Comparing multiple practical models in electricity demand forecasting.

Introduction

Electricity demand forecasting is essential to the comprehensive utilization and dispatch of the power system, so it does have crucial meaning to obtain a high-precision forecasting model of electricity demand. It is generally believed that electricity consumption in a region is inevitably influenced by many random factors such as population density, migration, electricity unit prices, and holiday policies. Therefore, the time series of electricity consumption has highly non-linear characteristics, such as seasonal and non-linear growth trends(Zhu et al., 2011). Although Taylor's research has shown that weather forecasting is an important input for many electric power forecasting models (James W. Taylor & Buizza, 2003), his following paper stated that multivariate modeling is usually not been considered in the real-time forecasting model, "because the weather variables tend to change in a smooth fashion, which will be captured in the demand series itself"(J W Taylor, 2003).

In this paper, I use the Household Electric Power Consumption data set published by the University of California, Irvine (UCI). The format of the data set is shown as follows

Table 1 Household Electric Power Consumption Dataset

DATE	TIME	GLOBAL_ACTIVE_POWER	GLOBAL_INTENSITY	SUB_METERING_1	SUB_METERING_2	SUB_METERING_3
16/12/2006	17:24:00	4.216	18.4	0	1	17
16/12/2006	17:25:00	5.36	23	0	1	16
16/12/2006	17:26:00	5.374	23	0	2	17
16/12/2006	17:27:00	5.388	23	0	1	17
16/12/2006	17:28:00	3.666	15.8	0	1	17
16/12/2006	17:29:00	3.52	15	0	2	17
16/12/2006	17:30:00	3.702	15.8	0	1	17
16/12/2006	17:31:00	3.7	15.8	0	1	17
16/12/2006	17:32:00	3.668	15.8	0	1	17

This archive contains 2075259 measurements gathered between December 2006 and November 2010 (47 months). The information of each column is as follows, *globalactivepower*: household global minute-averaged active power (in kilowatt)； *global_intensity*: household global minute-averaged current intensity (in ampere)； *submetering1*: energy sub-metering No. 1 (in watt-hour of active energy). It corresponds to the kitchen, containing mainly a dishwasher, an oven and a microwave (hot plates are not electric but gas powered); *submetering2*: energy sub-metering No. 2 (in watt-hour of active energy). It corresponds to the laundry room, containing a washing-machine, a tumble-drier, a refrigerator and a light; *submetering3*: energy sub-metering No. 3 (in watt-hour of active energy). It corresponds to an electric water-heater and an air-conditioner.

Since this dataset records electricity consumption hourly, I group them into the daily dataset and Figure 1 shows the daily global power demand from December 2006 to November 2010. A within-year seasonal cycle is apparent from the similarity of the demand profile from one year to the next. Therefore, using a simulation model that is capable to capture the seasonality is appealing in this pre-investigation.

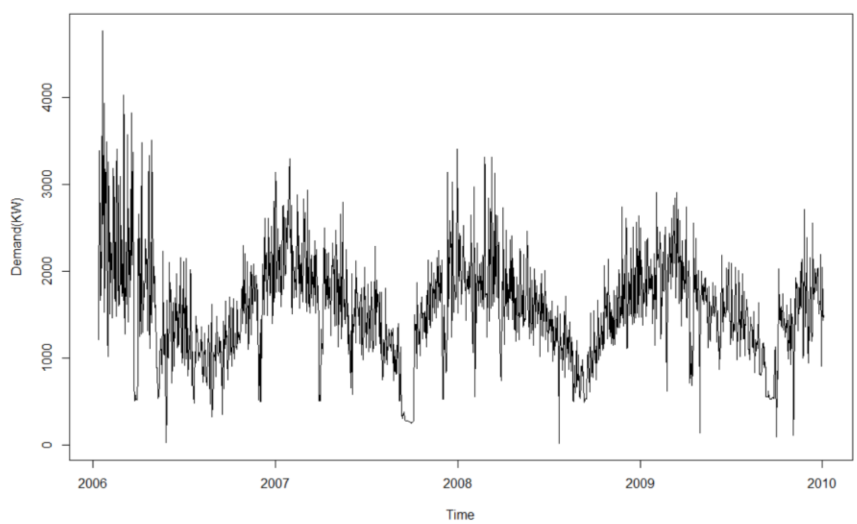


Figure 1 Daily electricity demand

Data Investigation

Constant mean testing

Firstly, I applied linear modeling to verify if the electricity demand has a constant expectation across time or not. The results show the independent variable are significantly correlated with time ($\beta_1 = -0.165_{(0.038)}$), which indicates the non-stationarity. Figure 2 shows the autoregressive correlation and partial autoregressive correlation of residuals after removing the time trend. We could find the ACF are persistent until lag 30, which implies the autoregressive process.

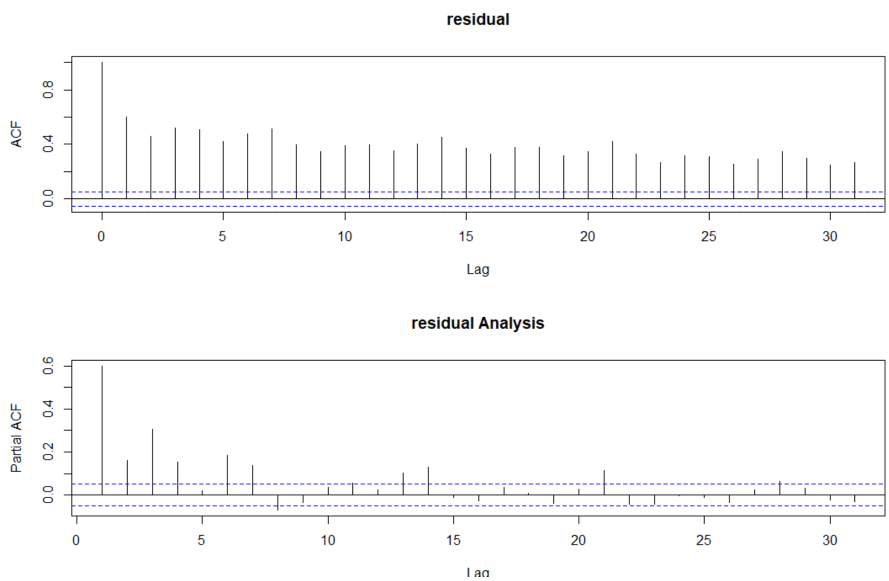


Figure 2 ACF and PACF of residuals after removing time trend

Furthermore, a change-point model is performed for detecting regions with significant variation of the expectation. Figure 3 shows the results from this model, which indicates 13 segment partitions are optimal. Since a total of 13 seasons are covered from December 2006 to November 2010, these results also fit our pre-assumption, the seasonal process is robust.

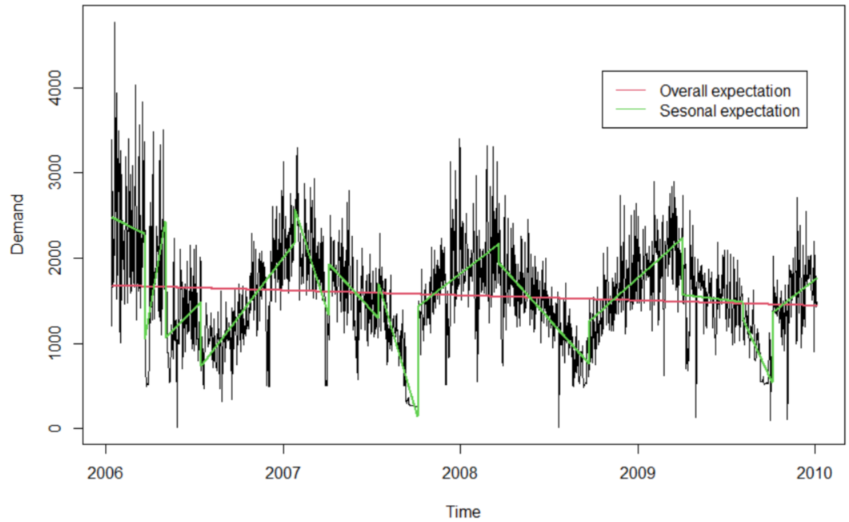


Figure 3 Changepoint detection of the electricity demand

Heterogeneity testing

After testing the constness of the expectation, I also want to find out whether the volatilities are varying across time. A fractionally integrated ARMA model is employed to check whether the electricity demand data are a long memory process. According to the change point analysis, I divided electricity demand data into 5 parts based on different years, including 2006 to 2007, 2007 to 2008, 2008 to 2009,2009 to 2010 respectively. And then, ARFIMA models with $p = 2$ and $q = 2$ are performed for each section, a total of five fractional different estimators are obtained: $d_{2006-2007} = 0.35713_{(0.02167)}$, $d_{2007-2008} = 0.48165_{(0.01840)}$, $d_{2008-2009} = 0.47749_{(0.02055)}$, $d_{2009-2010} = 0.46826_{(0.02148)}$, $d_{2010-} = 0.37709_{0.02554}$. All 5 estimators are statistically significant. Among them, the electricity demand records from 2007 to 2010 are complete, and the difference estimators d of those three years are close to each other, which demonstrates there is no apparent heterogeneity in the data. Figure 4 shows a comparison between the observed demands and the prediction from the ARFIMA model, with $Mean\ absolute\ percentage\ error_{(MAPE)} = 3.250001$, $Mean\ absolute\ error_{(MAE)} = 343.1269$, and $Root\ mean\ squared\ error_{(RMSE)} = 470.0043$.

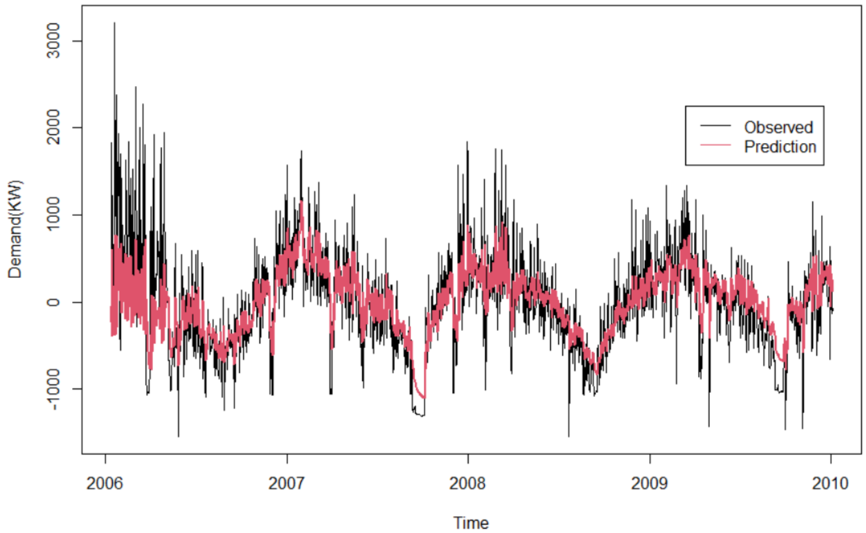


Figure 4 Comparison between observation and prediction (ARFIMA)

Modeling and forecasting

Precision indexes

The mean absolute error (MAE), root mean square error (RMSE) and mean absolute percentage error (MAPE) are three intuitive forecasting indexes, they have been widely utilized for evaluating the efficiency of forecasting models. In this paper, I also applied them as my measuring metrics. Assume we have the observed data y_t and prediction data $\hat{y_t}$ which given by our model, the definitions of MAE, RMSE and MAPE are respectively shown as follows:

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|y_t - \hat{y_t}|$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n}(y_t - \hat{y_t})^2}{n}}$$

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{|y_t - \hat{y_t}|}{n} * 100\%$$

SARMA and SARIMAX model

Since the seasonal component within our is robust, seasonal autoregressive moving average model (SARMA) and its integrated form become my first candidates. For removing the time trend within power consumption data, I added time data as the input variable for the SARIMA model and applied first-order differencing for the SARIMAX model. Figure 5 shows the ACF and PACF for the dependent variable after the first-order differencing.

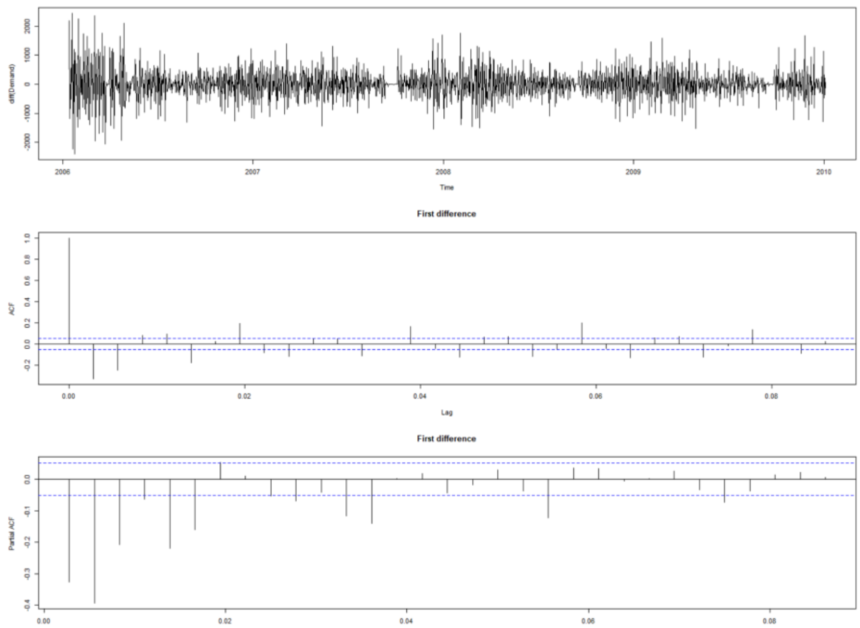


Figure 5 ACF and PACF for demands after first-order differencing

Based on the AIC and Log-likelihood criteria, I finally selected SARMAX model with $p = 4, q = 5$, and SARIMA model with $p = 3, d = 1, q = 3$. All results indicate the seasonal estimators are not significant. For the SARMAX model, $RMSE = 428.9939, MAE = 313.8097$, $MAPE = 0.36658$. And for the SARIMA model, $RMSE = 430.6007$, $MAE = 314.6841$, $MAPE = 0.361529$. Figure 6 shows the diagnostic for two models' residual. From the Ljung-Box statistic, there is no short-term serial correlation within residuals.

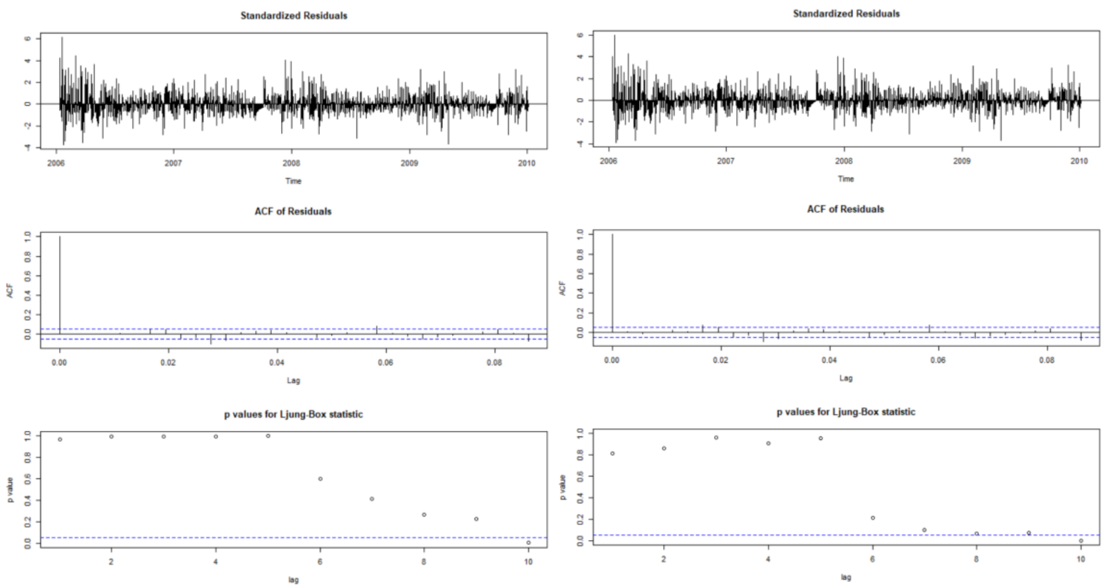


Figure 6 Comparison between observation and prediction (ARFIMA)

Dynamic Linear Regression model

Based on the results from the ARFIMA model, demands data has a long-term memory process. Therefore, using dynamic linear modeling with the higher-order lags as the input variables may also give us a good prediction. After removing non-significant input variables, the final model is $y_t = \beta_1 y_{t-1} + \beta_2 y_{t-3} + \beta_3 y_{t-4} + \beta_4 y_{t-6} + \beta_5 y_{t-7} + \beta_6 y_{t-8} + \beta_7 y_{t-9} + \beta_8 y_{t-13} + \beta_9 y_{t-14} + \beta_{10} y_{t-21} + w_t$., where w_t is white noise. Figure 7 shows the ACF and PACF of residuals from this model, which proves there is no obvious serial correlation within the residuals. And for this model, $RMSE = 410.331$, $MAE = 306.8899$, $MAPE = 0.3554943$.

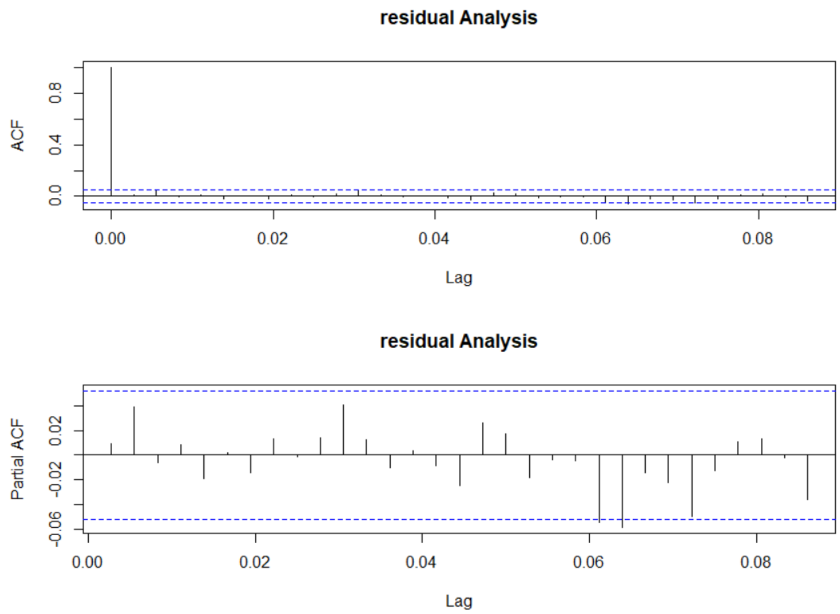


Figure 7 ACF and PACF of the residuals (dynamic linear model)

Holt-winters

Another practical approach for time series modeling named Holt-Winters, which also uses historical values to make the prediction. Compared with the normal moving average process, Holt-winters smoothing time series data using the exponential window function, which also very similar to the kernel moving average process. After performing this model, I get the smoothing parameter $\alpha = 0.2167$, $\beta = 0.0001$. BY checking the ACF and PACF of residuals from this model, we could find there still have some strong serial correlation(as shown in Figure 8). And for this model, I get the $MAPE = 0.3622604$, $RMSE = 456.119$, $MAE = 333.3419$.

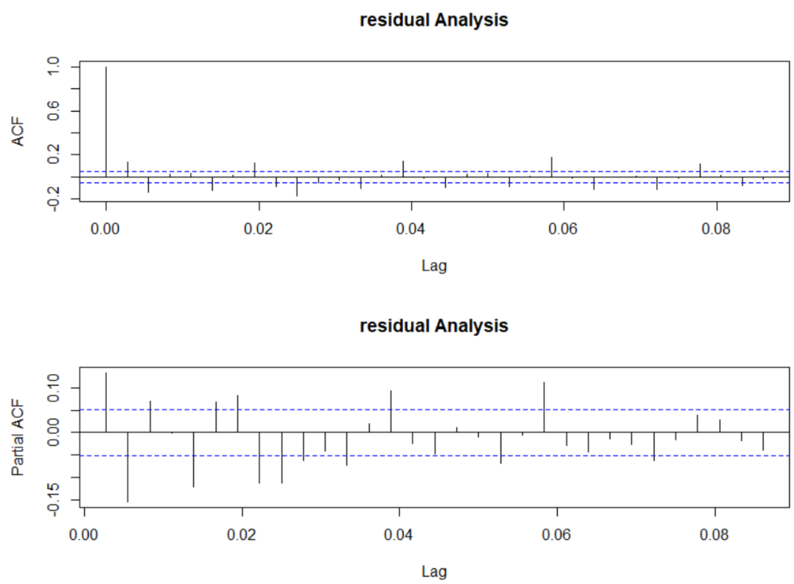
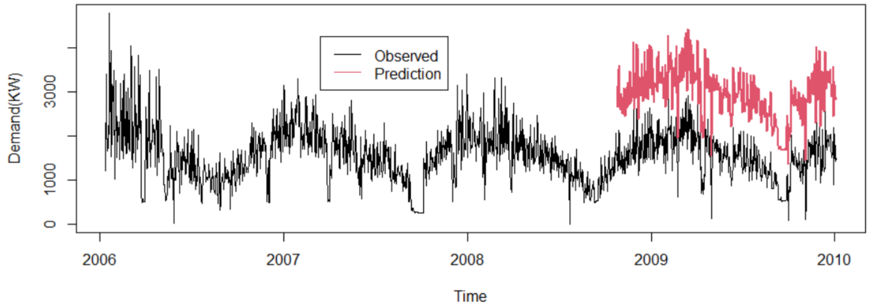


Figure 7 ACF and PACF of the residuals (Exponential smoothing)

Long short-term memory modeling (LSTM)

Long short-term memory modeling is a time series forecasting approach from the deep learning domain. Basically, it could dynamically justify the weight of the historical data using iteration, and a very long memory process could be applied (long lags input are possible). Figure 8 shows the comparison between the observation and prediction from my LSTM model. As shown below, there are obvious offsets between those two datasets. It may cause by the normalization approach(all inputs in the LSTM model are required to be scaled). If I only compare the variation from their mean, the results become much better. The other drawback of the LSTM model is we need to use enough records as the training dataset. Therefore, each LSTM model could only generate 30% predictions. Although we could use k-fold cross-validation to get the completed prediction. Since this approach is computational tremendous, I do not perform it here.

Under the original scale, for LSTM model, $MAPE = 1.047166$, $MAE = 1377.535$, $RMSE = 1381.806$. If we only compare the variation around mean, $MAE = 0.9263611$, $RMSE = 108.5662$, $MAE = 82.67138$.



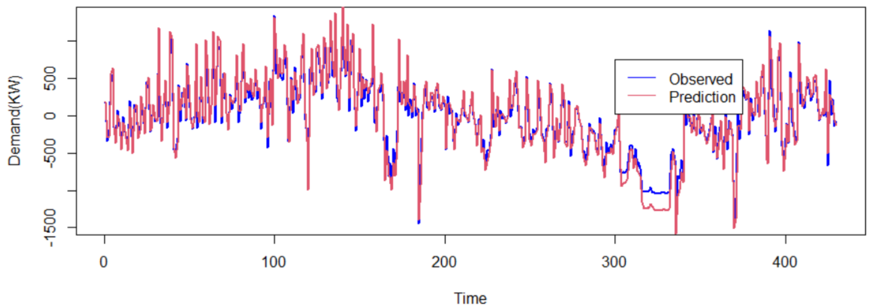


Figure 8 Original scale comparison (above), offsets from the mean comparison (below)

Conclusion

I have investigated 6 types of commonly utilized time series modeling on electricity demands. Table2 demonstrates the value of the metric from each model respectively.

Table 2 comparison of 6 models

Model Name	RMSE	MAE	MAPE
Dynamic Linear model	410.331	306.8899	0.355494
SARMAX	428.9939	313.8097	0.36658
SARIMA	430.6007	314.6841	0.361529
ARFIMA	470.0043	343.1269	0.387754
Exponential Smoothing	456.119	333.3419	0.36226
LSTM model (origin)	1381.806	1377.535	1.047166
LSTM model (offset)	108.5662	82.67138	0.926361

According to the results of this research, the dynamic linear model has a good performance for modeling the long memory process from the electricity demands. However, it is kind of hard to find the logical relationship between power demands and 10 input lags. On the other hand, The LSTM model is remarkable for modeling offsets. It could be utilized as a complementary tool for other models.

Reference

Taylor, J W. (2003). Short-term electricity demand forecasting using double seasonal exponential smoothing. *Journal of the Operational Research Society*, *54*(8), 799–805. https://doi.org/10.1057/palgrave.jors.2601589

Taylor, James W., & Buizza, R. (2003). Using weather ensemble predictions in electricity demand forecasting. *International Journal of Forecasting*, *19*(1), 57–70. https://doi.org/10.1016/S0169-2070(01)00123-6

Zhu, S., Wang, J., Zhao, W., & Wang, J. (2011). A seasonal hybrid procedure for electricity demand forecasting in China. *Applied Energy*, *88*(11), 3807–3815. https://doi.org/10.1016/j.apenergy.2011.05.005