$ ” and regression statistics conducted based on the formula (10), we get  $r^2 = 0.956$ , as shown in Figure 8. A larger value of  $r^2$  indicates better fitting of the regression line, that is, the effect of the linear regression is better, so the first regression model tested by the “determination coefficient  $r^2$ ” meets the requirements. Its scatter diagram of the regression fitting line is as shown in Figure 9. As the scattered points are on both sides of the line, the fitting effect is good.



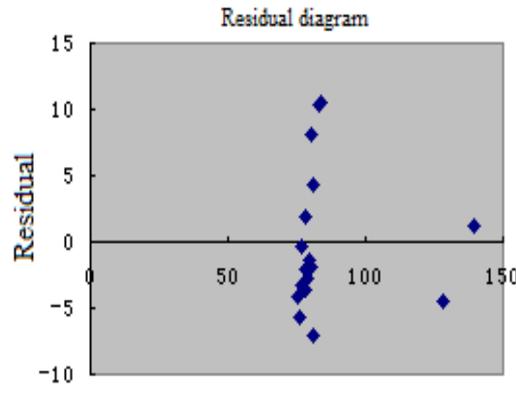
**Figure 9. Scatter Diagram of the Regression Fitting Line**

However, passing the above tests just shows that the linear relationship between variables X and Y is significant, or the linear regression equation is valid, but it cannot guarantee satisfying fitting of the data or rule out the data that is not entirely reliable due to unexpected reasons, such as abnormal values, periodic interference factors, *etc.*, Only when the assumptions related to the residual terms in the model are satisfied, can we safely use the regression model. Therefore, in the use of regression equations for analysis and forecast, we should use the residual diagram to help us test whether the model satisfies the basic assumptions in order to make further changes of the model.

According to the residual  $L_t = Y_t - \hat{Y}_t$ , residuals are obtained as shown in Figure 10. With the independent variable X used as the horizontal axis and the residual as the vertical axis, the corresponding residual residues are marked in the Cartesian coordinate system to get the residual diagram as shown in Figure 11.

Month	Forecast	Residual	Standardized residual
2013.1	102.756	-2.756	-0.509
2013.2	108.573	10.427	1.924
2013.3	170.438	-4.43	-0.819
2013.4	105.047	-7.047	-1.301
2013.5	105.637	4.363	-0.805
2013.6	98.64	-5.64	-1.041
2013.7	102.07	1.93	0.356
2013.8	99.216	-3.216	-0.594
2013.9	101.137	-2.137	-0.394
2013.10	104.91	-1.91	-0.53
2013.11	98.132	-4.132	-0.763
2013.12	103.401	-1.401	-0.259
2014.1	101.59	-3.59	-0.663
2014.2	109.424	10.576	1.952
2014.3	185.804	1.196	0.221
2014.4	99.326	-0.326	-0.06
2014.5	104.896	8.104	1.496

**Figure 10. Residual Analysis**



**Figure 11. Scatter Diagram of Residuals**

As shown in Figure 10, according to the residual analysis of the regression model, the scattered points of residuals of the model are on both sides of the residual based 0 axis, and the changes of the scattered points on both sides are basically the same, largely indicating that the regression model meets the basic assumptions. Therefore, we can use this model to forecast the sales. As  $\hat{Y}_t = -5.633 + 1.372X_t$ , by substituting the tobacco sales  $X_t$  of June 2014, namely 926,000 tons as shown in Figure 11, into the formula above, we can obtain the value  $\hat{Y}_t$  of tobacco sales forecast for June 2014, that is, 1,214,140 tons.

#### 4. Conclusions

Taking into full account of the characteristics of the target data source interdependencies of time series and regression analysis based forecasting techniques, we take the tobacco sales of a tobacco company in China reached in the past few years as the research model to set up a digital forecasting system of tobacco sales, which incrementally extracts data in real time from the data base of the tobacco company for simulation and model tests, thus calculating the sales of the next year to make the production plan. According to the comparison between the simulation results and the experimental results, there is little difference, which indicates the method proposed in this paper is feasible to some extent.

#### References

- [1] K. Wang, X. Zhou and T. Li, "Optimizing load balancing and data-locality with data-aware scheduling", *Big Data (Big Data)*, 2014 IEEE International Conference on. IEEE, (2014), pp. 119-128.
- [2] L. Zhang, B. He and J. Sun, "Double Image Multi-Encryption Algorithm Based on Fractional Chaotic Time Series", *Journal of Computational and Theoretical Nanoscience*, vol. 12, (2015), pp. 1-7.
- [3] T. Su, Z. Lv and S. Gao, "3d seabed: 3d modeling and visualization platform for the seabed", *Multimedia and Expo Workshops (ICMEW)*, 2014 IEEE International Conference on. IEEE, (2014), pp. 1-6.
- [4] Y. Geng, J. Chen, R. Fu, G. Bao and K. Pahlavan, "Enlighten wearable physiological monitoring systems: On-body rf characteristics based human motion classification using a support vector machine", *IEEE transactions on mobile computing*, vol. 1, no. 1, (2015) April, pp. 1-15.
- [5] Z. Lv, A. Halawani and S. Feng, "Multimodal hand and foot gesture interaction for handheld devices", *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, vol. 11, no. 1, (2014), pp. 10.
- [6] G. Liu, Y. Geng and K. Pahlavan, "Effects of calibration RFID tags on performance of inertial navigation in indoor environment", *2015 International Conference on Computing, Networking and Communications (ICNC)*, (2015) February.

- [7] J. He, Y. Geng, Y. Wan, S. Li and K. Pahlavan, "A cyber physical test-bed for virtualization of RF access environment for body sensor network", *IEEE Sensor Journal*, vol. 13, no. 10, (2013) October, pp. 3826-3836,
- [8] W. Huang and Y. Geng, "Identification Method of Attack Path Based on Immune Intrusion Detection", *Journal of Networks*, vol. 9, no. 4, (2014) January, pp. 964-971.
- [9] X. Li, Z. Lv and J. Hu, "XEARTH: A 3D GIS Platform for managing massive city information", *Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA)*, 2015 IEEE International Conference on. IEEE, (2015), pp. 1-6.
- [10] J. He, Y. Geng, F. Liu and C. Xu, "CC-KF: Enhanced TOA Performance in Multipath and NLOS Indoor Extreme Environment", *IEEE Sensor Journal*, vol. 14, no. 11, (2014) November, pp. 3766-3774.
- [11] N. Lu, C. Lu, Z. Yang and Y. Geng, "Modeling Framework for Mining Lifecycle Management", *Journal of Networks*, vol. 9, no. 3, (2014) January, pp. 719-725.
- [12] J. He, Y. Geng and K. Pahlavan, "Toward Accurate Human Tracking: Modeling Time-of-Arrival for Wireless Wearable Sensors in Multipath Environment", *IEEE Sensor Journal*, vol. 14, no. 11, (2014) November, pp. 3996-4006.
- [13] Z. Lv, A. Halawani and S. Fen, "Touch-less Interactive Augmented Reality Game on Vision Based Wearable Device", *Personal and Ubiquitous Computing*, vol. 19, no. 3, (2015), pp. 551-567.
- [14] G. Bao, L. Mi, Y. Geng, M. Zhou and K. Pahlavan, "A video-based speed estimation technique for localizing the wireless capsule endoscope inside gastrointestinal tract", 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), (2014) August.
- [15] D. Zeng and Y. Geng, "Content distribution mechanism in mobile P2P network", *Journal of Networks*, vol. 9, no. 5, (2014) January, pp. 1229-1236.
- [16] W. Gu, Z. Lv and M. Hao, "Change detection method for remote sensing images based on an improved Markov random field", *Multimedia Tools and Applications*, (2015), pp. 1-16.
- [17] Z. Chen, W. Huang and Z. Lv, "Towards a face recognition method based on uncorrelated discriminant sparse preserving projection", *Multimedia Tools and Applications*, (2015), pp. 1-15.
- [18] J. Hu and Z. Gao, "Distinction immune genes of hepatitis-induced hepatocellular carcinoma", *Bioinformatics*, vol. 28, no. 24, (2012), pp. 3191-3194.

## Authors

**Zhenyu Yang** received the M.S. degree in Electronic Engineering from Kyushu Kyoritsu University in Fukuoka, Japan. He is currently an Engineer in the China Tobacco Guangxi Industrial CO, LTD. His research interest is mainly in the area of Computer Software, Mechanical and Electrical Integration. He has published several research papers in scholarly journals in the above research areas and has participated in several books.