Time Series Analysis and Forecasting 2DD23

Xuqiang Fang x.fang@student.tue.nl
Hadi Sotudeh h.sotudeh@student.tue.nl

***Abstract*—Time series is a series data points with time order indexes. Researches have been focused on analyzing time series and building models in order to extract meaningful characteristics from it. Time series forecasting is to use models to predict future values according to previously observed values. This report consists of time series analysis and forecasting, using different models such as Holt-Winter's model and Box-Jenkins model, based on the same criterion, the best model is selected and used to forecast future values. It shows that even if the model can fit the time series well, there are still some incidents which can affect the precision of the forecast by the model.**

I. Introduction

Part I is an univariate time series analysis, starting with exploratory data analysis (EDA), both in time domain and frequency domain, based on the EDA we identified the possible trend and seasonality, then we performed Holt-Winter's exponential smoothing, automatic exponential smoothing as well as Box-Jenkins model to fit the time series. Using the criteria AICc (SSE for Holt-Winter's) we selected the best model, and then we forecasted the future with the model.

Part II is the analysis of multivariate time series. We also started with EDA, for the energy per capita time series, we identified that there was trend but no seasonal patterns, and we continued with frequency domain analysis. Then we tried to fit an ARIMA model for the energy per capita. For the dynamic regression, we tried to use one of the time series to forecast another. Finally we fit an exponential smoothing model to the energy per capita time series.

Verification and validation are useful when building models. We performed both verification and validation on all models and we forecast future with the model that performed the best.

II. Part I

Starting with the exploratory data analysis both in the time domain and the frequency domain, we first transformed the dataset into time series, then we plotted it in time domain, as shown in figure 1, we found out that there were seasonal patterns and missing values (from March 1995 to September 1996), and there were maybe trends within the time series. Each year, the level goes up first and then goes down, and that is the seasonal pattern.

To see if the successive observations are correlated, we plotted the autocorrelation plot at time lag 1, and calculated the coefficient. We obtained $R = 0.734108908$, which shows that at time lag 1, the successive observations are positively correlated as shown in figure 2.

A. Exploratory Data Analysis

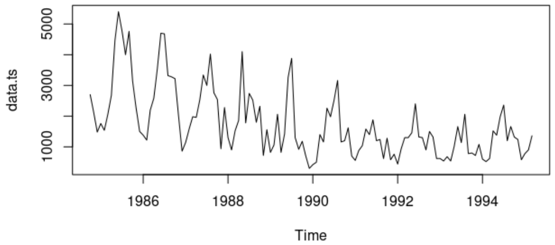


Figure 1. Time Series Plot

Then we continued with the analysis both in time domain and frequency domain.

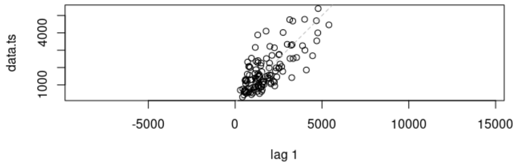


Figure 2. Autocorrelation at time lag 1

In the time domain, as shown in figure 3, we can identify seasonality in ACF plot because after each 12 months, the pattern is repeated (but it is not clear whether or not it is an additive seasonality). For the trend, we concluded that there was a weak decreasing trend if any trend existed, because we did not see a lot of high values for the ACF plot. Also from the PACF plot, we can see that successive observations are positively correlated. Also from the log transformed time series plot, we found out an irregular fluctuations around year 1990.
In the frequency domain, we started with the raw periodogram and then followed by the smoothed periodogram. From the raw periodogram we can see that there are multiple peaks within the periodogram but there seems no dominant frequencies within the periodogram.

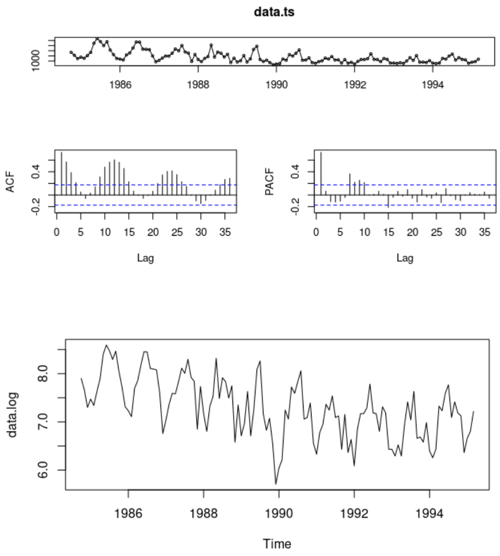


Figure 3. Time Domain: tsdisplay and log transformed data

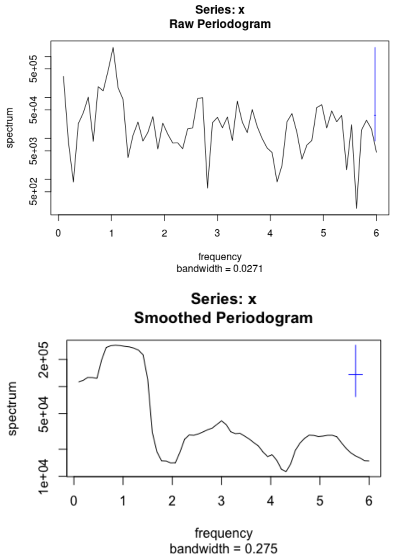


Figure 4. Frequency Domain: Raw and smoothed periodogram

Then we continued with the smoothed periodogram and we found that there is a dominant frequency around 1, which is the monthly record of the data, which suggests that there are seasonal patterns within the time series because of the the low frequency part in the spectrum which is high.

B. Seasonal Decomposition

Based on Hyndman's book, 'If the data have a strong seasonal pattern, we recommend that seasonal differencing be done first because sometimes the resulting series will be stationary and there will be no need for a further first difference.' [2].

By looking at the time series, we found out there was a yearly seasonal pattern, within each year, the level goes up in January and then goes down after June, it reaches a low point in December. The level goes up quite smoothly with barely some small vibrations. Since a seasonal pattern was detected, we proceeded with seasonal decomposition of the time series. First we performed one time seasonal differencing, and we obtained the results as shown in figure 5. From the ACF and PACF plots, we identified that there were still some significant autocorrelations within the plots, so we continued with another seasonal differencing and we found out that the significant patterns still existed and were about the same. So we figured out that one time seasonal decomposition was enough.

To decide if the seasonal pattern is additive or multiplicative, we looked at the seasonal variation to see if it was constant over time or not. And we found out that the seasonal variation was proportional to the trend, as shown in figure 1, so we concluded that the seasonal pattern was **multiplicative**.

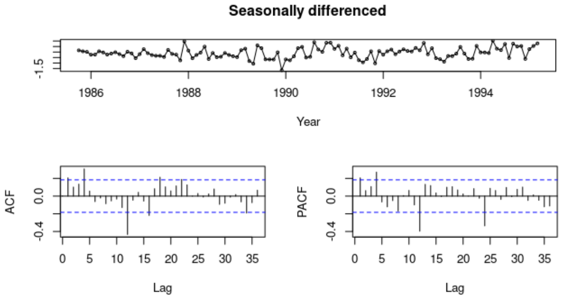


Figure 5. Time Domain: One step seasonal differencing

C. Exponential Smoothing Models

In this part, we used both Holt-Winter and Automatic Exponential Smoothing models to fit the time series, we performed both verification and validation on the models. The best model is used to compare with the best ARIMA model and we selected the better one to forecast.

1) Holt-Winter's Exponential Smoothing: There are three types of exponential smoothing models in general and they are applied to different occasions. The simple exponential smoothing is used for time series without trend and seasonality, Holts' exponential smoothing model is used for time series with trend and Holt-Winter's exponential smoothing is used for time series with both trend and seasonality. Since we found that there were seasonality within the time series and possible trend within it, we would choose to use Holt-Winter's exponential smoothing.

Table I
TABLE: HOLT-WINTER'S MODEL PARAMETERS

a 1229.9643432	b 13.9814393	s1 0.9426255	s2 1.3041867	s3 1.6756330
s4 1.5390151	s5 1.3767354	s6 1.0922448	s7 1.0534098	s8 0.7964503
s9 0.6024360	s10 0.4964780	s11 0.5579722	s12 0.8657817	

As concluded above, the seasonal pattern was 'multiplicative', so we used Holt-Winter's exponential smoothing with

’multiplicative’ mode of seasonality, and we obtained results as shown in figure 6. The level, trend and seasonal parameters are shown in table I.

And the smoothing parameters are: $\alpha = 0.006805899$, $\beta = 1$, and $\gamma = 0.1324159$. As we can see that alpha is close to zero, which means that the level update depends on much more on the past instead of recent, so the level line is smooth. But beta=1 means that the trend update depends just on its last value, so the trend line is not smooth but it vibrates, $\gamma = 0.1324159$ shows that seasonal update depends on both past and recent seasonal levels, but it depends on more on the past.

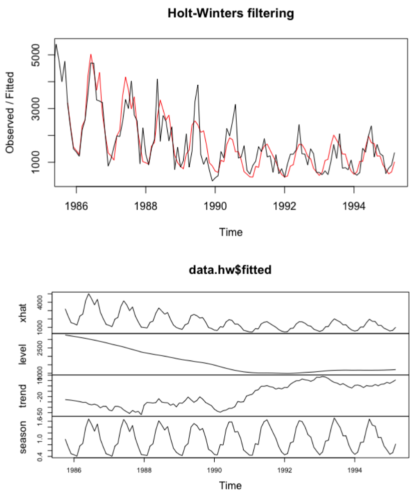


Figure 6. Time Domain: Holt-Winter’s multiplicative model

To evaluate the model we obtained using Holt-Winters exponential smoothing method, we first performed verification of the model and then the validation of the model. To verify the model, we looked at the in-sample accuracy and the in-sample forecast errors. The in-sample accuracy is shown in table II, note that RMSE=552.40. The in-sample forecast errors is shown in figure 7.

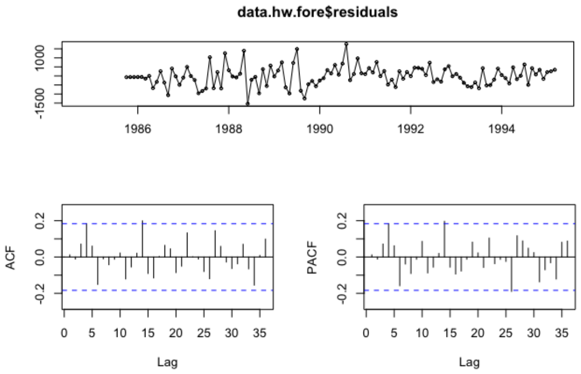


Figure 7. Holt-Winter’s In-sample Forecast Residuals

The ACF and PACF plots both showed that there were still some patterns within the residuals, it was not a white noise. To see if those autocorrelations are significant or not, we performed Box-Ljung test, and from the ACF plot, there is a spike at lag=4, so we performed the test setting lag=4.

Table II
IN-SAMPLE ACCURACY: HOLT-WINTER’S

ME	RMSE	MAE	MPE	MAPE	ACF1	Theil’s U
-6.455	552.40	419.11	-8.015	31.05	0.0116	0.8196

Table III
BOX-LJUNG TEST

Box-Ljung test
data: data.hw.fore$residuals
X-squared = 4.6836, df = 4, p-value = 0.3213

The test results suggest not reject null hypothesis, meaning that the autocorrelation is not significant. Then we performed normality test of the residuals. The Shapiro-Wilk normality test suggested that not reject normality. So the residuals could be considered as white noise.

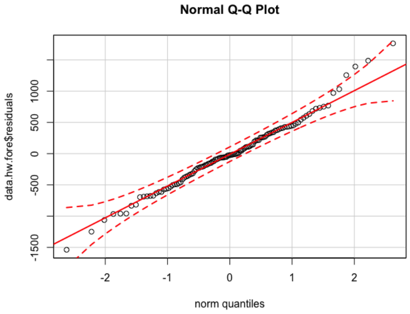


Figure 8. Shapiro-Wilk Normality Test for Holt-Winter’s Forecast Residuals

Shapiro-Wilk normality test
data: data.hw.fore$residuals
W = 0.98584, p-value = 0.2759

To validate the model, we use the last two years (19%) data, we fit the model using data from March 1993 to March 1995. For the Holt-Winter’s exponential smoothing model, the **RMSE** is 641.1747.

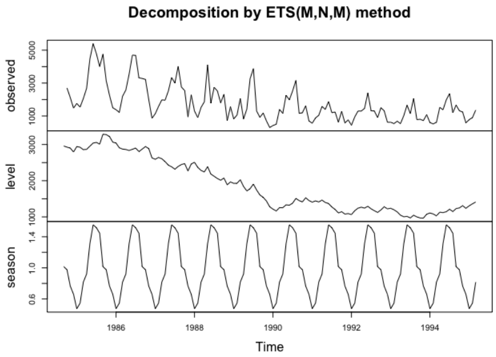


Figure 9. ETS Decomposition

2) Automatic Exponential Smoothing: As another part of exponential smoothing method, we also applied automatic exponential smoothing model to compare with the Holt-Winter's model we obtained in previous part.

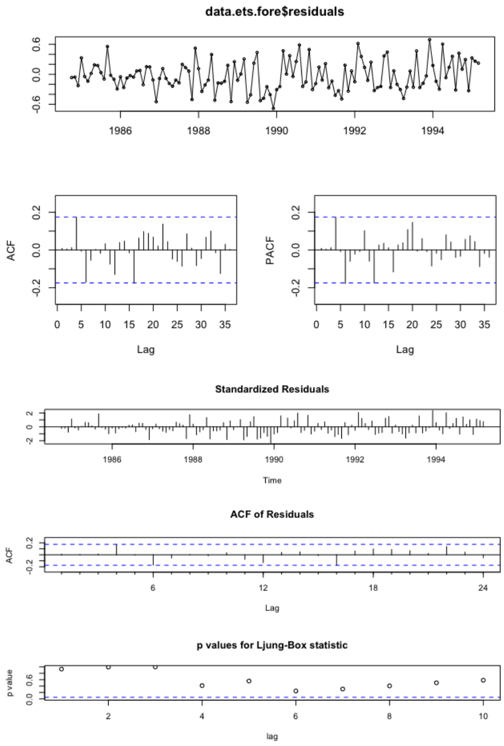


Figure 10. ETS Model Residuals and Box-Ljung test

The ETS suggests a model of ETS(M,N,M), which is multiplicative error, no trend and multiplicative seasonal pattern, and the seasonality is 12. The estimated parameters are: $\alpha = 0.1633, \gamma = 1e - 04$.For the initial states are: $l = 2959.0364$ $l = 1.0155, , 1.4449, 1.5166, 1.5531, 1.3099, 0.9209, 0.8169, 0.5462, 0.4739, 0.66, 0.7661, 0.9761$, $\sigma = 0.2986$

We also performed verification for the ETS model.

From the ACF plot we did not see any significant autocorrelations. And the Ljung-Box test statistics also showed that all autocorrelations were not significant. The normality test for the residuals also suggests not reject null hypothesis, so the residuals could be considered as white noise.

Box-Ljung test
data: data.ets.fore$residuals
X-squared = 3.9444, df = 4, p-value = 0.4136

Shapiro-Wilk normality test
data: data.ets.fore$residuals
W = 0.98269, p-value = 0.1078

Table IV
IN-SAMPLE ACCURACY:ETS

ME	RMSE	MAE	MPE	MAPE	ACF1	Theil's U
-67.92	536.14	417.80	-14.42	30.634	0.6881	-0.0232

As shown in table IV, the in-sample accuracy shows the ETS model has RMSE of 536.14, which is better than the in-sample accuracy for Holt-Winter's model. To validate the ETS model, we also used the last two years' data. The **RMSE** for the ETS model is 368.6765.

3) Forecast: Through verification and validation of the two exponential models, we find that ETS model performed better based on both in in-sample accuracy RMSE and validation RMSE. so we used the ETS model to forecast the future, the grey region in figure 11 is the confidence interval, the blue line is the forecast levels, the level still keeps the seasonal pattern as it goes up first and then goes down within each year, and the confidence interval reaches the biggest at peak, the exact forecast values can be found in appendix (see figure 26). The exact formula for the final model is as follow, the forecast values are generated according to these formulas.

$$y_t = l_{t-1}s_{t-12}(1 + \epsilon_t)$$

$$l_t = l_{t-1}(1 + 0.1633\epsilon_t)$$

$$s_t = s_{t-12}(1 + 0.0001\epsilon_t)$$

where ϵ_t is the white noise with standard deviation of $\sqrt{0.2986} = 0.5464$. As we can see from the formula, the seasonal pattern does not change much, the level changes, since 0.1633 is quite small, we can see the level does not vibrate a lot.

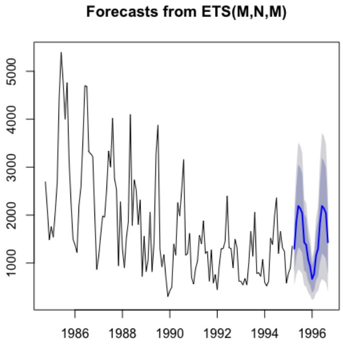


Figure 11. ETS Model Forecasting

D. Box-Jenkins Model

In this section, we used ARIMA models to fit the time series and we selected the best model according to criterion AICc. After comparing the best ARIMA model to the model we obtained in previous section, we chose the better one to forecast. We followed the following flowchart from Hyndman's book[1] (figure 12) to build a Box-Jenkins model for the time series.

The first three steps are same as the previous section in exponential smoothing models, so we started from the fourth step. In the fourth step, we estimated the ARIMA model parameters and also came up with several candidates for it. In the plots of the seasonally differenced data, there are spikes in the PACF at lags 12 and 24, and a seasonal lag in the ACF at lag 12. This may suggest a seasonal AR(2) term. In the non-seasonal lags, there are two significant spikes in the PACF suggesting a possible AR(4) or AR(1) term. The pattern in the ACF is not indicative of any simple model. Consequently, this initial analysis suggests that a possible model for these data is an $\mathrm{ARIMA}(4, 0, 0)(2, 1, 0)_{12}$ or $\mathrm{ARIMA}(1, 0, 0)(2, 1, 0)_{12}$.

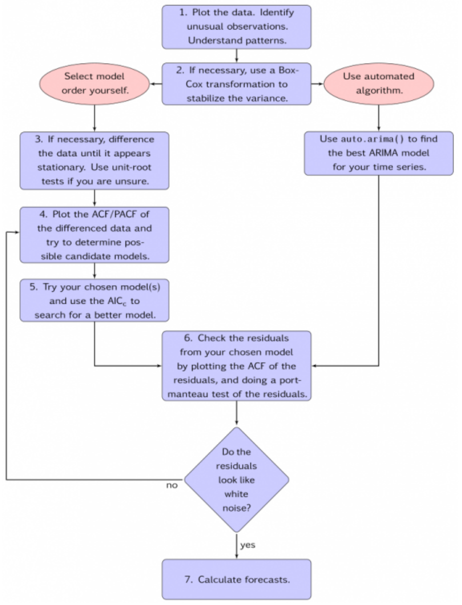

1. Plot the data. Identify unusual observations. Understand patterns.
2. If necessary, use a Box-Cox transformation to stabilize the variance.
Select model order yourself.
Use automated algorithm.
3. If necessary, difference the data until it appears stationary. Use unit-root tests if you are unsure.
Use auto.arima() to find the best ARIMA model for your time series.
4. Plot the ACF/PACF of the differenced data and try to determine possible candidate models.
5. Try your chosen model(s) and use the AICc to search for a better model.
6. Check the residuals from your chosen model by plotting the ACF of the residuals, and doing a portmanteau test of the residuals.
Do the residuals look like white noise?
no
yes
7. Calculate forecasts.

Figure 12. Box-Jenkins Model: General process for forecasting using an ARIMA model

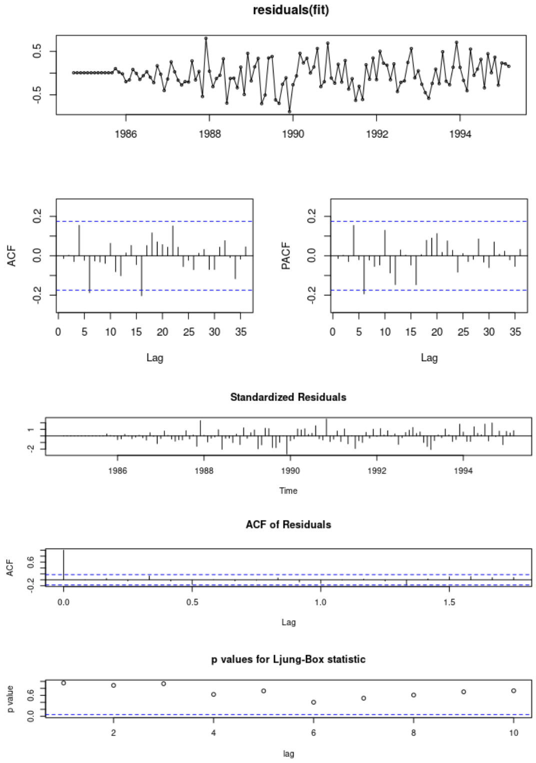


Figure 13. Box-Jenkins Residuals Plot

We fitted these models along with some variations on them and computed their AICc values to verify our models which were shown in the following table V.

From the table V, we selected $\mathrm{ARIMA}(1, 0, 1)(2, 1, 0)_{12}$ which has the smallest AICc, then we look at its variants in the seasonal part, which is shown in the following table VI.

Among all these models, the best model is $\mathrm{ARIMA}(1, 0, 1)(0, 1, 1)_{12}$ because it has the lowest value of AICc, which is equal to 107.61. Having obtained the model, we then looked at the residuals plot, as shown in figure 13.

Table V
CANDIDATE ARIMA MODELS AND AICC

Model	AICc
$\mathrm{ARIMA}(4, 0, 0)(2, 1, 0)_{12}$	123.73
$\mathrm{ARIMA}(4, 0, 1)(2, 1, 0)_{12}$	126
$\mathrm{ARIMA}(4, 0, 2)(2, 1, 0)_{12}$	128.11
$\mathrm{ARIMA}(2, 0, 3)(2, 1, 0)_{12}$	126.76
$\mathrm{ARIMA}(1, 0, 3)(2, 1, 0)_{12}$	125.07
$\mathrm{ARIMA}(1, 0, 2)(2, 1, 0)_{12}$	123.16
$\mathrm{ARIMA}(1, 0, 1)(2, 1, 0)_{12}$	120.93
$\mathrm{ARIMA}(2, 0, 2)(2, 1, 0)_{12}$	125.05
$\mathrm{ARIMA}(3, 0, 2)(2, 1, 0)_{12}$	128.16
$\mathrm{ARIMA}(3, 0, 0)(2, 1, 0)_{12}$	129.03
$\mathrm{ARIMA}(2, 0, 0)(2, 1, 0)_{12}$	129.53

Table VI
ARIMA (1,0,0) WITH VARIANTS IN SEASONAL PART

ARIMA Model (1,0,1)	AICc
$\mathrm{ARIMA}(1, 0, 1)(1, 1, 0)_{12}$	134.56
$\mathrm{ARIMA}(1, 0, 1)(0, 1, 1)_{12}$	107.61
$\mathrm{ARIMA}(1, 0, 1)(1, 1, 2)_{12}$	111.81
$\mathrm{ARIMA}(1, 0, 1)(1, 1, 1)_{12}$	109.61

As it is shown, there are significant spikes in both the ACF and PACF, but the model doesn't fail a Ljung-Box test which means the residuals are not correlated. There are other candidates (variants of the first estimation) which we didn't look at their AICc, so we tried running auto.arima() with differencing specified to be d=0 (trend degree of differencing) and D=1 (seasonality degree of differencing), and allowing larger models than usual. This led to an $\mathrm{ARIMA}(1, 0, 1)(2, 1, 0)_{12}$ model (note that this model has the lowest AICc among all models with D=1 and d=0 which are invertiable too). All of its coefficients are shown in table VII, as shown, all coefficients are significant and the obtained model did pass all the tests and residuals are like a white noise.

Table VII
AUTO ARIMA COEFFICIENTS

Coefficients	ar1	ma1	sar1	sar2
Value	0.9636	-0.7863	-0.6839	-0.3734
s.e.	0.0376	0.1022	0.0906	0.0869

In addition, we tried using the automatic ARIMA algorithm. Running auto.arima() with arguments left at their default values led to an $\mathrm{ARIMA}(0, 1, 5)(2, 0, 0)_{12}$ model. However, residuals of this model are not like white noise and its ACF and PACF have different spikes. All in all, its AICc=135.51, which is very high compared to the previous models we studied, so we ignored this model.

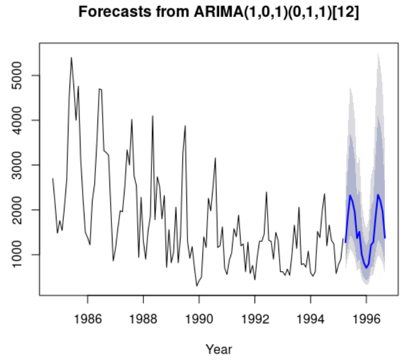


Figure 14. Box-Jenkins Forecasting

Then we performed validation for the Box-Jenkins models, we still used the test set consisting of the last two years (19%) of data. Actually we performed validation for all the models listed in table V and VI, we found out that the **RMSE** for all models were basically the same, which means all of them are performing the same as other candidates. Based on AICc (Verification part) and RMSE (Validation part), ARIMA$(1,0,1)(0,1,1)_{12}$ is selected. In practice, we would normally use the best model we could find, even if it did not pass all tests.

Then we used the model ARIMA$(1,0,1)(0,1,1)_{12}$ and its coefficients (which all are significant) (Table VIII) to forecast the next 18 months, which is shown in figure 14. The exact forecast values can be found in appendix. (see figure 27). The forecasted values have the seasonal pattern with a very weak increasing trend. Their prediction intervals are small at the begining, but they increase a lot at peaks, then they decrease which is obvious in both the figure 14 and in appendix (figure 27).

Table VIII
ARIMA$(1,0,1)(0,1,1)_{12}$ COEFFICIENTS

Coefficients	ar1	ma1	sma1
Value	0.9985	-0.7929	-1.0000
s.e.	0.0153	0.0635	0.1615

The specific model written in formula is as follow:

$$(1-0.9985B)(1-B^{12})x_t = (1-0.7929B)(1-B^{12})e_t$$

$$(1-B^{12}-0.9985B+0.9985B^{13})x_t = (1-B^{12}-0.7929B+0.7929B^{13})e_t$$

$$x_t = 0.9985x_{t-1}+x_{t-12}-0.9985x_{t-13}+ e_t-0.7929e_{t-1}-e_{t-12}+0.7929e_{t-13}$$

where ϵ_t is the white noise with standard deviation of $\sqrt{0.114} = 0.33763886032$.

This formula has 7 parameters and means that forecasted variable at time t depends on variable's real values and white noise at time t-1 (previous month), t-12 (previous year, the same month), and t-13 (previous year, one month before this month).

E. Comparison between Exponential Smoothing and Box-Jenkins models

We used the two years' data (19%) for validation for both exponential smoothing models and ARIMA models. Then we chose the same criterion to see which one performed the best. As shown in table IX and figure 15, ETS model performed the best.

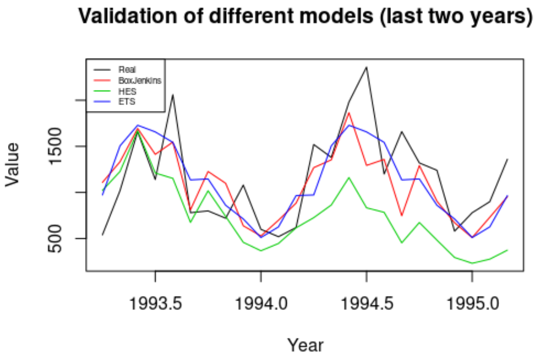


Figure 15. Validation Plot for Different Models

Table IX
VALIDATION FOR EXPONENTIAL SMOOTHING MODELS AND ARIMA MODELS

metric	Box-Jenkins	Holt Winters	ETS
RMSE	1257.881	641.1747	368.6765

III. PART II

A. Exploratory Data Analysis

In this section we looked at the fingerprints of the time series. We analyzed the time series both in the time domain and the frequency domain.

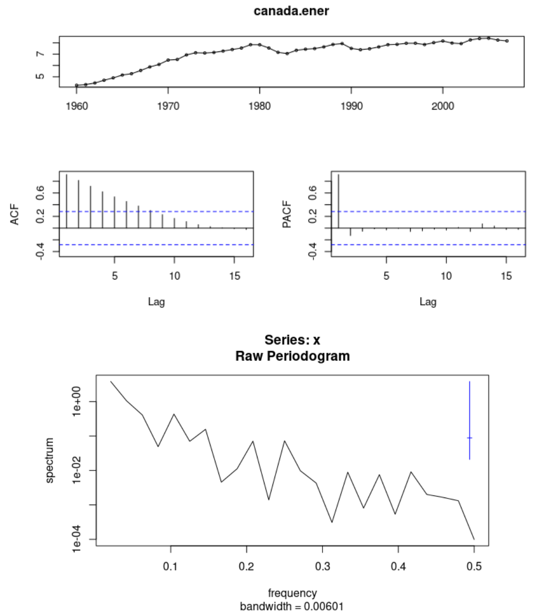


Figure 16. Time Series Plot (energy per capita)

Time Domain (energy per capita): The energy per capita plot shows a clearly increasing trend, but we are not sure if there is any seasonal pattern, then we looked at the tsdisplay of the capita series. The ACF plot shows a clear sign of increasing trend, but it does not suggest any seasonal pattern either. Frequency domain (energy per capita): The raw periodogram shows that the dominant frequencies are in the low frequency zone, mainly smaller than 0.1, which indicates a large period. So for this time series, there is obviously a trend but there is no obvious seasonal pattern. And since there is no high frequency spikes, the time series is quite smooth.

Time domain (gdp): The gdp plot also shows an increasing trend, and it shows no obvious seasonal pattern. For the ACF and PACF plots, ACF plot has many consecutive significant correlations, which suggests the existence of an increasing trend, and it does not show any sign of a seasonal pattern.

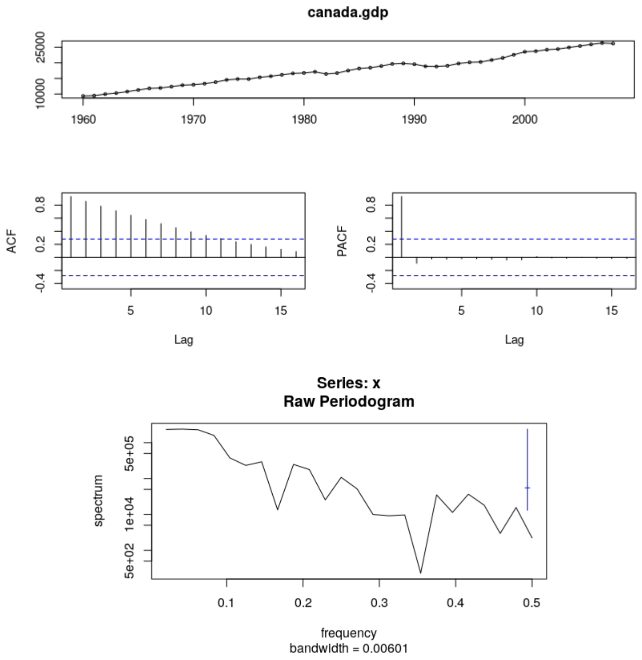


Figure 17. Time Series Plot (gdp)

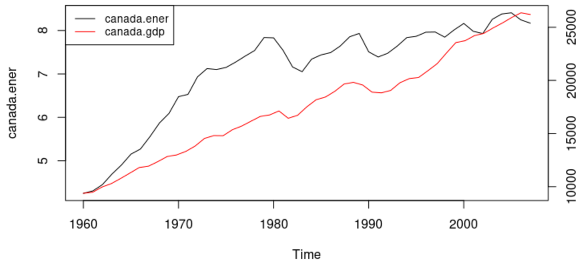


Figure 18. Time Series Plot (energy per capita & gdp)

Frequency domain(gdp): The raw periodogram suggests the dominant frequencies located in the low frequency zone, which indicates a very large period for the time series. So, as for the time series available, we can ignore the effect of seasonal pattern. Also because of the peaks in the low frequency zone, it shows that the time series has an increasing trend.

The time series plot for both energy per capita and gdp are shown in figure 18, it shows that both have increasing trend, this indicates a possible correlation between the two time series.

B. *Univariate Box-Jenkins Model(energy per capita)*

For the Box-Jenkins model, we still followed the flowchart in figure 12. The fingerprint analysis we concluded that the energy per capita time series had an increasing trend but no obvious seasonal pattern. Also since there was not much variance so we did not transform the time series. Also the time series kept increasing until 1980, after 1980, there was a slight decrease. There was no unusual changes in the time plot. (As shown in figure 18)

Since the time series is not stationary, we started with one-step finite differencing. After the one-step finite differencing, we can see that the time series seems to be stationary. So for the ARIMA model, we choose d=1. Based on the ACF and PACF plots, there are spikes both at time lag 1, so we figure that the candidate for p and q should be p=1 and q=1. Since we conclude that there is no seasonal pattern, so the proposed model is ARIMA$(1, 1, 1)$, we also considered all other combinations of p and q and we compared all the proposed models in order to find the best one.

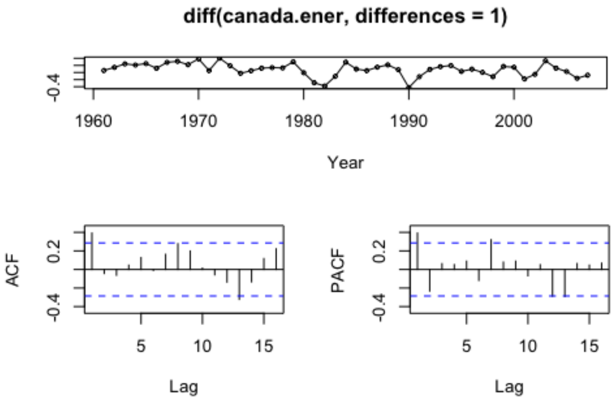


Figure 19. tsdisplay after one-time finite differencing(energy per capita)

Table X
UNIVARIATE ARIMA MODELS

Model	AICc
Arima(1,1,1)	-26.45
Arima(1,1,0)	-27.43
Arima(1,1,2)	-26.68
Arima(0,1,1)	-28.33
Arima(0,1,0)	-16.1
Arima(0,1,2)	-26.3
Arima(2,1,0)	-25.99
Arima(2,1,1)	-25.99
Arima(2,1,2)	-24.76

According to the AICc, ARIMA(0,1,1) gives the best model. Also by looking at the ACF and PACF plots, we could see that none of the correlations are significant. Looking at the residuals, we used qqplot as shown in figure 21 to check the residuals and performed Shapiro-Wilk normality test, test results suggest not reject null hypothesis, so the residuals indeed can be considered as white noise.

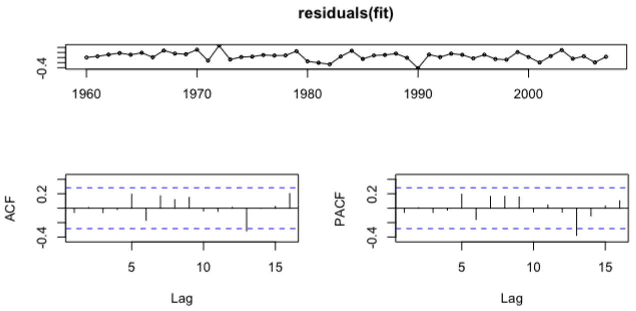


Figure 20. ARIMA: tsdisplay for the residuals (energy per capita)

Shapiro-Wilk normality test
data: fit.arima$residuals
W = 0.98058, p-value = 0.6029

Table XI
ARIMA MODEL COEFFICIENTS

Coefficients	ma1	σ^2
Values	0.5660	0.02965
s.e.	0.1357	

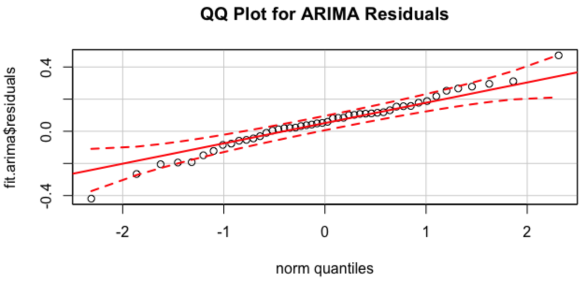


Figure 21. ARIMA: QQ plot for the residuals

The specific formula for the ARIMA (0,1,1) is as follow:

$$y_t = y_{t-1} + e_t + 0.566e_{t-1}$$

For the non-seasonal ARIMA model, the auto regressive part is 0 and there is a one-time finite differencing, the e_t is the white noise with the standard deviation of $\sqrt{0.02965} = 0.17219$. As we can see from the formula, the update level depends on previous value and the values of white noises at time t and time t-1 with corresponding coefficients of 1 and 0.566.

Table XII
IN-SAMPLE ACCURACY: ARIMA

ME 0.05235639	RMSE 0.1685781	MAE 0.1318204	MPE 0.8583297	MAPE 1.909194
MASE 0.7940359	ACF1 -0.05877094			

We also performed validation for the model. We found the real data for year 2008 to 2013. [3] And we chose the criterion **RMSE**, for the ARIMA model we obtained, the RMSE=0.3676267.

We also applied the auto.arima to select the best model automatically, and the best model suggested by auto.arima was also ARIMA(0,1,1). (See the code)

C. Dynamic Regression

We chose GDP value of Canada as the predictor variable. We took the following steps:

- We applied differencing until all variables (forecast and predictor) are stationary. One time differencing was enough to have stationary variables.
- We fitted the regression model with AR(2) errors.
- We calculated the errors (nt) from the fitted regression model and identified an appropriate ARMA model for them.
- We re-fitted the entire model using the new ARMA model for the errors.
- Finally, we checked that the et series looks like white noise.

Possible candidate ARIMA models include an MA(5) and AR(1). However, further exploration reveals that an ARIMA(1,1,0) has the lowest AICc value among possible candidates. We refit the model with ARIMA(1,1,0) errors to obtain the results in table XIII.

Table XIII
ARIMA MODEL COEFFICIENTS

Coefficients	ar1	canada.gdp
Values	0.3108	2e-04
s.e.	0.1485	2e-04

The exact mathematical formula is:

$$y_{t'} = 0.0002x_t^{'} + n_t^{'}$$

$$n_t^{'} = 0.3108n_{t-1}^{'} + e_t$$

we know that $y_{t'} = y_t - y_{t-1}$, $x_{t'} = x_t - x_{t-1}$, and $n_{t'} = n_t - n_{t-1}$. We just replace them in the formula and we will have:

$$y_t - y_{t-1} = 0.0002(x_t - x_{t-1}) + 0.3108(n_t - n_{t-1}) + e_t$$

The final expanded formula will be:

$$y_t = y_{t-1} + 0.0002x_t - 0.0002x_{t-1} + 0.3108n_t - 0.3108n_{t-1} + e_t$$

and

$$n_t - n_{t-1} = 0.3108(n_{t-1} - n_{t-2}) + e_t$$

which is equal to:

$$n_t = 1.3108n_{t-1} - 0.3108n_{t-2} + e_t$$

where ϵ_t is the white noise with standard deviation of $\sqrt{0.02464} = 0.15697133496$.

This formula has 6 parameters, this means that forecasted variable at time t depends on real value of forcast variable at time t-1 and value of predictor variables at time t and t-1 and white noise at time t and t-1.

Using this model to forecast, the results is shown in figure 22.

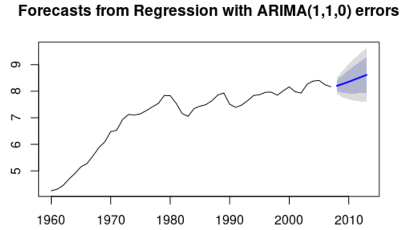


Figure 22. Forecasts from Regression with ARIMA(1,1,0)

Then we looked at the lagged version and followed the steps mentioned at section 9.1 in Hyndman's book [2]. We got the results shown in table XIV.

Table XIV
LAGGED VERSION ARIMA

Coefficients	ar1	Adlag0	Adlag1	Adlag2
Values	0.3592	2e-04	1e-04	-1e-04
s.e.	0.1786	2e-04	2e-04	2e-04

The chosen model includes GDP only the last three months and has AR(1) errors. The model can be written as:

$$y_t^{'} = 0.0002x_t^{'} + 0.0001x_{t-1}^{'} - 0.0001x_{t-2}^{'} + n_t^{'}$$

$$n_t^{'} = 0.3592n_{t-1}^{'} + e_t$$

Notice that $y_t^{'} = y_t - y_{t-1}$ and so on. So the expanded version of the formula is as following:

$$y_t - y_{t-1} = 0.0002(x_t - x_{t-1}) + 0.0001(x_{t-1} - x_{t-2}) - 0.0001(x_{t-2} - x_{t-3}) + (n_t - n_{t-1})$$

which is equal to:

$$y_t = y_{t-1} + 0.0002x_t - 0.0001x_{t-1} - 0.0002x_{t-2} - 0.0001x_{t-3} + n_t - n_{t-1}$$

and

$$n_t - n_{t-1} = 0.3592(n_{t-1} - n_{t-2}) + e_t$$

which is equal to:

$$n_t = 1.3592n_{t-1} - 0.3592n_{t-2} + e_t$$

where ϵ_t is the white noise with standard deviation of $\sqrt{0.02318} = 0.15224979474$.

This formula means that forecasted variable at time t depends on real value of forcast variable at time t-1 and value of predictor variables at time t, t-1, t-2, and t-3. In addition, it depends on white noise at time t and t-1.

Using this model to forecast, the results is shown in figure 23.

For the forecasting part, first we forecasted the next 8 years of the predictor variable using ets command in R, then we used the result to forecast the forecast variable using dynamic regression.

Validation and verification of these two models are shown in table XV

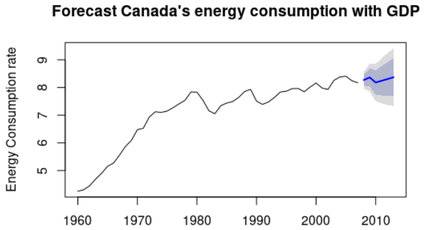


Figure 23. Forecasts Canada's Energy with GDP

Table XV
VERIFICATION AND VALIDATION OF DYNAMIC REGRESSION MODELS

Models	Dynamic Regression (Simple)	Dynamic Regression (Lagged)
Verification	-36.06 (AICc)	-32.14 (AICc)
Validation	0.613 (RMSE)	0.472 (RMSE)

Based on the outputs, there is no single model which outperforms in both verification and validation parts, but we choose the lagged version as our final model here because it perfroms better in the validation part.

D. Exponential Smoothing models

We used non-seasonal exponential smoothing model for energy indicator of Canada. There are two ways to do this in R, we could choose either HoltWinters command or ets one. We did both and chose the best one. First, we ran with HoltWinters with gamma=FALSE because the time series doesn't contain seasonality. Other parameters are obtained using optimizing automatically. We obtained $\alpha = 1, \beta = 0, \gamma = FALSE$

For level, it only takes the recent value. For trend, it only takes the difference of levels into account.

Table XVI
IN-SAMPLE ACCURACY: HOLT-WINTER'S ES

ME	RMSE	MAE	MPE	MAPE	ACF1	Theil's U
0.02761	0.1851	0.1467	0.5440	2.113	0.3937	0.8849

Next we fit the time series with ETS model. We obtained $\alpha = 0.9999, \beta = 0.1217$. ETS gave us a non-seasonal model with additive trend and multiplicative error. Its alpha is close to 1 which means the last level and last trend have more weights when it wants to predict next level and trend.

Table XVII
IN-SAMPLE ACCURACY: ETS

ME	RMSE	MAE	MPE	MAPE	ACF1	Theil's U
-0.027	0.1761	0.1327	-0.359	1.8535	0.7993	0.2757

Based on RMSE metrics, ETS performs better, so we choose ETS one and forecast next 6 years, as shown in figure 24.

The exact mathematical formulas for ETS model are:

$$y_t = (L_{t-1} + T_{t-1})(1 + \epsilon_t)$$

$$L_t = (L_{t-1} + T_{t-1})(1 + 0.9999\epsilon_t)$$

$$T_t = T_{t-1} + 0.1217(L_{t-1} + T_{t-1})\epsilon_t$$

where ϵ_t is the white noise with standard deviation of $\sqrt{0.0235} = 0.15329709716$.

Here, we have a Holts linear method or ETS(M,A,N) with multiplicative errors which means a multiplicative error, an additive trend, and no seasonality. The forecasted variable is sum of the last trend and level multiplied by $1+\epsilon_t$. The level depends on the previous level and trend multiplied by $1 + 0.9999\epsilon_t$. The trend depends on the previous level and trend and error value at time t.

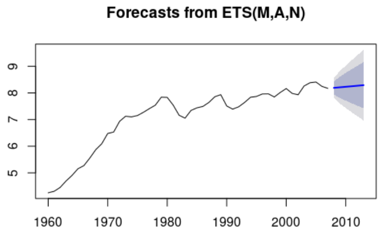


Figure 24. Forecasts Using ETS

E. Comparison between the models

We used the real values from 2008 to 2013 to validate all models [2]. Based on RMSE, the Box-Jenkins model performed the best. However when we plotted the forecast by all models with the real values in the same figure (figure 25), we found out that all models did not work, all predictions are basically increasing while the real value is decreasing.

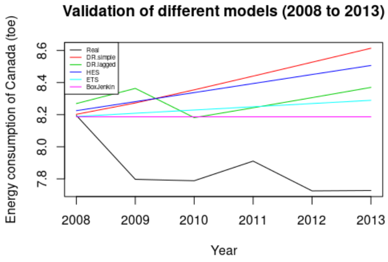


Figure 25. Validation plot of all models

metric	Box-Jenkins	DR(Simple)	DR(Lagged)	ES(HW)	ETS
RMSE	0.367	0.613	0.472	0.563	0.426

We tried to think about the reason why the forecast did not work and we thought about the global financial crisis. Obviously, due to unusual financial crisis in 2008, energy per capita and gdp both decreased while the predictions can not capture this.

So as we can see, even if we tried to build the best model possible, still there are some incidents that may affect the real time series data, and we can not build perfect model using different time series models.

IV. DISCUSSION

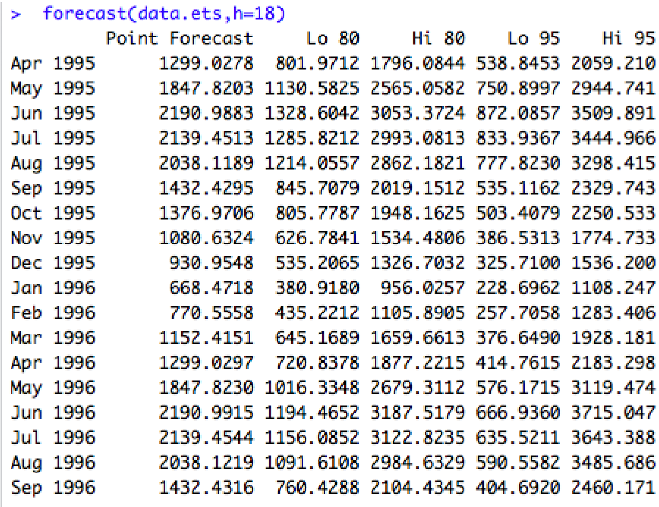
Time series analysis is technically flexible in the sense that there are several methods to build a model for a time series. Among all these models, we do follow some 'principles', for example, we always start analysis with the exploratory data analysis of the time series, choosing the appropriate parameters according to ACF and PACF, verification and validation are both necessary when building the models.

After writing this report, we found out that there was no golden hammer for all datasets. In addition, the tools available sometimes do not work as the way we thought they would be. For example, in the energy consumption part, we tried different models but the error was not acceptable and always very big. We also found out that even if we found the best model possible, sometimes it was still very difficult for the predictions to comply with the real data, simply because there are some patterns or incidents which the models can not capture.

REFERENCES

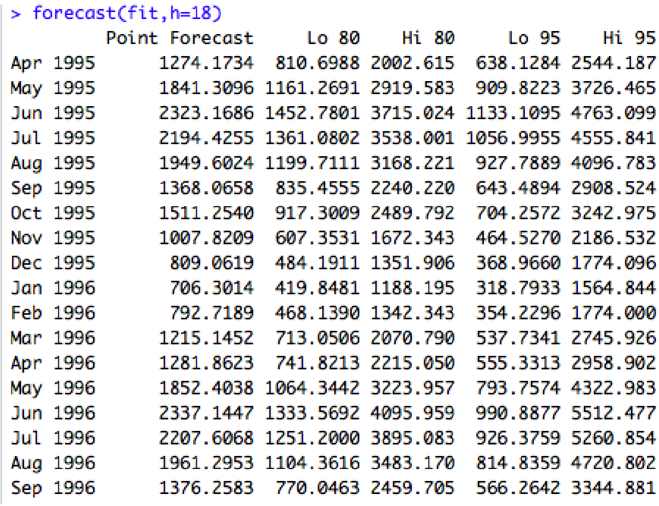
[1] https://en.wikipedia.org/wiki/Time_series
[2] Hyndman, Rob J., and George Athanasopoulos. Forecasting: principles and practice. OTexts, 2014.
[3] http://data.worldbank.org/indicator/EG.USE.PCAP.KG.OE?locations=CA,accessedJuly26,2017.

APPENDIX



```
>  forecast(data.ets,h=18)
         Point Forecast       Lo 80     Hi 80    Lo 95    Hi 95
Apr 1995      1299.0278  801.9712 1796.0844 538.8453 2059.210
May 1995      1847.8203 1130.5825 2565.0582 750.8997 2944.741
Jun 1995      2190.9883 1328.6042 3053.3724 872.0857 3509.891
Jul 1995      2139.4513 1285.8212 2993.0813 833.9367 3444.966
Aug 1995      2038.1189 1214.0557 2862.1821 777.8230 3298.415
Sep 1995      1432.4295  845.7079 2019.1512 535.1162 2329.743
Oct 1995      1376.9706  805.7787 1948.1625 503.4079 2250.533
Nov 1995      1080.6324  626.7841 1534.4806 386.5313 1774.733
Dec 1995       930.9548  535.2065 1326.7032 325.7100 1536.200
Jan 1996       668.4718  380.9180  956.0257 228.6962 1108.247
Feb 1996       770.5558  435.2212 1105.8905 257.7058 1283.406
Mar 1996      1152.4151  645.1689 1659.6613 376.6490 1928.181
Apr 1996      1299.0297  720.8378 1877.2215 414.7615 2183.298
May 1996      1847.8230 1016.3348 2679.3112 576.1715 3119.474
Jun 1996      2190.9915 1194.4652 3187.5179 666.9360 3715.047
Jul 1996      2139.4544 1156.0852 3122.8235 635.5211 3643.388
Aug 1996      2038.1219 1091.6108 2984.6329 590.5582 3485.686
Sep 1996      1432.4316  760.4288 2104.4345 404.6920 2460.171
```

Figure 26. Forecast Using ETS model



```
> forecast(fit,h=18)
         Point Forecast       Lo 80    Hi 80     Lo 95    Hi 95
Apr 1995      1274.1734  810.6988 2002.615  638.1284 2544.187
May 1995      1841.3096 1161.2691 2919.583  909.8223 3726.465
Jun 1995      2323.1686 1452.7801 3715.024 1133.1095 4763.099
Jul 1995      2194.4255 1361.0802 3538.001 1056.9955 4555.841
Aug 1995      1949.6024 1199.7111 3168.221  927.7889 4096.783
Sep 1995      1368.0658  835.4555 2240.220  643.4894 2908.524
Oct 1995      1511.2540  917.3009 2489.792  704.2572 3242.975
Nov 1995      1007.8209  607.3531 1672.343  464.5270 2186.532
Dec 1995       809.0619  484.1911 1351.906  368.9660 1774.096
Jan 1996       706.3014  419.8481 1188.195  318.7933 1564.844
Feb 1996       792.7189  468.1390 1342.343  354.2296 1774.000
Mar 1996      1215.1452  713.0506 2070.790  537.7341 2745.926
Apr 1996      1281.8623  741.8213 2215.050  555.3313 2958.902
May 1996      1852.4038 1064.3442 3223.957  793.7574 4322.983
Jun 1996      2337.1447 1333.5692 4095.959  990.8877 5512.477
Jul 1996      2207.6068 1251.2000 3895.083  926.3759 5260.854
Aug 1996      1961.2953 1104.3616 3483.170  814.8359 4720.802
Sep 1996      1376.2583  770.0463 2459.705  566.2642 3344.881
```

Figure 27. Forecast Using Box-Jenkins model

```
PART I
########################
library('fpp')
library('ggplot2')
library('forecast')
library(car)

##### importing and transforming the dataset
data.ts <- ts(group04_data_1$month04, start=c(1984,10),
end = c(1995,3),frequency=12)

######## exploratory data analysis part
###time domain
plot(data.ts)

tsdisplay(data.ts)

seasonplot(data.ts, season.labels = TRUE, year.labels = TRUE, col=rainbow(12))

monthplot(data.ts)

data.log <- log(data.ts)
plot(data.log)

######### seasonal decomposition
data.sd1 <- diff(data.ts, lag=12, differences = 1)
plot(data.sd1)

data.sd2 <- diff(data.ts, lag=12, differences = 2)
tsdisplay(data.sd2)
plot(data.sd2)

data.deco <- decompose(data.ts,type = 'additive')
plot(data.deco)

data.deco.mul <- decompose(data.ts,type = 'multiplicative')
plot(data.deco.mul)


###frequency domain
spectrum(data.ts)
spectrum(data.ts, span=5)
#spectrum(data.sd1)
#spectrum(data.sd2)


######### exponential smoothing model
#Holt-Winter's exponential smoothing
data.hw <- HoltWinters(data.ts,seasonal = 'multiplicative')
plot(data.hw)

#model parameter
data.hw$alpha
data.hw$beta
data.hw$gamma
data.hw$coefficients

data.hw$fitted
plot(data.hw$fitted)
data.hw$SSE
with(data.hw,accuracy(fitted,x))



#diagnostics
data.hw.fore <- forecast(data.hw)
tsdisplay(data.hw.fore$residuals)
#from the ACF plot, it looks like the lag is 4.
Box.test(data.hw.fore$residuals,lag = 4, type = 'Ljung-Box')
qqPlot(data.hw.fore$residuals, main = 'Normal␣Q-Q␣Plot')
shapiro.test(data.hw.fore$residuals)


#validation
exvalidate <- function(x, h,...)
{
  train.end <- time(x)[length(x)-h]
  test.start <- time(x)[length(x)-h+1]
  train <- window(x,end=train.end)
  test <- window(x,start=test.start)
  fit <- HoltWinters(train,...)
  fc <- forecast.HoltWinters(fit,h)
  return(accuracy(fc,test)[2,"RMSE"])
  #return(fc)
}
exvalidate(data.ts, h=24, alpha=0.08586768,beta=0,gamma=0.4719598,
seasonal='multiplicative')

#HES_plot <- exvalidate(data.ts, h=24, alpha=0.08586768,beta=0,
gamma=0.4719598, seasonal='multiplicative')$mean


#automatic exponential smoothing
data.ets <- ets(data.ts)
summary(data.ets)
plot(data.ets)
accuracy(data.ets)
tsdiag(data.ets)
data.ets$aicc
data.ets.fore <- forecast.ets(data.ets,h=18)
#data.ets.fore
#plot(data.ets.fore)
#normality test
tsdisplay(data.ets.fore$residuals)
Box.test(data.ets.fore$residuals, lag = 4, type='Ljung-Box')
qqPlot(data.ets.fore$residuals)
shapiro.test(data.ets.fore$residuals)

with(data.ets,accuracy(fitted,x))
#validation
etsvalidate <- function(x,h)
{
  train.end <- time(x)[length(x)-h]
  test.start <- time(x)[length(x)-h+1]
  train <- window(x,end=train.end)
  test <- window(x,start=test.start)
  fit <- ets(train)
  fc <- forecast.ets(fit,h)
  return(accuracy(fc,test)[2,"RMSE"])
  #return(fc)
}
etsvalidate(data.ts,h=24)

#ets_plot <- etsvalidate(data.ts,h=24)$mean

#forecast
data.ets.fore <- forecast(data.ets,h=18)
plot(data.ets.fore)

#log transform
data.log <- log(data.ts)
plot(data.log)
data.log.hw <- HoltWinters(data.log,seasonal = 'additive')
plot(data.log.hw)
data.hw <- HoltWinters(data.ts,seasonal = 'additive')



######### Box—Jenkins model

plot(data.ts)

tsdisplay(data.ts)

data.log <- log(data.ts)

plot(data.log)

tsdisplay(data.log)

tsdisplay(diff(data.log,12), main="Seasonally␣differenced", xlab="Year")

### choosing the best model AICc
fit <- Arima(data.ts, order=c(1,0,1), seasonal=c(0,1,1),lambda = 0)
qqPlot(fit$residuals)
accuracy(fit)
tsdiag(fit)

tsdisplay(residuals(fit))
Box.test(residuals(fit),lag=36, fitdf = 4,type="Ljung—Box")

#data.d1 <- diff(data.ts, differences = 1)

#data.d1.sd2 <- diff(data.d1, lag=12, differences = 2)


#tsdisplay(data.d1)

#data.d2 <- diff(data.ts, differences = 2)

#data.d2.sd1 <- diff(data.d2, lag=12, differences = 1)

#tsdisplay(data.d2.sd1)

auto.arima(data.ts)

fit <- auto.arima(data.ts, lambda=0, d=0, D=1,ic="aicc" ,
max.order=10, stepwise=FALSE, approximation=FALSE)
tsdisplay(residuals(fit))
Box.test(residuals(fit), lag=36, fitdf=8, type="Ljung")


####################
getrmse <- function(x,h,...)
{
  train.end <- time(x)[length(x)—h]
  test.start <- time(x)[length(x)—h+1]
  train <- window(x,end=train.end)
  test <- window(x,start=test.start)
  fit <- Arima(train,...)
  fc <- forecast(fit,h=h)
  return(accuracy(fc,test)[2,"RMSE"])
  #return(fc)
}

getrmse(data.ts,h=24,order=c(3,0,0),seasonal=c(2,1,0),lambda=0)

#BoxJenkins_plot <- getrmse(data.ts,h=24,order=c(3,0,0),
seasonal=c(2,1,0),lambda=0)$mean

### null—hypothesis = stationary time series
kpss.test(data.ts)

### ndiffs
ns <- nsdiffs(data.ts)
if(ns > 0) {
  xstar <- diff(data.ts,lag=frequency(data.ts),differences=ns)
} else {
  xstar <- data.ts
}
nd <- ndiffs(xstar)
if(nd > 0) {
  xstar <- diff(xstar,differences=nd)
}

plot(xstar)

x<-data.ts
h=24
test.start <- time(x)[length(x)—h+1]
real <- window(x,start=test.start)

#Plot (validation of different models)
ts.plot(real,BoxJenkins_plot,HES_plot,ets_plot,col=1:4,
gpars = list(ylab="Value",xlab="Year"))
#col=c("black","red","green","blue","pink"))
title(main = "Validation␣of␣different␣models␣(last␣two␣years)")

legend("topleft",
       legend=c("Real","BoxJenkins","HES","ETS"),
       col=1:4,
       lty=1,
       cex=0.5)



PART II
########################
library(fpp)
library(ggplot2)
library(forecast)
library(reshape)
library(car)
##Exploratory Analysis

canada.ener <- ts(group04_data_2$ener_canada, start=1960,end=2007)

plot(canada.ener, main='Energy␣per␣Capita␣Plot')

tsdisplay(canada.ener)

spectrum(canada.ener)
spectrum(canada.ener, span=5)
canada.ener.d1 <- diff(canada.ener, differences = 1)
tsdisplay(canada.ener.d1)
plot(canada.ener.d1,main='Finite␣Differencing␣Energy␣per␣Capita')

#canada.ener.d2 <- diff(canada.ener, differences = 2)
#tsdisplay(canada.ener.d2)
#plot(canada.ener.d2,main='Two Steps Finite Differencing Energy per Capita')

plot(cbind(canada.ener,canada.gdp), main='')
########

canada.gdp <- ts(group04_data_2$gdp_canada, start=1960,end=2007)

plot(canada.gdp)

tsdisplay(canada.gdp)

spectrum(canada.gdp)

canada.gdp.d1 <- diff(canada.gdp, differences = 1)
tsdisplay(canada.gdp.d1)
plot(canada.gdp.d1)

#canada.gdp.d2 <- diff(canada.gdp, differences = 2)
#tsdisplay(canada.gdp.d2)
#plot(canada.gdp.d2)

### Plot together
plot(canada.ener, las=0)
par(new=TRUE)
plot(canada.gdp,
     col=2,
     bty='n',
     xaxt="n",
     yaxt="n",
     xlab="", ylab="")

axis(4, las=0)

legend("topleft",
       legend=c("canada.ener","canada.gdp"),
       col=1:2,
       lty=1,
       cex=0.85)
############# Box—Jekins ARIMA model
plot(canada.ener) #variance not much, do not need to transform
tsdisplay(canada.ener)
tsdisplay(diff(canada.ener,differences = 1), xlab="Year")

# choose the best model according to AICc, other models are omitted,
# see the results table in the report
fit.arima <- Arima(canada.ener, order=c(0,1,1), method = 'ML')
summary(fit.arima)

qqPlot(fit.arima$residuals)
shapiro.test(fit.arima$residuals)
accuracy(fit.arima)
tsdisplay(residuals(fit.arima))

fit <- auto.arima(canada.ener,d=1,D=0,ic='aicc')
summary(fit)
tsdisplay(residuals(fit))

canada.ener.boxjenkin.fore <- forecast.Arima(fit.arima,h=6)
canada.ener.boxjenkin.fore$mean



############# Dynamic Regression
#first, we should remove trend and seasonality? d=1

#First check
(fit <- Arima(canada.ener, xreg=canada.gdp,order=c(2,1,0)))
tsdisplay(arima.errors(fit), main="ARIMA␣errors")
tsdisplay(residuals(fit)) #seems like a white noise

##check candidates
(fit_simple <- Arima(canada.ener, xreg=canada.gdp,order=c(1,1,0)))
tsdisplay(residuals(fit_simple)) #seems like a white noise
Box.test(residuals(fit_simple),fitdf=1,lag=5,type="Ljung")
canada.gdp.ets <- ets(canada.gdp)
canada.gdp.ets.fore <- forecast(canada.gdp.ets, h=6)
plot.forecast(canada.gdp.ets.fore)

canada.gdp.next <- canada.gdp.ets.fore$mean

DR.simple.fore <- forecast(fit_simple,xreg=canada.gdp.next, h=6)
plot.forecast(DR.simple.fore)

# Lagged predictors. Test 0, 1, 2 or 3 lags.
Advert <- cbind(canada.gdp,
                c(NA,canada.gdp[1:47]),
                   c(NA,NA, canada.gdp[1:46]),
                   c(NA,NA,NA, canada.gdp[1:45]))
colnames(Advert) <- paste(”AdLag”,0:3,sep=””)

# Choose optimal lag length for advertising based on AIC
# Restrict data so models use same fitting period
fit1 <- auto.arima(canada.ener[4:48], xreg=Advert[4:48,1], d=1)
fit2 <- auto.arima(canada.ener[4:48], xreg=Advert[4:48,1:2], d=1)
fit3 <- auto.arima(canada.ener[4:48], xreg=Advert[4:48,1:3], d=1)
fit4 <- auto.arima(canada.ener[4:48], xreg=Advert[4:48,1:4], d=1)
fit5 <- auto.arima(canada.ener[4:48], xreg=Advert[4:48,2], d=1)
fit6 <- auto.arima(canada.ener[4:48], xreg=Advert[4:48,2:3], d=1)
fit7 <- auto.arima(canada.ener[4:48], xreg=Advert[4:48,2:4], d=1)
fit8 <- auto.arima(canada.ener[4:48], xreg=Advert[4:48,3], d=1)
fit9 <- auto.arima(canada.ener[4:48], xreg=Advert[4:48,3:4], d=1)

# Best model fitted to all data (based on AICc)
# Refit using all data
(fit_lagged <- auto.arima(canada.ener, xreg=Advert[,1:3], d=1))

#forecasting
DR.lagged.fore <- forecast(fit_lagged, xreg=cbind(canada.gdp.next[1:6],
c(Advert[47,1],canada.gdp.next[2:6])
,c(Advert[46,2],Advert[47,2],canada.gdp.next[3:6])), h=6)

plot(DR.lagged.fore, main=”Forecast␣Canada's␣energy␣consumption␣with␣GDP”,
ylab=”Energy␣Consumption␣rate”)

##CCF part
ccf(canada.ener,canada.gdp,na.action = na.omit)
ccf(canada.ener.d1,canada.gdp.d1,na.action = na.omit)


###################################
#Exponential Smoothing
plot(canada.ener)
canada.ener.hes <- HoltWinters(canada.ener, gamma = FALSE)

plot(canada.ener.hes$fitted)
plot(canada.ener.hes)

canada.ener.hes$SSE
with(canada.ener.hes, accuracy(fitted,x))

canada.ener.hes.fore <- forecast.HoltWinters(canada.ener.hes,h=6)
plot.forecast(canada.ener.hes.fore)

ggtsdisplay(canada.ener.hes.fore$residuals)
accuracy(canada.ener.hes.fore)

#automatic exponential smoothing
canada.ener.ets <- ets(canada.ener)
summary(canada.ener.ets)
plot(canada.ener.ets)
accuracy(canada.ener.ets)
tsdiag(canada.ener.ets)
canada.ener.ets.fore <- forecast.ets(canada.ener.ets,h=6)
canada.ener.ets.fore
plot(canada.ener.ets.fore)
#normality test
tsdisplay(canada.ener.ets.fore$residuals)
Box.test(canada.ener.ets.fore$residuals, lag = 4, type='Ljung-Box')
qqPlot(canada.ener.ets.fore$residuals)
shapiro.test(canada.ener.ets.fore$residuals)

###Validation part

RMSE = function(m, o){ #m stands for model and o stands for observation
  sqrt(mean((m-o)^2))
}

canada.ener_future<-read.csv(”Dropbox/TSFproject/canada␣-␣Sheet1.csv”)
canada.ener_future <- canada_Sheet1
RMSE(canada.ener.boxjenkin.fore$mean,canada.ener_future$energy)
RMSE(DR.simple.fore$mean,canada.ener_future$energy)
RMSE(DR.lagged.fore$mean,canada.ener_future$energy)
RMSE(canada.ener.hes.fore$mean,canada.ener_future$energy)
RMSE(canada.ener.ets.fore$mean,canada.ener_future$energy)

#Plot (validation of different models)
ts.plot(canada.ener_future$energy,DR.simple.fore$mean,DR.lagged.fore$mean,
canada.ener.hes.fore$mean,canada.ener.ets.fore$mean,canada.ener.boxjenkin.fore$mean,
col=1:6,gpars = list(ylab=”Energy␣consumption␣of␣Canada␣(toe)”,xlab=”Year”))
#col=c(”black”,”red”,”green”,”blue”,”pink”))
title(main = ”Validation␣of␣different␣models␣(2008␣to␣2013)”)

legend(”topleft”,
        legend=c(”Real”,”DR.simple”,”DR.lagged”,”HES”,”ETS”,”BoxJenkin”),
        col=1:6,
        lty=1,
        cex=0.48)
```