# Rural Grain Production Forecast Based on Combination Forecasting Model

**Ming-Chang Lee[1,*], Li-Er Su[2]**

[1]National Kaohsiung University of Applied Science, Taiwan
[2]Shih Chien University, Kaohsiung Campus, Kaohsiung Taiwan
*Corresponding Author

Abstract: Econometrics forecasting exists accuracy problem, general rural grain production forecast method have Cobb-Douglas production function model, multiple linear regression model. With the difference of forecasting results between these two methods, therefore, in this paper introduces combination forecasting model to solve the above problem. Combination forecasting model adopted to improve prediction accuracy. The contributions of this paper build Cobb-Douglas production function model, multiple linear regression models, and combination forecasting model and offered procedures in its entirety solutions. This paper indicates combination forecasting model optimal solution and uses the calculated error metrics to measure the accuracy of these three prediction methods. By analyzing different indicators, we find that the combination forecasting model is superior to two other models in application. This paper can be used as an example about the forecast theory and practice problem.

**Keywords:** Cobb-Douglas production function model; multiple linear regression model; Combination Forecasting Mode; Quadratic programming model ; Lagrange function

## 1. Introduction

In the complex socio-economic system, there will be many complex factors influence. Forecasters often on the same prediction, in different forecasting methods, there have different conditions hypothesis. Forecasters use different forecasting methods and establish variety of forecasting models. Selection the best statistical hypothesis testing results of a variety of model, and to exclusion of other prediction models; this is not the best way to improve forecast accuracy. This study use combination forecasting model, and verify that it is the best way to predict accurate.

Regional Farmers grain production forecast methods are Cobb-Douglas production function model (Cobb and Douglas, 1928 ), ARIMA model (Sohail et al., 1994,), multiple linear regression model (Mehnatkesh et al. 2010), Artificial Neural Networks (Cheng et al. 2011), Grey Neural Network Model (Zhu et al. 2011), ARIMA Model (Raghavender, 2010), Combination Forecasting Model (Sohail et al. 1994; Ding, 2007). Xiang and Zhang (20213) use weighted Markov chain model, regression model, and grey method to predict grain production. Chen et al. (2011) utilizes ARIMA model, BP neural network and the combination of stepwise regression analysis with BP neural network to predict the grain production. Dong et al. (2010) use regression method based on SVM classification in grain production forecasting. Hossain et al. (2012) considers Cobb-Douglas production function with error residual and Multicollinearity error term. Yang and Li (2015) use the combination forecasting of grain production based on stepwise regression method and RBF Neural network. Prediction model based on econometric methods can be broadly divided into two categories. One is structure prediction model, which uses a priori laws of economic theory or mathematical statistical model established between the numbers of variable relationships. Another is non-structure prediction model, namely the use of economic variables among the number of time-series. In this study, two structures prediction model, Cobb-Douglas production function model and multiple linear regression models.

## 2. The specific modeling steps of OLS estimate

The specific modeling 7 steps of ordinary least squares (OLS) estimate as follows:

Step 1: Initial smooth test. Selection the predictive variables, use time series of the scatter or line chart of the sequence to determine the initial smooth.

Step 2: Testing multi-collinearity

Multicollinearity is said to be problem when variance inflation factors of one or more predictors become large. If some covariates are collinear, the ordinary least square (OLS) estimates of these parameters will have a large variance. (Kennedy 2008). There are two common ways to examine the phenomenon (multicollinearity). (1) Scatter, plot this way a scatter plot is constructed and if multicollinearity exists the measurements should be scattered around a straight line. (2) A formal detection-tolerance or the variance inflation factor (VIF) for multicollinearity:

$$VIF_i = \frac{1}{1-R_i^2} \tag{1}$$

Where $R_i^2$ is the coefficient of determination of a regression of the i^th predictor on all other predictors, and $VIF_i$ is the variance inflation factor associated from the i^th predictor. When $VIF_i \geq 10$, it shows that $x_i$ is almost linear combination of several other independent variables. So, $x_i$ can be considered removed from the model (Haan, 2002).

Step 3 Model parameters estimation use EViews software

(1) Regression coefficient test (t-statistic, p-value)

Regression coefficient test is the use of t-statistic. Select the regression coefficients, all explanatory variables are included in the model, and then one by one the ability to explain not significant explanatory variables (that is, after the explanatory variables are not significant, namely, the value of p-value less than 0.05).

(2) Goodness of fit mode -- $R^2$ ( R-squared)

R-squared is to predict the goodness of fit mode degree. R-square can take on any value between 0 and 1, with a value closer to 1 indicating that a greater proportion of variance is accounted for by the model. The bigger value of R has a better fit model. Goodness of fit mode is the use of F-statistic.

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} \tag{2}$$

$$R_a^2 = 1 - (1 - R^2) \times [(n - 1)/(n - p)] \tag{3}$$

(3) Autocorrelation of residuals--(DW test)

The DW statistics test for autocorrelation of residuals, specifically Lag-1 autocorrelation. Assume the residuals flow a first- order autoregressive process $e_i = pe_{t-1} + n_t$, where $n_t$ is random and $p$ is the first-order autocorrelation coefficient of the residuals. The test for positive autocorrelation of residuals, the hypotheses for the D-W test can be written:

$$H_0 : p = 0 \tag{4}$$

$$H_1 : p > 0$$

The DW statistics is given by

$$d = \frac{\sum_{t=1}^{n} (e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2} \tag{5}$$

Ostrom (1990) showed that the residuals flow a first- order autoregressive process, d is the related to the first-order autocorrelation coefficient, p, as follows

$$d = 2(1 - p) \tag{6}$$

(4) Regression coefficients are equal to zero tests (F-statistic, Probe (F-statistic).

$H_0$ : All regression coefficients are equal to zero

$H_1$: At least one does not. An equivalent null hypothesis is that R-squared equals zero. (7)

The F-test determines whether the proposed relationship between the response variable and the set of predictors is statistically reliable, and can be useful when the research objective is either prediction or explanation

$$F = \frac{SSR/(p-1)}{SSE/(n-p)} \tag{8}$$

(6) Statistical model goodness fit --Akaik info criterion (AIC)

AIC is to assess the complexity of the statistical model and the statistical model to measure the "goodness fit". The value of AIC or SC is as small as possible. The AIC value of the model is the following.

$$AIC = 2p - 2\ln(L) \tag{9}$$

Where p is the number of parameters, L is the likelihood function. It assumes that error model of independent obey normal distribution. Let n be observation number and RSS is the residual sum of squares. The AIC value of the model can define as: Akaike, 1973; Aho et al. 2014)

$$AIC = 2p + nLog(RSS/n) \tag{10}$$

The Schwarz Information Criterion (SIC) is a measure to help in the selection between candidate models. Using this criterion, the best model is the one with the lowest SC. This criterion takes into account both the closeness of fit of the points to the model and the number of parameters used by the model (Schwarz, 1978).

$$SIC = -2(\frac{Log\ L}{n}) + \frac{pLog\ n}{n} \tag{11}$$

Step 4 Calculate the prediction error matrix E.

Step 5 Establish the convex quadratic programming model.

Step 6 (a) Does not Consider the weighting coefficient is non-positive, calculate $L = \frac{E^{-1}R}{R^T E^{-1} R}$

(b) Consider the weighting coefficient is non-negative, use Mathlib software to solve model (24).

Step 7 Calculate the common error indicators, it measure combination forecasting accuracy.

(a) Sum of the squared errors (SSE)

(b) Mean squared error (MSE)

(c) Mean absolute error (MAE)

(d) Mean absolute percentage error (MAPE)

(e) Square root of the average of the summing square forecasting errors (RMSE)

## 3. Building mode

### 3.1 Cobb–Douglas production function

Cobb–Douglas production function is a particular functional form of the production function, widely used to represent the technological relationship between the amounts of two or more inputs, particularly physical capital and labor, and the amount of output that can be produced by those inputs. Cobb–Douglas production function (Cobb and Douglas ,1928) model：

$$Y = AL^{\alpha} K^{\beta} \tag{12}$$

, where L, K, respectively, for the production process of labor and capital inputs, Y is the largest product.

Consider actual regional grain production, grain production selected as output variable Y. Crops sown area (S)**,** Agricultural machinery total power (K), amount of fertilizer (H), Irrigated area (G), Agricultural labor (L) is input variables. The model is:

$$Y_t = AS_t^{b1} K_t^{b2} H_t^{b3} G_t^{b4} L_t^{b5} u_t$$

This formulation implies that technical change is exogenous and disembodied. The above equation is usually estimated as

follows: $\log y_t = \log A + b_1 \log S_t + b_2 \log K_t + b_3 \log H_t + b_4 \log G_t + b_5 \log L_t + \log u_t$ (13)

Where $u_t$ is an error term.

**3.2 Multi- linear regression model**

**3.2.1 Multiple linear regression method**

Some research on standard reference of multiple linear regression method are a monograph on regression in time series context (Ostrom, 1990), a statistical text on regression (Weisberg, 1985). Multiple linear regression method used to model to linear relationship between the dependent variable and one or more independent variables. The model is:

$$y_t = b_0 + b_1 x_{1t} + b_2 x_{2t} + \ldots + b_p x_{pt} + e_t \qquad (14)$$

$x_{it}$ = value of i$^{th}$ predictor in year t

$b_0$ = Regression coefficient

$bi$ = Coefficient on the i$^{th}$ predictor

$p$ = Total number of predictors

$y_t$ = Predicted in year t

$e_t$ = Error term

The fitted values $\tilde{b}_0, \tilde{b}_1, \tilde{b}_2, \ldots, \tilde{b}_p$ estimate the parameters $b_0, b_1, b_2, \ldots, b_p$ of the population regression line. The residuals $e_t$ are equal to $y_t - \tilde{y}_t$ the difference between the observed and fitted values. The sum of the residuals is equal to zero.

The variance $\partial^2$ may be estimated by $s^2 = \dfrac{\sum e_t^2}{n-p-1}$, also known as the mean-squared error (or MSE).

In multiple linear regression method have six basic assumptions for regression model (Ostrom, 1990).

(1) Linearly: multiple linear regression models applied to a linear relationship. If the relationship is not linear, there are two resources: (1) transformed data, so that a linear relationship, or (2) the use of other statistical models (for example, neural networks, binary classification tree).

(2) Nonstochastic X: The errors are uncorrelated with the individual predictors. i. e. $E(e_t, x_{it}) = 0$.

(3) Zero mean: The expected value of the residuals is zero. i. e. $E(e_t) = 0$.

(4) Constant Variance: the variance of the residuals is constant. i. e. $E(e_t^2) = \sigma^2$.

(5) Nonautoregression: the residuals are random, or uncorrelated in time. i.e., $E(e_i, e_{i+m}) = 0, m \neq 0$

(6) Normality: the error term is normal distributed.

### 3.3 Combination forecasting model

Let n be the kind of prediction methods individual arithmetic weighting factor $l_i$ (i = 1,2, ..., n). The combination for forecasting of theoretical models:

$$y_t = l_1 x_{1t} + l_2 x_{2t} + \dots\dots + l_n x_{nt}, \text{ and } \sum_{i=1}^{n} l_i = 1 \quad (15)$$

Weighting factor $l_i$ for the selection method have arithmetic average, standard deviation method, variance reciprocal method, AHP method, Delphi method, the optimal weighting and combination forecasting method. Combination forecasting method is the best theoretical prediction method.

Let $x_{it}$ be the solution of i$^{th}$ period in t$^{th}$ model (method), $i = 1,2,3, \quad m, \quad t = 1, 2, .., N$.

$(l_1, l_2, ..., l_n)$: Arithmetic weighting coefficient of n kinds of single forecasting methods.

$\sum_{i=1}^{n} l_i = 1$

$x_t$ = combination forecasting value of t$^{th}$ period.

$e_{it} = (x_t - x_{it})$

The square sum of combination prediction error:

$$\begin{aligned} \delta &= \sum_{t=1}^{N} \sum_{i=1}^{m} \sum_{j=1}^{m} l_i \, l_j \, e_{it} \, e_{jt} \\ &= \sum_{i=1}^{m} \sum_{j=1}^{m} [l_i \, l_j (\sum_{t=1}^{M} e_{it} \, e_{jt})] \\ &= \sum_{i=1}^{m} \sum_{j=1}^{m} [l_i \, l_j \, E_{ij}] = L^T EL \end{aligned} \quad (16)$$

Where, $L = (l_1, l_2, ..., l_m)^T$,

$$E = \begin{bmatrix} e_1^2 & e_1 e_2 & & e_1 e_m \\ e_2 e_1 & e_2^2 & & e_2 e_m \\ & & & \\ e_1 e_m & e_2 e_m & & e_m^2 \end{bmatrix} \quad (17)$$

Assume $R = (1, 1, \dots\dots,1)^T$, $e_i = (e_{i1}, e_{i2}, \dots\dots, e_{im})^T$

$L$ is a column vector of combination weighting coefficients, $R$ is an m-dimensional row vector whose elements are 1. The following quadratic programming model：

$$\min \delta = L^T EL \quad (18)$$

s. t. $\quad R^T L = 1$

By Appendices A1 and A2

The solution of simultaneous equations is $L = \frac{E^{-1}R}{R^T E^{-1} R}$ . (19)

The stationary point of model is optimum solution.

$$\delta = L^T EL = (\frac{E^{-1}R}{R^T E^{-1}R})^T E(\frac{E^{-1}R}{R^T E^{-1}R}) = \frac{1}{R^T E^{-1}R} \quad (20)$$

The above prediction mode combination of computing weighting coefficients may occur weighting coefficient is negative. In forecasting practice, the negative interpretation of weighting coefficients in academia not achieved consensus. Therefore, combined forecasting model must be considered negative coefficients.

$$\min \delta = L^T EL \quad (21)$$

$$\text{s. t.} \quad \begin{aligned} R^T L &= 1 \\ L &\geq 0 \end{aligned}$$

This model is a nonlinear problem.

Eq. 21 is a convex quadratic programming problem. The solution of convex quadratic programming problem by uses Kuhn-Tucker (K-T) conditions. The K-T necessary and sufficient conditions are the optimum solution:

$$\begin{aligned} 2EL - \lambda R - U &= 0 \\ R^T L &= 1 \\ U^T 1 &= 0 \\ L \geq 0,\ U &\geq 0 \end{aligned} \quad (22)$$

Where, $\lambda$ is the Lagrange's multiplier of constraint of $R^T L = 1$. $U$ is k-T multiplier of the constraint of $L \geq 0$. The above conditions transform into linear programming model is:

$$\min v$$

$$\textbf{s. t.} \begin{cases} 2EL - (\lambda_1 - \lambda_2)R - U = 0 \\ R^T L + v = 1 \\ L \geq 0 \;\; U \geq 0 \;\; \lambda_1 \geq 0 \;\; \lambda_2 \geq 0 \end{cases} \quad (23)$$

The solution of Eq. 21 can be used SPSS, STAT, MATLAB software package etc.

## 4. Illustration

1994 - 2011 grain production selected as output variable Y, Crops sown area (S), Agricultural machinery total power (K), amount of fertilizer (H), Irrigated area (G), Agricultural labor (L) is input variables(see appendix).

### 4.1 Cobb–Douglas production model (C-D)

The sample data 1994-2004 input into (3), and use of econometric software EViews. The result of calculation is found that agricultural machinery total power (K), agricultural labor (L) is not significant at the reliability of 95%. We remove variable agricultural machinery total power (K) and agricultural labor (L). The outputs of econometrics' software are showed on table 1.

Table 1: Data fitting of Cobb–Douglas production model

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|---|---|---|---|---|
| Log A<br>Log S<br>Log G<br>Log H | 1.1458<br>1.8325<br>0.6274<br>0. 4663 | 0.15483<br>0.12335<br>0.16058<br>0.04142 | 2.38540<br>9.6165<br>3.9077<br>11.2584 | 0.00023<br>0.0000<br>0.0000<br>0.0000 |
| R-squared | 0.93864 | Mean dependent var. | | 10.36575 |
| Adjusted R-squared | 0.93373 | S. D. dependent var. | | 0.12358 |
| S. E. of regression | 0.03317 | Akaike info criterion | | -3.5482 |
| Sum square residual | 0.02751 | Schwarz criterion | | -3.4302 |
| Log likelihood | 52.7715 | F-statistic | | 142.492 |
| Durbin –Watson stat. | 1.0356 | Prob (F-statistic) | | 0.0000 |

From the table 1, we observe that all the coefficients are significant at the reliability of 95%. The R-squared is 093864 and Adjusted R-squared is 0.93864. Akaike info criterion is smaller than the Schwarz criterion. This shows that the model is good fitting. Log A, Log S, Log G and Log H Regression coefficients are not equal to zero tests.

$$\log y_t = 1.1458 + 1.8325 \log S_t - 0.6247 \log G_t + 0.4663 \log H_t \qquad (24)$$

**4.2 Multi- linear regression model (MLR)**

Since the correlation coefficient for each input variable (Crops sown area (S)**,** Amount of fertilizer (H), Irrigated area (G)) with output variable $Y_t$ could approximate linear representation. We can build a Multi- linear regression model to analysis and prediction.

$$y_t = \beta_0 + \beta_1 S_t + \beta_2 G_t + \beta_3 H_t \qquad (25)$$

Where, $\beta_0$, $\beta_1$, $\beta_2$, $\beta_3$ are multi- linear regression coefficients of this model.

The sample data 1994-2004 input into Eq. 25, and use of econometric software EViews. The result of calculation is found that agricultural machinery total power (K), agricultural labor (L) is not significant at the reliability of 95%. The outputs of econometrics' software are showed on table 2.

**Table 2:** : Data fitting of multi- linear regression production model

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|---|---|---|---|---|
| $\beta_0$<br>$\beta_1$ (S)<br>$\beta_2$ ( G)<br>$\beta_3 (H)$ | 14.284<br>0.0207<br>0.382<br>1.752 | 15.48325<br>0.12335<br>0.16058<br>0.04142 | 2.3704<br>10.6927<br>3.4561<br>10.2844 | 0.00023<br>0.0000<br>0.0000<br>0.0000 |

| | | | |
|---|---|---|---|
| R-squared | 0.93830 | Mean dependent var. | 10.36575 |
| Adjusted R-squared | 0.92548 | S. D. dependent var. | 0.12358 |
| S. E. of regression | 0.03456 | Akaike info criterion | -2.4785 |
| Sum square residual | 0.02815 | Schwarz criterion | -2.2178 |
| Log likelihood | 50.2146 | F-statistic | 159.19 |
| Durbin –Watson stat. | 1.12453 | Prob (F-statistic) | 0.0000 |

The R-squared is 093830 and Adjusted R-squared is 0.92548. Akaike info criterion is smaller than the Schwarz criterion. This shows that the model is good fitting. Constant, S, G and H of regression coefficients are not equal to zero tests.

$$y_t = 14.284 + 0.207S_t + 0.382G_t + 1.752H_t \tag{26}$$

4.3 **Combination Forecast model (CFM)**

The forecast results of Cobb-Douglas production model and Multi- linear regression model by use 2004-2014 data shown in Table 3. The error vector of Cobb–Douglas production model is $E_1 =$ (-20.1, -97.9, -37.1,106.2, -40.6, -13.5,-33.2 , -13.9, -13.1, 12.8). The error vector o Multi- linear regression model is $E_2 =$ (-0.7, -5.4, 1.9, 107.1, -137.4, 18.7, -16.6, 5.5. -3.5, -5.9) ◦

Table 3: The error vector of Cobb–Douglas production model and Multi-linear regression model

| Year | Actual output ① | C-D model Data fitting ② | MLR model Data fitting ③ | $E_1$=②-① | $E_2$=③-① |
|---|---|---|---|---|---|
| 2005 | 45649 | 45628.9 | 45648.3 | -20.1 | -0.7 |
| 2006 | 44510 | 44412.1 | 44504.6 | -97.9 | -5.4 |
| 2007 | 46662 | 46624.9 | 46663.9 | -37.1 | 1.9 |
| 2008 | 50414 | 50520.2 | 50521.1 | 106.2 | 107.1 |
| 2009 | 49457 | 49416.4 | 49319.6 | -40.6 | -137.4 |
| 2010 | 51210 | 51196.5 | 51228.7 | -13.5 | -18.7 |
| 2011 | 50839 | 50805.8 | 50822.4 | -33.2 | -16.6 |
| 2012 | 51228 | 51214.1 | 51233.5 | -13.9 | 5.5 |
| 2013 | 50828 | 50814.9 | 50834.5 | -13.1 | -3.5 |
| 2014 | 46217 | 46229.8 | 46211.1 | 12.8 | -5.9 |

$$E = (E_1, E_2)^T (E_1, E_2) = \begin{bmatrix} 17617.25 & 31084.99 \\ 31084.99 & 26104.78 \end{bmatrix} \quad E^{-1} = \begin{bmatrix} -3.47904E-05 & 2.17227E-08 \\ 2.17227E-08 & -5.5733E-05 \end{bmatrix}$$

$$\text{From Eq.19} \quad L = \frac{E^{-1}R}{R^T E^{-1}R} = \begin{bmatrix} 0.4029864 \\ 0.5970136 \end{bmatrix}$$

$$l_1 = 0.4029864, l_2 = 0.5970136 \tag{27}$$

Combination Forecast model is:

$$y = 0.402984 * (C-D) + 0.5970136 * MLR$$

## 4. Results and Discussion

Two types of error measurement are used in this study, which are MAPE, and RMSE. The MAE is percentage of the mean ratio of the error to the original data, which has been widely used as a performance measure in forecasting. RMSE is square root of the average of the summing square forecasting errors. The related formulas are expressed as followings:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y})^2} \tag{28}$$

$$MApE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \tilde{y}}{y_i} \times 100\right|$$

The actual grain production and the forecasting production from 2005-2014 are compared. Contrasting the results of Combination forecasting model which is based on the Cobb-Douglas model and multiple regression models, the result are listed below. The comparison results are shown in Table 4.

Table 4: Forecasting results of different models

| Year | Actual grains production | Cobb-Douglas model | Multiple regression model | Combination forecasting model |
|---|---|---|---|---|
| 2005 | 45649 | 45628.9 | 45648.3 | 45640.46 |
| 2006 | 44510 | 44412.1 | 44504.6 | 44467.31 |
| 2007 | 46662 | 46624.9 | 46663.9 | 46648.16. |
| 2008 | 50414 | 50520.2 | 50521.1 | 50520.71 |
| 2009 | 49457 | 49416.4 | 49319.6 | 49358.58 |
| 2010 | 51210 | 51196.5 | 51228.7 | 51215.70 |
| 2011 | 50839 | 50805.8 | 50822.4 | 50815.66 |
| 2012 | 51228 | 51214.1 | 51233.5 | 51225.88 |
| 2013 | 50828 | 50814.9 | 50834.5 | 50820.61 |
| 2014 | 46217 | 46229.8 | 46211.1 | 46218.61 |

Contrast analysis of the Cobb-Douglas model, multiple regression model, and Combination forecasting results and the error test. The results are listed below. Table 5 is denoted as comparison of three models.

Table 5: Comparison of three models

| Model | Cobb-Douglas model | multiple regression model | Combination forecasting |
|---|---|---|---|
| Root mean squared error (RMSE) | 51.092 | 55.753 | 48.784 |
| MAPE | 0.08085 | 0.0607 | 0.0603 |

All single model's RMSE and MAPE is large than Combination forecasting model's RMSE. It says that the combined forecasting model more effectively improves the prediction accuracy

## 5. Conclusion

This paper respectively makes use of Cobb-Douglas model and multiple regression model and combination forecasting model to establish forecasting model. Use train data (1994-2004) to establish Cobb-Douglas model and multiple regression model. We found that three major factors that influence grain production are sown area, irrigated area, and the amount of fertilizer. Increasing grain production is important to ensure that rural grain safety measures. Based on combination forecasting theory, the novel combination forecasting model of rural grain production has been established and we apply this kind of model to forecast the grain production during the 2005-2014. The numerical example proved that the combined forecasting model is superior to Cobb-Douglas model and multiple regression modes. The combined forecasting mode can ensure that rural grain safety measures.

**Reference**

[1] Aho, K.; Derryberry, D.; Peterson, T. (2014), Model selection for ecologists: the worldviews of AIC and BIC, Ecology **95**: 631–636,doi:10.1890/13-1452.1.

[2] Akaike, H. (1973), Information theory and an extension of the maximum likelihood principle, in Petrov, B.N.; Csáki, F., 2nd International Symposium on Information Theory, Tsahkadsor, Armenia, USSR, September 2-8, 1971, Budapest: Akadémiai Kiadó, p. 267-281.

[3] Bates, J. M. and Granger, C, W, J, (1969), the combination of forecasts, Operation Research Quarterly, 20(4): 451-488. DOI: 10.2307/3008764

[4] Chen, Z., Wang, H.,Song, T., Rong, B., Xiang, Y. and Hong,T. (2011), The contrast of several ways in grain production prediction, Journal of Convergence Information Technology (JCIT) 6(12): 248-256. doi:10.4156/jcit.vol6.issue12.31

[5] Cheng, L., Liu, Y., Cheng, H., Zhang, Y., Si, X. and Zhang, C. (2011), Growth Trends Prediction of Online Forum Topics Based on Artificial Neural Networks, JCIT: Journal of Convergence Information Technology, 6(10): 87-95, 2011.

[6] Cobb, C. W. and Douglas, P. H. (1928), A Theory of Production, American Economic Review **18** (Supplement): 139–165.

[7] Ding, C. F. (2007), Application of Combination Forecasting Method in Grain Production Forecast, Research of Agricultural Modernization, 28(1): 101-103.

[8] Dong, Y. Cheng, M. Zhang, Y. P. and Zhao, S. (2010), Regression method based on SVM classification and its application in production forecast, Journal of Computer Application, 30: 2310-2313.

[9] Haan, C. (2002), Statistical methods in hydrology, second edition. Iows State University Press, Ames, Iowa.

[10] Hossain, Md. M., Majumder, A. K., Basak, T. (2012), An application of Non-Linear Cobb-Douglas production function to selected Manufacturing industries in Bangladesh, Open Journal of Statistics, 2: 460-468.

[11] Kennedy, P. (2008). A Guide to Econometrics 6ed. Malden: Wiley-Blackwell.

[12] Mehnatkesh, A., Ayoubi, S. Jalalian A. and Dehghani, A. A. (2010), Prediction of rained wheat Grain yield and biomass using Artificial Neural Networks and Multiple Linear Regressions and determination the most factors by sensitivity analysis, Information Technology, Automation and

Precision Farming. International Conference of Agricultural Engineering - CIGR-AgEng 2012: Agriculture and Engineering for a Healthier Life, Valencia, Spain, 8-12 July 2012. 2010.
[13] Ostrom, C. W. (1990), Time series analysis, Regression Techniques, Second Edition" Quantitative Applications in the social Science, V. 07-009: Newbury Park, Sage Publications.
[14] Raghavender, M., (2010), Forecasting Paddy Production in Andhra Pradesh with ARIMA Model, International Journal of Agricultural and Statistical Sciences, 6(1): 251-258.
[15] Schwarz, G., (1978). Estimating the dimension of a model. The Annals of Statistics 6, 461–464.
[16] Sohail, A., Sarwar, A. and Kamran, M., (1994), Forecasting total food grains in Pakistan. Department of Mathematics and Statistics, University of agriculture, Faisalbad, Journal of Engineering and Applied Sciences. 13: 140- 146.
[17] Weisberg, S. (1985), Applied linear regression, 2nd ed., John Wiley, New York, 324 pp.
[18] Wen J. and Lei, L. (2010), A Combined Forecasting Method of Grain Yield in China based on GM (1,1) and BP network, Third International Conference on Information and Computing, 4-6 June, 2010.
[19] Xiang, C. S. and Zhang, L. F. (2013), Grain yield prediction model based on grey theory and Markov, Computer Sciences, 40: 245-248.
[20] Yang, L. and Li, B. (2015), The combination forecasting model of grain production based on step regression method and RBF neural network, Advanced Journal of food Science and Technology, 7(11): 891-895.
[21] Zhu, C. Sha Li, S., Wu, L., and Hao, Z., (2011), A Combined Grey Neural Network Model and its Application in Data Mining, JCIT: Journal of Convergence Information Technology, 6(7):94-101, 2011

Appendix: 1994-2014 Related Data

| Year | Actual grains production (Y) | Crops sown area (S) | Irrigated area (G) | amount of fertilizer (H) | Agricultural labor (L) | Agricultural machinery total power (K) |
|---|---|---|---|---|---|---|
| 1994 | 32502 | 11723 | 4488 | 1269 | 14745 | 7956 |
| 1995 | 35450 | 11495 | 4457 | 1406 | 15679 | 7990 |
| 1996 | 38728 | 11346 | 4417 | 1513 | 16614 | 8017 |
| 1997 | 40731 | 11404 | 4464 | 1659 | 18022 | 8073 |
| 1998 | 37911 | 112884 | 4445 | 1739 | 19497 | 8034 |
| 1999 | 40298 | 10884 | 4403 | 1775 | 20912 | 8075 |
| 2000 | 39408 | 11093 | 4422 | 1930 | 22950 | 8114 |
| 2001 | 40755 | 11126 | 4440 | 1999 | 24836 | 8162 |
| 2002 | 44624 | 11012 | 4437 | 2141 | 26575 | 8236 |
| 2003 | 43528 | 11220 | 4491 | 2357 | 28707 | 8316 |
| 2004 | 44266 | 11346 | 4740 | 2590 | 29388 | 8413 |
| 2005 | 45649 | 11231 | 4782 | 2805 | 30308 | 8462 |
| 2006 | 44510 | 11056 | 4859 | 2930 | 31816 | 8499 |
| 2007 | 46662 | 11050 | 4872 | 3151 | 33802 | 8534 |
| 2008 | 50414 | 10954 | 4875 | 3317 | 36118 | 8668 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 2009 | 49457 | 11006 | 4958 | 3593 | 38546 | 85947 |
| 2010 | 51210 | 11250 | 5038 | 3827 | 42015 | 8508 |
| 2011 | 50839 | 11291 | 5123 | 3980 | 45207 | 8475 |
| 2012 | 51228 | 11378 | 5229 | 4083 | 48996 | 8315 |
| 2013 | 50828 | 11316 | 5315 | 4124 | 52575 | 82038 |
| 2014 | 46217 | 10608 | 5382 | 4146 | 55172 | 8083 |

## Appendices

**A1**： $E$ is positive definite, then $E$ has an inverse matrix.

Proof: Assume $A=(e_1,e_2,....,e_m)$, then

$$A^T A=\begin{bmatrix} e_1^2 & e_1e_2 & & e_1e_m \\ e_2e_1 & e_2^2 & & e_2e_m \\ & & & \\ e_1e_m & e_2e_m & & e_m^2 \end{bmatrix}=E$$

Since $E$ is symmetric matrix. Assume $x=(x_1,x_2,...,x_m)$, $x\neq 0$ then $Ax\neq 0$ otherwise, it exists $x_1,x_2,...,x_m$ not all zero, such that $Ax=x_1e_1+x_2e_2+....,+_m e_m=0$

This is contradicted that $x_1,x_2,...,x_m$ is linearly independent.

When $x\neq 0$ and

$$x^T Ex=x^T A^T Ax=(Ax)^T(Ax)>0$$

Thus $x^T Ex$ is positive definite, this say the combination forecasting matrix $E$ is positive definite matrix.

**A2**：The Hesse matrix oft he objective function $\delta=L^T EL$ is positive definite, then this model have an optimum solution.

Proof:

The model is a quadratic programming problem containing constraints condition. Use Lagrange function

$$F(L,\lambda) = L^T EL + \lambda(R^T L - 1) \quad (1)$$

The necessary conditions for Eq.1:

$$\frac{\partial F}{\partial L} = 2EL + \lambda R = 0 \quad (2)$$
$$\frac{\partial F}{\partial \lambda} = R^T L - 1 = 0$$

From theorem 1, we know that $E$ is positive definite, then there have $E^{-1}$

From Eq.2, the solution of simultaneous equations is $L = \dfrac{E^{-1}R}{R^T E^{-1}R}$ (3)

The Hesse matrix of objective function $\delta = L^T EL$ is

$$H(\delta) = (\frac{\partial^2 \delta}{\partial l_i \partial l_j})_{m\times n} = \begin{bmatrix} e_1^2 & e_1e_2 & & e_1e_m \\ e_2e_1 & e_2^2 & & e_2e_m \\ & & & \\ e_1e_m & e_2e_m & & e_m^2 \end{bmatrix} = E$$

The stationary point of model is optimum solution.

$$\delta = L^T EL = (\frac{E^{-1}R}{R^T E^{-1}R})^T E(\frac{E^{-1}R}{R^T E^{-1}R}) = \frac{1}{R^T E^{-1}R} \quad (4)$$