

Title	タイのキャッサバ輸出予測に応用した時系列データ予測のハイブリッドモデルに関する研究
Author(s)	Pannakkong, Warut
Citation	
Issue Date	2017-09
Type	Thesis or Dissertation
Text version	ETD
URL	http://hdl.handle.net/10119/14821
Rights	
Description	Supervisor:Huynh Nam Van, 知識科学研究科, 博士

氏名	WARUT PANNAKONG		
学位の種類	博士(知識科学)		
学位記番号	博知第 208 号		
学位授与年月日	平成 29 年 9 月 22 日		
論文題目	A study on hybrid models for prediction of time series data with application to forecasting Thailand's cassava export (タイのキャッサバ輸出予測に応用した時系列データ予測のハイブリッドモデルに関する研究)		
論文審査委員	主査	Huynh Nam Van	北陸先端科学技術大学院大学 准教授
		神田 陽治	同 教授
		藤波 努	同 教授
		由井 隆也	同 准教授
		中森 義輝	同 非常勤講師
		村井 哲也	千歳科学技術大学 教授

論文の内容の要旨

Time series forecasting is an active research area that plays important role in planning and decision making in several practical applications. The main task of this research area is to improve the prediction accuracy.

This thesis proposes three novel hybrid forecasting models which are significantly extended from the Zhang's model and the Khashei and Bijari's model by involving the clustering algorithm (i.e. k -means) and the discrete wavelet transform (DWT) for inputs pre-processing. Additionally, instead of including only the lagged values of time series as the input variables, additional variables such as moving averages and annual seasonal index are included into the proposed model. The experiments are conducted comprehensively with several hybridization scenarios in term of structures and variables to find the most suitable forecasting model for Thailand's cassava export.

The first proposed hybrid model (so called ARIMA+ARIMA/ANN+ k -means/ARIMA/ANN) is the hybrid forecasting model involving autoregressive integrated moving average (ARIMA), artificial neural network (ANN) and the k -means clustering. These single models and the k -means clustering are used to build the forecasting models in different level of complexity (i.e. ARIMA; hybrid model of ARIMA and ANN; and hybrid model of k -means, ARIMA, and ANN). To obtain the final forecasting value, the forecasted values of these three models are combined with the weights generated from discount mean square forecast error (DMSFE) method.

The second proposed hybrid model (so called DWT/ARIMA/ANN) is the hybrid forecasting model of the DWT, the ARIMA, and the ANN without linear or nonlinear assumption on the approximation and the detail. The proposed model starts with decomposing the time series by the DWT to get the approximation and the detail. Then, the approximation and the detail are separately analyzed by the Zhang's model involving the ARIMA and the ANN in order to capture both linear and nonlinear components of the approximation and the detail. Finally, the linear and nonlinear components are additively combined for the final forecasting value.

These two novel hybrid models are applied to three well-known data sets: Wolf's sunspot, Canadian lynx, and exchange rate (British pound to US dollar) to evaluate the prediction capability in three measures (i.e. MSE, MAE, and MAPE). The prediction performance of the proposed models is compared to both the traditional single and hybrid models. The results imply that the proposed models give the best performance in MSE, MAE, and MAPE for all three data sets.

Then, the proposed hybrid models are implemented to Thailand's cassava export as the case study. In addition, we also propose the third novel hybrid model (so called ARIMA/ANN with pre-processed variables), which is the hybrid model of the ARIMA and the ANN with pre-processed variables for Thailand's cassava export forecasting. The experimental results indicate that the DWT/ARIMA/ANN model is the best model for the native starch and the sago. On the other hand, the ARIMA/ANN with pre-processed variables model is the best model for the modified starch.

In conclusion, all three proposed hybrid models have shown their forecasting capability over both the traditional single and hybrid models. Therefore, they can be used as the alternative models for time series prediction. Moreover, the stakeholders involves in the cassava supply chain can apply the proposed models specified for each type of the cassava export to obtain more accurate prediction results of Thailand's cassava export. The proposed models for cassava forecasting can be applied to other commodity products sharing the similar characteristic with the cassava as well.

Keywords: hybrid time series forecasting model, autoregressive integrated moving average (ARIMA), artificial neural network (ANN), k -means, discrete wavelet transform (DWT)

論文審査の結果の要旨

Time series forecasting is an active research field that has received considerable attention in a variety of application areas. During the last decade, several hybrid forecasting models have been proposed in order to take advantage of unique strength of conventional models such as autoregressive integrated

moving average (ARIMA) and artificial neural network (ANN) in linear and non-linear modeling. More recently, in order to improve prediction accuracy, self-organizing map (SOM) has been used for forming the time series into clusters before applying the hybrid model of ARIMA/ANN (Ruiz-Aguilar et al., 2015). However, the final prediction made by summation of the prediction values of individual clusters would potentially cause the overprediction problem. Moreover, the discrete wavelet transform (DWT) has been also applied to decompose the time series into approximation (trend) and detail (noise) before further analysis with the hybrid model of ARIMA/ANN (Khandelwal et al., 2015). The approximation, nevertheless, is assumed to be composed of only nonlinear patterns, which is somewhat impractical.

The research of this dissertation is aimed at developing novel hybrid models in order to overcome such limitations of the previously developed models. The research problem is well motivated and clearly formulated. An adequate overview of relevant theoretical and practical background literature is also provided in the dissertation. The theoretical contribution of this research is to propose two new hybrid models for time series forecasting: the first proposed model is a combination of k -means clustering, ARIMA and ANN, while the second proposed model is a hybridization of ARIMA, ANN and DWT without the assumption of linearity on the detail and the approximation. A set of comprehensive experiments have been conducted on well-known datasets such as Wolf's sunspot, Canadian lynx and Exchange rate (British pound to USD) to evaluate the prediction capability of the proposed models. Finally, a case study of forecasting Thailand's cassava export was conducted to demonstrate the practical applicability of these proposed models.

This dissertation has made an interesting contribution to theoretical and experimental developments within the area of time series forecasting and application. The research work presented has eventually resulted in two journal papers (one accepted and one under review), one book chapter and three refereed conference papers.

In summary, Mr. Warut Pannakkong has completed all the requirements in the doctoral program of the School of Knowledge Science, JAIST and finished the examination on August 1, 2017. All the committee members unanimously decided to pass the candidate.