

A Chinese New Year Aware ARIMA Model for Electricity Consumption Forecast

Tiao Cai, Xin Wang, Sui Huang

Department of Computer Science, JiNan University, Guangzhou 510632, China
ths@jnu.edu.cn

Abstract - The forecast of electricity consumption has a significant impact on both economics and society. Various models have been developed to make this forecast accurate. Among them is the highly successful ARIMA model which is based on time series analysis. However, when applied local in China, ARIMA forecast is not as accurate as expected, due to the moving holiday effect of Chinese New Year festive period. This paper proposes a revised ARIMA model which adjusts its forecast by the use of a China New Year (CNY) coefficient. Preliminary experimental results on real data show that this revised model outperforms the traditional one.

Index Terms - consumption forecast, time series, seasonal effect, ARIMA model

1. Introduction

Both the economic and population growths generate an ever-increasing demand for electric power. It is widely believed that the electricity generation and consumption can be used as an indicator to local economic trend. Therefore, it is of vital importance to make the electricity consumption forecast as accurate as possible.

Various approaches[1][5] have been proposed for electricity consumption forecast. They include regression analysis[4], time series model, wavelet analysis[3], growth rate analysis, and auto-regressive integrated moving average (ARIMA) model[8]. The ARIMA model is capable of analyzing non-linear periodic data and has been proved successful in practice.

There is a lack of balance in mainland China in terms of the demand and supply of electric power. One typical case is that its power consumption demonstrates some sharp escalation in its periodicity, which occurs during the traditional Chinese New Year (CNY) festive period. For simplicity, we call it CNY period in this paper. However, as the Chinese calendar is lunar-based, this CNY period may vary from late January to early March of a year. An ARIMA-model-based electricity consumption forecast has difficulties in dealing with this above uncertainty and is inaccurate during this CNY period[8].

To tackle this above problem, we propose a Chinese-new-year-aware ARIMA model. A time series model is first developed without the CNY period concern, before its prediction is adjusted with the CNY period factors for final output. We have evaluated the performance of our proposed approach with data of the electricity consumption in various cities between 2005 and 2011. The results show that it clearly outperforms the original ARIMA model.

The paper is organized as follows. Section II presents a brief review of time series analysis and the ARIMA model. The Chinese-new-year-aware ARIMA model is then explained in depth in section III. For conciseness concern, we summarize our experimental results in section IV, before giving some concluding remarks in section V.

2. Time series analysis and the ARIMA model

Time series can be roughly classified into two categories, stable and unstable. Linear models can be used to analyze stable time series, whilst auto-regressive integrated moving average (ARIMA) model can be used for unstable ones which are more common in practice.

ARIMA models are the most general class of models for forecasting a time series which can be stationarized by transformations such as differencing and logging.

Mathematically, the ARIMA model can be expressed as below

$$\begin{cases} \Phi(B)\nabla^d x_t = \Theta(B)\varepsilon_t \\ E(\varepsilon_t) = 0, \text{Var}(\varepsilon_t) = \sigma_\varepsilon^2, E(\varepsilon_t \varepsilon_s) = 0, s \neq t \\ Ex_s \varepsilon_t = 0, \forall s < t \end{cases} \quad (1)$$

In practice, some of the unstable time series demonstrate periodic variations, i.e. similarities arise between the data items which are S (the step size) apart. It can be fitted into a seasonal ARIMA model as

$$\nabla^d \nabla_S^D x_t = \frac{\Theta(B)\Theta_S(B)}{\Phi(B)\Phi_S(B)} \varepsilon_t \quad (2)$$

where $\{\varepsilon_t\}$ is the zero-mean white noise series and D is the seasonal difference order and S is the step size.

3. The Chinese-new-year-aware ARIMA model

Historical data show that there is a low spell of power consumption during the CNY period in China. We aim to develop a modified ARIMA model which can predict the power consumption more accurately, as explained below.

A closer look at the raw power consumption data shows that the low spell usually occurs one month after the Chinese new year. We denote this month as *isCNY* and calculate the average of the power consumptions of the two months before and after this *isCNY* month. We then use this average (denoted as $AVG_{prev+next}$) to replace the actual power consumption of this *isCNY* month in order to smooth the time series of a calendar year.

A coefficient α is then introduced to eliminate the impact of this CNY period in the refined model.

$$\alpha = (APC - AVG_{prev+next}) / AVG_{prev+next} \quad (3)$$

where APC is the actual power consumption of the $isCNY$ month.

Consider the fact that the impact of CNY period has been eliminated from the new time series given in the previous subsection, we must adjust it in the following manner.

- mark the predecessor and successor month of this $isCNY$ month as pre and suc , and use the new time series to predict the power consumptions of pre and suc respectively, and denote them as F_{pre} and F_{suc}
- calculate the average of F_{pre} and F_{suc} , and denote it as $average(pre, suc)$, and calculate the average as for all the CNY periods given in the time series.
- the final prediction for any given month is as follows:

$$\begin{aligned} Forecast_{month} &= Forecast_{ARIMA} \\ &\text{if } month \neq isCNY, \text{ or} \\ &= Forecast_{ARIMA} + \\ &\quad average(pre, suc) * \alpha_{avg} \\ &\text{if otherwise} \end{aligned} \quad (4)$$

4. Experimental results and analysis

We used the data of overall power consumption in various cities between 2005 and 2011 to evaluate the accuracy of the new forecast method proposed above. The following subsection A explains the experiment and results of one city in Guangzhou, which is followed by a summary of other similar results for other cities. For conciseness, we leave detailed explanation of the results in our technical reports.

We first denoted the CNY month for each year, i.e. March of 2005 and February of 2006, etc. as suggested in the previous section. We then calculated the CNY coefficient α_{2005} , α_{2006} , ... and α_{2009} . After replacing the real power consumption of the $isCNY$ month with $AVG_{prev+next}$, we obtained a new time series of power consumption is illustrated in Figure.1(left). We found that, once we replaced the real power consumption of the $isCNY$ month with $AVG_{prev+next}$, the impact has been eliminated from the time series, i.e. the sharp escalation in the curve has been removed, which resulted in a smooth curve during the corresponding festival periods. The resulting time series is still unstable although its periodicity is obvious, with annual increments, as illustrated. It can then be stabilized by applying differentiation with a lag of 12, and the resulting time series was also given in Figure.1 (right).

Auto-correlation analysis was then applied to this new time series and the resulting autocorrelogram and partial autocorrelogram are given in Figure.2 respectively. The results of white noise detection are $P=0.0424$ when lag is 6, and $P=0.0156$ when lag is 15, both below the pre-defined significance level of 0.05. This suggests that this series after differentiation did not satisfy the pure randomness prerequisite.

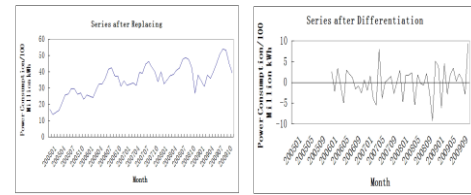


Fig.1 The time series after replacing the real power consumption of the $isCNY$ month with $AVG_{prev+next}$ (left) and the series after differentiation (right)

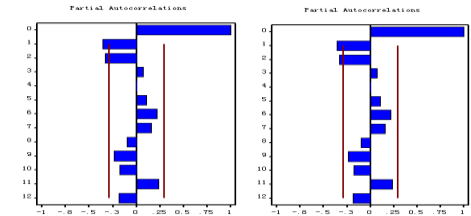


Fig.2 The autocorrelogram (left) and partial autocorrelogram (right)

In order to forecast the power consumption of 2010 whose Chinese new year is February and the $isCNY$ month is March, we must first calculate both the $average(pre, suc)$, i.e. $average(2, 4)$ and the $\alpha_{avg} = (\alpha_{2005} + \alpha_{2006} + \dots + \alpha_{2009}) / 5$, before we have the forecasts of this CNY period in Table.I.

Table.1 error rate of the proposed forecast approach

Month	January 2010	February 2010	March 2010
Real power consumption (kWh)	4023802791	4014180568	2951428008
Forecast(kWh)	4562036086	4269023188	3223675232
error rate	0.133762	0.063486	0.092243

To illustrate the effectiveness of our proposed CNY aware forecast approach, we compared its predictions against those made by the traditional ARIMA model which does not consider the CNY seasonal effect. The results are given in Figure.3, which suggests that the error rate falls below 0.1 for March 2010 whose power consumption was affected by the CNY period most significantly. This suggested that the traditional ARIMA model has been outperformed by

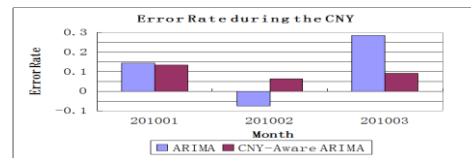


Fig.3 Comparison of error rates of the ARIMA model and the Chinese new year aware ARIMA model.

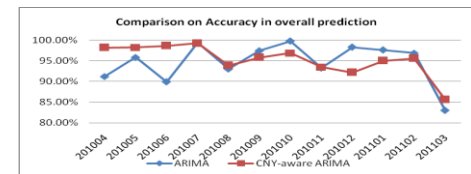


Fig.4 Comparisons on the overall power consumption forecasts made by ARIMA and CNY-aware ARIMA between 04/2010 and 03/2011

Our CNY-aware model in power consumption forecast.

Similar results are given in Figure.4 below. The results show that the proposed CNY-Aware ARIMA forecast outperformed the original ARIMA one, except in 201007, 201108, 201012 and 201103 cases where the differences are relatively modest.

5. Conclusions

We propose in this paper a tailor-made ARIMA model which can effectively deal with the seasonal effect of time series analysis. The experimental results show that this Chinese-new-year-aware variant significantly outperforms the traditional ARIMA model in power consumption forecast.

References

- [1] Energy Saving Forum, Stanford University, Electric load forecast, probing the issues with models, Energy Saving Forum report 3, vol.1, Apr 1979.
- [2] J.D. Hamilton, Time Series Analysis, Princeton University Press, 1994.
- [3] Y. Huang and Y. Lu, "Electricity consumption forecasting based on wavelet analysis and grey model", Power System and Clean Energy, Vol.27, No.3, March 2011. (in Chinese).
- [4] A. Lendasse, J. Lee, V.Wertz and M.Verleysen, "Forecasting electricity consumption using nonlinear projection and self-organizing maps", Neurocomputing, Vol.48, 2002.
- [5] Z. Mohamed and P.Bodger, "Forecasting electricity consumption: a comparison of models for New Zealand", Electricity Engineers' Association of New Zealand Annual Conference, June 2004. (references).
- [6] F.Nogales, J. Contreras, A.Conejo and R. Espinola, "Forecasting next day electricity prices by time series models". IEEE Trans. On Power System, Vol.17, No.2, 2002.
- [7] D. Shao, H. Lin, Z. Hao, H. Zhong and Z. Tao, "Forecating method of medium-term electricity consumption under the impact from the spring festival", Electric Machines and Control, Vol.11, No.5, Sep 2007. (in Chinese).
- [8] S. Wang and J. Jiang, "Seasonal ARIMA models based confidence intervals forecast model for monthly electricity consumption", Guangdong Power Transmission Technology, No.5, 2010. (in Chinese).