Determining the Numbers of Time Series Observation for Rice Crisis Forecasting

Yuyun Hidayat, Titi Purwandari
Department of Statistics, Faculty of Mathematics and Natural Sciences,
Universitas Padjadjaran, Indonesia
yuyun.hidayat@unpad.ac.id, titipurwandari@yahoo.com

Sri Rijati Wardiani
Department of Literature and Cultural Studies, Faculty of Cultural Sciences,
Universitas Padjadjaran, Indonesia
sri.rijati@unpad.ac.id

Sukono
Department of Mathematics, Faculty of Mathematics and Natural Sciences,
Universitas Padjadjaran, Indonesia
sukono@unpad.ac.id

Abdul Talib Bon
Department of Production and Operations,
University Tun Hussein Onn Malaysia, Malaysia
talibon@gmail.com

Abstract

An objective visibility analysis has been developed as an integral part of the research "Identifying Indicators for Rice Crisis Forecasting in Asia". Analysis of visibility is a topic that received less attention in the study of forecasting models. Yet without a visibility analysis predicted results would be questionable whether the forecasting results can still be trusted. Analysis of visibility is required for assessing the ability of a model to forecast future events. In this paper we introduce visibility error which has different concept from the definition of error obtained during model building. Visibility error is the forecasting performance statistic which is needed as base in determining how long period of forecasting can be done. Without this statistics the forecaster can predict the next thousand years without clear statistical reasoning. Specifically this study developed error time variant visibility using inductive logic. One important finding raised in this paper is a formula for determining the numbers of data testing required to perform forecasting work based on specific forecasting specifications. The formula is: N = V + 2.

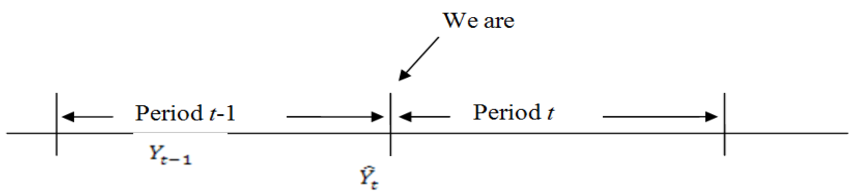
Keywords:
Visibility, error visibility, forecasting, error time variant

1. Introduction

Rice crisis is defined as a situation in which a prolonged uncontrollable riot occurs because of rice scarcity, (Hidayat, 2015). This definition is inspired by the thought that crisis can be stated compactly as a situation when intolerable bad things are happened, (Hidayat, 2013; Torbick et al., 2017). The severity of the situation during those years was widely reported. A summary report of food riots that erupted across the globe in 2008 outlines state responses to the food riots and sketches the state of democracy in countries where riots occurred (Mindi Schneider, 2008; Boansi, 2014).

Rice crisis Study in this paper having population as describred on working paper of (Hidayat, et al., 2017) while the underlying spirit of the rice crisis indicator validity are inspired by the main problem in modelling (Sukono, et al. , 2016; Choudhury and Jones, 2014).

Visibility is the ability of a model to predict the future. Measure of the accuracy and precision of forecasting methods will vary depending on how far the method can predict the future. Visibility is an important indicator of forecasting models. Because it demonstrates the ability of a model to predict the future. Visibility forecast performance shown by the statistics. For example, visibility two period ahead, show the eligibility of a model to forecast the next two periods. Without the forecast performance statistics we have no basis for forecasting the future. Because without these statistics so we can predict the next 1000 years without foundation. This means that through the study of visibility, a forecasting activity will avoid the problem of speculation. No visibility analysis makes us wonder whether the prediction still valid now.It is about the validity period of a forecasting model.



2. Framework

Differentiation of the developed model is given in the equation (1).

$$\mathbf{Y_i = \hat{Y}_i + \varepsilon_i,} \tag{1}$$

Y_i is the observation for variable response period i, and ε_i is a random experimental error, i=1,2.3...,n. Where i and n is an indicator that is distinctive compared with existing models because $\boldsymbol{\varepsilon_i}$, different for each forecasting period. In the existing model εi is derived from the model built for 1% forecasting accuracy. But can we really believe that this model gives 1% accuracy in predicting future value? Obviously not. There is no reason to believe that past accuracy can be used as a reliable indicator for the accuracy of future forecasts. That is why in this study εi was identified at the time of model testing.

Proposed time variant error models to solve problems of overfitting. Overfitting is less of a concern in forecasting. Overfitting associated with building a model in which the model pursued very fit with historical data to make the smallest possible $\boldsymbol{\varepsilon_i}$, the forecasting model in this research introduces n as the length of visibility, and this is a new issue in forecasting. $\hat{Y}_i$= f (R), is a forecasting model for predicting Y. The model is identified at the time of the model building. Equation (6) and forecasting models in general, assume error, e_t is obtained during the model building, while the errors in equation (1) are obtained during the testing model.

$$F_{t+1} = \alpha X_t + (1-\alpha)F_t \tag{2}$$

$$F_{t+1} = \alpha X_t + (1-\alpha)[\alpha X_{t-1} + (1-\alpha)F_{t-1}] \tag{3}$$

$$= \boldsymbol{\alpha X_t + \alpha(1-\alpha)X_{t-1} + (1-\alpha)^2 F_{t-1}]}$$

$$F_{t+1} = \alpha X_t + \alpha(1-\alpha)X_{t-1} + \alpha(1-\alpha)^2 X_{t-2} + \alpha(1-\alpha)^3 X_{t-3}$$
$$+\alpha(1-\alpha)^4 X_{t-4} + \alpha(1-\alpha)^5 X_{t-5} + \cdots + \alpha(1-\alpha)^{N-1} X_{t-(N-1)} \tag{4}$$

$$F_{t+1} = F_t + \alpha(X_t - F_t) \tag{5}$$

$$F_{t+1} = F_t + \alpha e_t \tag{6}$$

3. Model Development: Visibility wit Error Time Variant Approach

The concept of time variant visibility error is illustrated by the following time series data: t-3,t-2,t-1,t0; is the data set for the building model, and t1,t2,t3,t4; is 4 data used as data testing.

3.1 Eligibility visibility of two future periods:

For t-3, t-2, t-1, t0 are utilized for forecasting purposes for the two future periods (forecasting the value of t2) so as to obtain the forecast error value for the actual data t2, the first two-step ahead forecast performance is formulated as follows:

$$F_{t2} = F(\text{ t-3,t-2,t-1, t0})$$
$$e_{t2} = F_{t2} - t2$$

The second forecast performance two step is obtained by sequencing time series data: t-3, t-2, t-1, t0, t1 to predict the value of t3 so as to obtain forecast error for actual data t3 as follows:

$$F_{t3} = F(\text{ t-3,t-2,t-1, } t_0,\ t_1)$$
$$e_{t3} = F_{t2} - t3$$

The last thing to do is to use the time series data: t-3, t-2, t-1, and t0, and t1, t2 to forecast the value of t4 and get forecast error for the actual data t4 as follows:

$$F_{t4} = F(\text{ t-3,t-2,t-1, t0,t1,t2})$$
$$e_{t4} = F_{t2} - t4$$

So there are 3 errors forecast for two-step ahead forecast performance based on 4 data testing that is e_{t2}, e_{t3}, e_{t4}. Thus now we already have a statistic forecast performance if we want to forecast two step ahead forecasts that are average 3 error data obtained by exploring 4 data testing.

In this research the average information is declared eligible if there are at least 3 data, since one data is considered by chance, two data are considered too many possibilities, and three data are considered sufficient to recognize the pattern. As a consequence of this procedure, we obtain an average of 4 data errors for one step ahead forecast. This procedure calculates different errors for each forecasting period. Error for one step ahead period of forecast performance, error for two step ahead period of forecast performance, and so on. Since forecast errors are different for each forecast period then this approach is called an error time variant. This procedure calculates different errors for each forecasting period. Error for one step ahead period of forecast performance, error for two step ahead period of forecast performance, and so on. Since forecast errors are different for each forecast period then this approach is called an error time variant. The consequence of this procedure is that if data testing only provides 2 data then it is not eligible to forecast two-step ahead period of forecast.

4. Inductive Process: Visibility with Error Time Variant Approach

Based on case elaboration four data testing can be done induction as follows:

- Case of Two data test: provides forecasting statistical performance
  2 forecast of one step ahead period
  1 forecast of two step ahead period.
- Case of Three data: provides forecasting statistical performance
  3 forecast of one step ahead,
  2 forecast of two steps ahead,
  1 forecast of three steps ahead,
- Case of four data provides forecast statistical performance
  4 forecast of one step ahead,
  3 forecast of two steps ahead,
  2 forecast of three step ahead
  1 forecast of four steps ahead
- Case of five data provides forecast statistic performance
  5 forecast of one step ahead,
  4 forecast of two steps ahead,
  3 forecast of three steps ahead,
  2 forecast of four steps ahead,
  1 forecast of five steps ahead
- Case of six data provides forecasting statistical performance
  6 forecast of one step ahead,
  5 forecast of two steps ahead,
  4 forecast of three steps ahead,
  3 forecast of four steps ahead,
  2 forecast of five steps ahead,
  1 forecast of 6 steps ahead,

- Case of seven data provides forecast statistic performance
  7 forecast of one step ahead,
  6 forecast of two steps ahead,
  5 forecast of three steps ahead,
  4 forecast of four steps ahead,
  3 forecast of five steps ahead,
  2 forecast of 6 steps ahead,
  1 forecast of 7 steps ahead
- Case of eight data test data provides forecast statistic performance
  8 forecast of one step ahead,
  7 forecast of two step ahead,
  6 forecast of three step ahead,
  5 forecast of four step ahead,
  4 forecast of five step ahead,
  3 forecast of six ahead,
  2 forecast of seven step ahead
  1 forecast of eight step ahead

- Case of nine test data provides forecasting statistical performance as follows:
  9 forecast of one step ahead,
  8 forecast of two steps ahead,
  7 forecast of three steps ahead
  6 forecast of four steps ahead
  5 forecast of five steps ahead
  4 forecast of six steps ahead
  3 forecast of seven steps ahead
  2 forecast of eight steps ahead
  1 forecast of nine steps ahead

  And so on

5. Conclusion

Based on the above result of the induction process has been obtained a new technique to decide the validity period of a forecasting method. Forecaster cannot do forecasting without knowing the statistics forecasting performance of a forecasting model. Please note, forecasting performance statistics is only a necessary condition and not a sufficient requirement to know the substance of prediction accuracy. One important finding of this framework is that we now have a formula for determining how much data testing is required to perform forecasting work based on specific forecasting specific lengths. The formula is: $N = V + 2$, where N = Number of test data required, and V = visibility or validity period. According to the formula we need N=3 to obtain one step ahead forecast or one validity period. We need 5 data test to get three steps ahead forecasting.

Acknowledgements

Acknowledgments are conveyed to the Rector, Director of DRPMI, and the Dean of FMIPA, Universitas Padjadjaran, who has granted an Internal Grant program of Universitas Padjadjaran, which is a means of enhancing research and publication activities for researchers at Universitas Padjadjaran.

References

Boansi, D. Yield Response of Rice in Nigeria: A Co-Integration Analysis, *American Journal of Agriculture and Forestry*, 2014; 2(2): 15-24, Published online January 20, 2014 (http://www.sciencepublishinggroup.com/j/ajaf). doi: 10.11648/j.ajaf.20140202.11.

Choudhury, A. and Jones, J., Crop Yield Prediction Using Time Series Models, *Journal of Economic and Economic Education Research,* Volume 15, Number 3, 2014, pp. 53-68.

Hidayat, Y., Ismail Bin Mohd, I.B., Mamat, M., and Titi Purwandari, T. Developing Unprecedented Restlessness As The New Strong Indicator of Rice Crisis at National Level, *Global Journal of Pure and Applied Mathematics*, ISSN 0973-1768 Volume 11, Number 6 (2015), pp. 4139-4160, © Research India Publications, http://www.ripublication.com

Hidayat, Y., Purwandari, T., Sukono, and Y.D. Ariska, Y.D. Countries Population Petermination to Test Rice Crisis Indicator at National Level Using k-means Cluster Analysis, *Journal Article Open Access*, , 2017 1757-899X 166 012023 doi:10.1088/1757-899X/166/1/012023

Hidayat, Y., Mohd, I.B., Mamat, M., and Sukono, Forecasting Framework of Rice Crises Time in Bandung-Indonesia, *Applied Mathematical Sciences*, Vol. 7, 2013, no. 123, 6125 – 6140 HIKARI Ltd, www.m-hikari.com, http://dx.doi.org/10.12988/ams.2013.310555

Mindi Schneider , "We are Hungry!" A Summary Report of Food Riots, Government Responses, and States of Democracy in 2008– Cornell University.

Sukono, Hidayat, Y., Suhartono , Sutijo, B., Bon, A.T.B., Supian, S. Indonesian Financial Data Modeling and Forecasting by Using Econometrics Time Series and Neural Network, Global Journal of Pure and Applied Mathematics, Volume 12, Number 4 (2016), pp. 3745–3757 © Research India Publications, http://www.ripublication.com/gjpam.htm

Torbick, N., Chowdhury, D., Salas, W., and Jiaguo Qi, Monitoring Rice Agriculture across Myanmar Using Time Series Sentinel-1 Assisted by Landsat-8 and PALSAR-2, Remote Sens. 2017, 9, 119; doi:10.3390/rs9020119. http://www.mdpi.com/journal/remotesensing

Biographies

Yuyun Hidayat is a lecturer at the Department of Statistics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran. He currently serves as Deputy of the Head of Quality Assurance Unit, a field of Statistics, with a field of Quality Control Statistics. His doctoral thesis was on topic Identifying Indicator at National Level for Rice Crisis Forecasting

Titi Purwandari is a lecturer at the Department of Statistics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, the field of applied statistics, with a field of concentration of Quality Control Statistics.

Sri Rijati is a lecturer at the Department of Literature and Cultural Studies Faculty of Cultural Sciences Universitas Padjadjaran, Indonesia,

Sukono is a lecturer in the Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran. Currently serves as Head of Master's Program in Mathematics, the field of applied mathematics, with a field of concentration of financial mathematics and actuarial sciences.

Abdul Talib Bon is a professor of Production and Operations Management in the Faculty of Technology Management and Business at the Universiti Tun Hussein Onn Malaysia since 1999. He has a PhD in Computer Science, which he obtained from the Universite de La Rochelle, France in the year 2008. His doctoral thesis was on topic Process Quality Improvement on Beltline Moulding Manufacturing. He studied Business Administration in the Universiti Kebangsaan Malaysia for which he was awarded the MBA in the year 1998. He's bachelor degree and diploma in Mechanical Engineering which his obtained from the Universiti Teknologi Malaysia. He received his postgraduate certificate in Mechatronics and Robotics from Carlisle, United Kingdom in 1997. He had published more 150 International Proceedings and International Journals and 8 books. He is a member of MSORSM, IIF, IEOM, IIE, INFORMS, TAM and MIM.