Predicting Canadian Tourism Demand

By: Julia Dedic

Introduction

The tourism industry is one of the most important parts of any countries economy. It has a great influence on the overall spending from the citizens and/or the government as well as the foreign income that will be spent within the country. This in-turn impacts employment by creating new and/or more opportunities, provides improvement to the country's infrastructure and services, and allows for better allotment of resources and improvement to the country's GDP.

Due to many technological advances, the act of travelling has become extremely convenient and accessible. Canada is one of the countries that benefits greatly from this industry. In particular, the tourism industry is represented more in Canada's national GDP than the agriculture, forestry, and fisheries sectors combined.[1] With this in mind, Canada's financial growth is largely dependent on its tourism demand. In this report, an analysis of the tourism-related spending in Canada will be conducted and further used to predict Canada's future growth in the industry as a measure of its economic conditions.

Motivation

To predict the economic growth of Canada for the next five year, tourism demand plays a significant role. Predicting the amount of tourism spending in the country in the next five years as well its growth rate is a very important financial indicator. Many groups such as local businesses, the government and economists can benefit from this information. Small businesses such as "Mom and Pop" shops, privately owned hotels and restaurants, and taxi services can use this information to predict future sales. With this information, the government can then allocate funds accordingly and further invest in tourism related infrastructures, advertisings, and promotions to encourage and meet the needs of Canadian tourism demand and support its growth. Finally, economists can use this information to help predict future economic activity in the country and further support how resources should be spent within the country to optimize future prediction.

Data and Methodology

The data was collected from CANSIM, Statistics Canada's socioeconomic database. It consists of 127 quarterly observations from the years 1986 to 2017.[2] The amount of spending per millions of Canadian dollars on tourism-related expenditures in Canada was recorded at each quarter. The data was split into training and testing sets. The quarterly spending from quarter one of 1986 to quarter four of 2011 was used as the training set and the spending from quarter one of 2012 until quarter three of 2017 was used as the testing set. Four different modelling techniques were explored on the data and compared to see which method yielded the best results on the testing set. The best method was then used to make a prediction for the next five years.

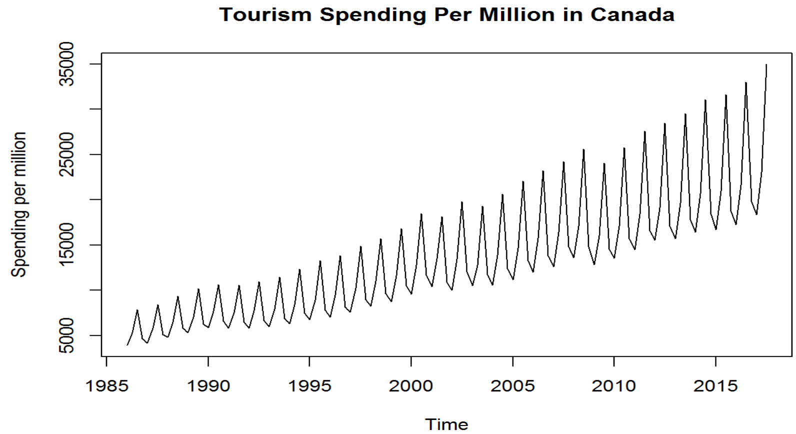

Figure 1. *Quarterly Tourism Spending Per Million of Canadian from 1896 to 2017*

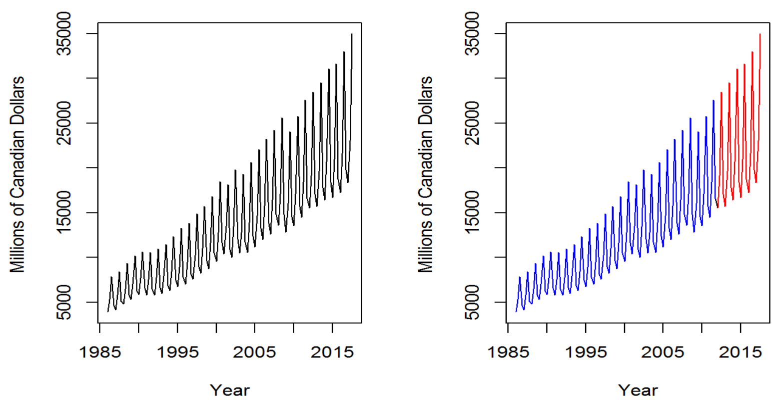

Figure 2*. Quarterly data set (****left****) split into training (****blue****) and testing (****red****)*

The four modelling methods used were: a Holt-Winters smoothing model, a linear regression model with SARIMA residuals, a Box-Jenkins approach comparing SARIMA models and a Bayesian Structural model. Analysis was conducted on each model which compared the residuals and predictions on the testing data. Mean Absolute Percentage Error or MAPE, was used as a measure of prediction accuracy for each model. MAPE was chosen since it expresses accuracy as a percentage and hence, it is the most natural statistic to interpret. As well, each model was also compared on other grounds including its fit to the data, normality and correlation of residuals, and prediction interval widths.

By looking at **Figures 1** and **2**, the data appears to have an increasing linear trend with strong multiplicative seasonality. Seasonality peaks seem to occur in the third quarter and troughs occur in the first quarter of each year. This may be due to the fact that tourism spikes coincide with Canada's favourable weather conditions (summer) which occurs roughly in the third quarter and unfavourable weather conditions in the first quarter (winter). There are no apparent outliers in the data.

By digging a bit further and decomposing the data, **Figure 3**, reveals two timestamps where the increasing linear trend seems to deviate. The timestamps are roughly 2003 and 2008. In 2003, the data shows a drop of roughly 5% in the second quarter of spending from 1.34 billion to 1.27 billion. This decline in spending continued for three more quarters until the first quarter of 2004.

After some research regarding the possible influences on the Canadian economy and tourism during this time period was conducted, it was found that this change point was most likely the cause of the SARS epidemic. This epidemic hit Canada around March 2003, with serious, widespread outbreak in the Toronto, Ontario area. This would have a huge influence on the Canadian tourism industry since Toronto is one of Canada's most visited cities.

The second change point appears to occur around the fourth quarter of 2008. Spending in the fourth quarter of 2008 was roughly unchanged from the year before, however, declines in tourism in every quarter were exhibited in 2009. The second and third quarter of 2009 in particular, saw the largest declines with over 6% declines in each quarter respectively. These declines would result in tourism spending in 2009 equivalent to pre-2007 levels. Again, some research pointed to the conclusion that this change may have coincided with the late 2008 global economic recession and U.S subprime crisis. Individuals were facing difficult financial situations and hence, reduced travelling funds would be prominent. Tourism spending in Canada did not recover until the third quarter of 2010. These change points were verified using the "changepoint" package in R.

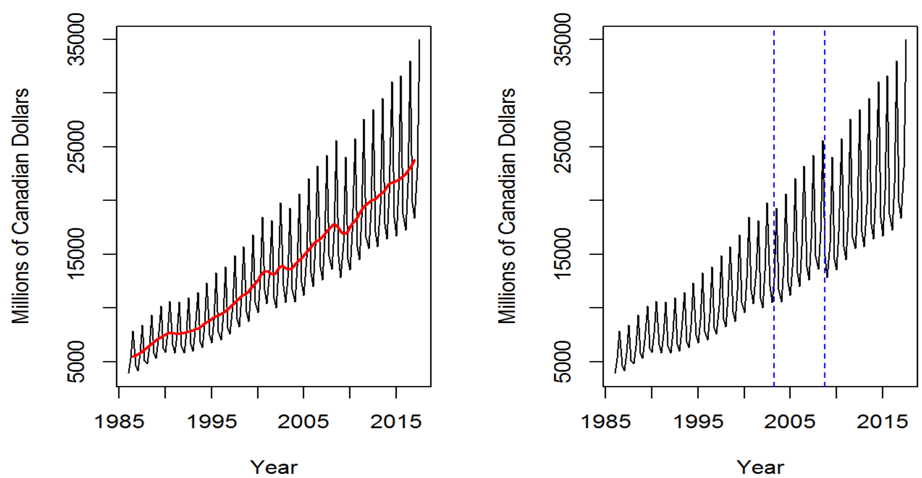

Figure 3. *Decomposed trend of data set (**left**), Change points in Q2, 2003 and Q4,2008 in data set (**right**)*

Since these change points seem to be very significant on the overall trend of the data, they will be incorporated in each of the four methods wherever possible. The change points can be seen in the plot on the right in **Figure 3** by the two blue vertical lines.

Data Analysis

Method 1: Holt-Winters Smoothing Model

Since the data exhibits increasing seasonal variance as the trend increases, the Holt-winters method with multiplicative seasonality is appropriate as a first method. **Figure 4** shows the fitted values versus the actual values of the training set using the Holt-Winters method.

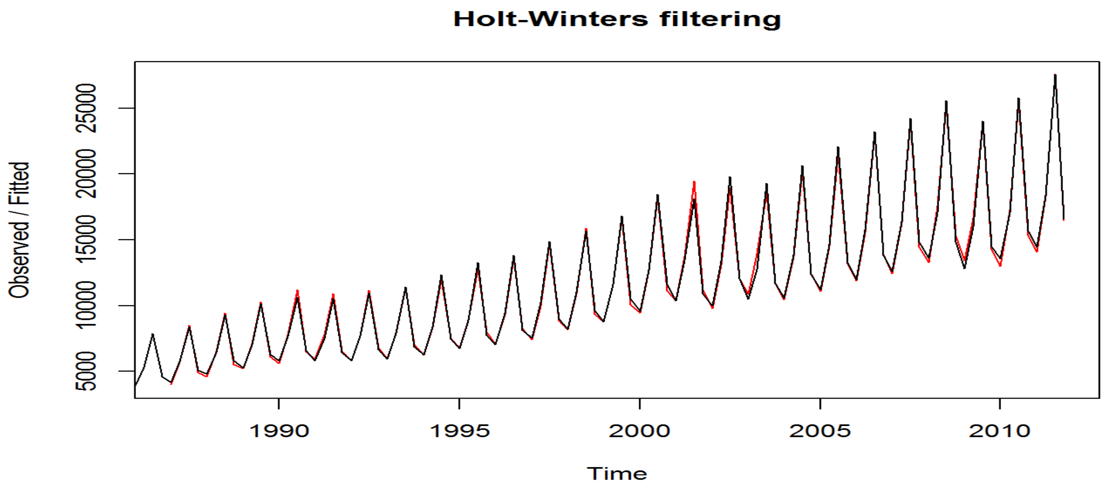

Figure 4. *Fitted (**red**) vs. Actual (**black**) values for training set using Holt-Winters*

Figure 4 shows that the Holt-Winters method does a fairly good job at fitting the data, although there are noticeable residuals after 2000. This is most apparent at peaks and troughs from 2000-2011.

Looking at the plots of residual diagnostics in **Figure 5**, notice that there are several spikes around 2003 and late 2008-2009. This provides evidence for the change points discussed earlier to exist. However, this is a major limitation of smoothing methods such as the one used. Smoothing methods such as the Holt-Winters method take into account past data, with the most emphasis on the most recent data and hence, tend to believe the trend will continue even when it does not. This means when the data exhibits change points, smoothing methods will have difficulty adjusting. By looking at the QQ plot, notice there is a departure from normality since we have heavy tails. This was confirmed with the Shapiro-Wilks test which gave a p-value close to 0, implying rejecting the normality assumption on the residuals. By looking at the ACF and PACF plots, notice there is only a bit of correlation at some lags but there does not seem to be any major issues with correlation.

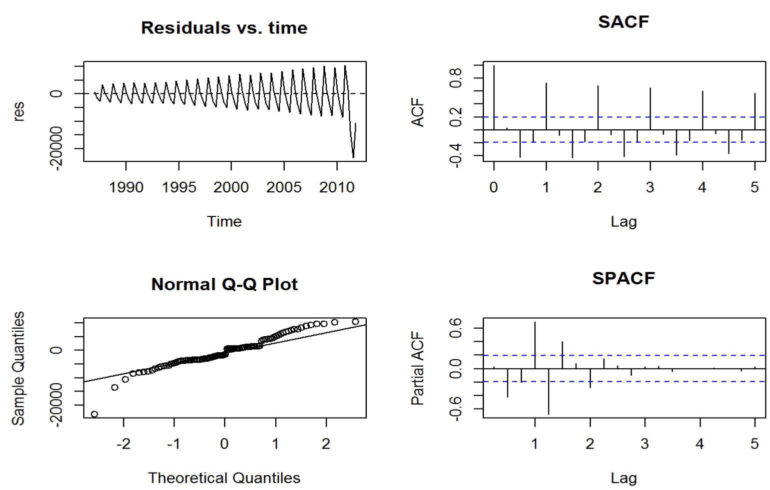

Figure 5: Residual Diagnostics plots using Holt-Winters Method

Using the testing data set, the predicted testing values from the Holt-Winters method are compared to the actual data set values from 2012-2017 in **Figure 6**.

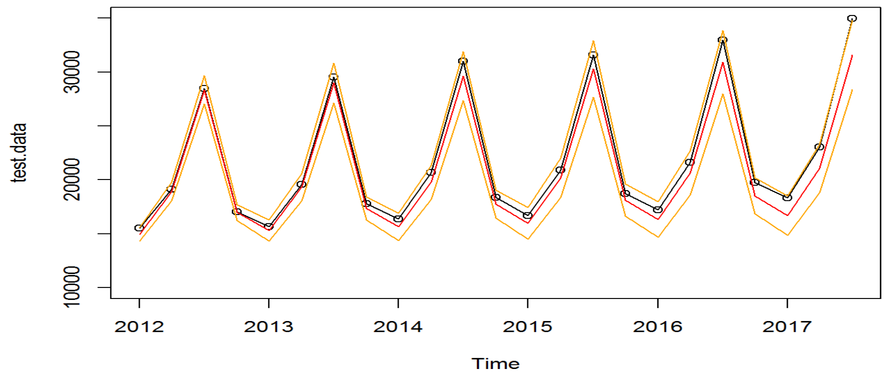

Figure 6. *Prediction (**red**) Vs. Actual Values (**black**) for testing set using Holt-Winters method*

By looking at **Figure 6**, notice that all 24 predictions fall within the 95% prediction interval (orange lines) so the Holts-Winters method seems to have done a fairly good job. However, comparing the prediction and actual values shows that the predictions are always below the actual values meaning that the model is constantly underestimating the actual values. As time increases, the difference between the actual and predicted values gets larger and the actual values approach the upper bound of the 95% prediction interval. This is most apparent between the years 2016 and 2017. Note that the prediction interval may be useless because we rejected the normality assumption of the residuals. The mean absolute percentage error is 4.11%. In conclusion, the Holts-Winters method seems to fit the data fairly well, however, prediction accuracy seems to be a bit questionable since the predictions are consistently underestimated. Another method is thus considered to take into account the change points and hence, improve prediction accuracy.

Method 2: Regression with SARIMA residuals using a Box-Jenkins approach

To start with a linear regression on the data set, a classical decomposition model was first fitted consisting of repressors for time and seasonality using indicator variables. As observed with the Holt-Winters method, since the data has increasing variance, a log transformation is appropriate to stabilize the data before the model is applied.

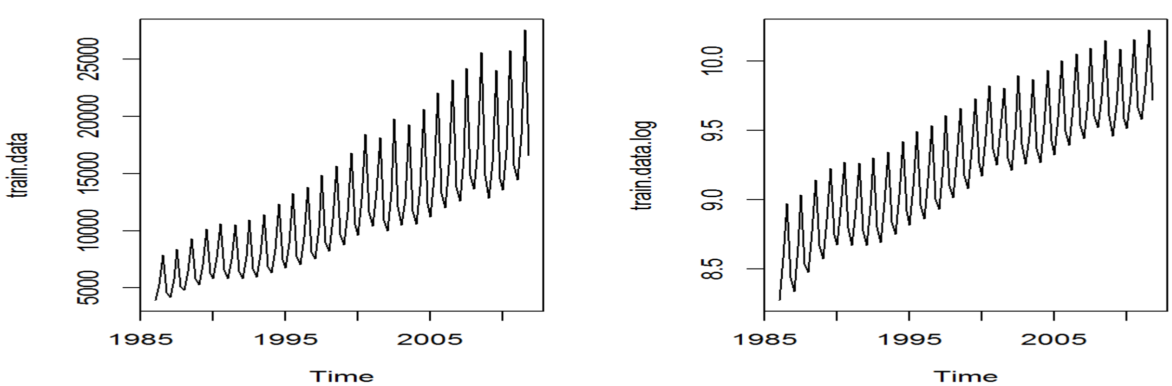

*Figure 7. Transformation of the original data (**left**) to the log transformed data (**right**)*

By taking the log transformed data and fitting it to the classical decomposition model, notice that the fit appears to be worse than the Holt-Winters method. The red line (fitted line) in **Figure 8** appears to have some significant departures from the black line (log training data) around 2001-2003 and 2009-2010. Not so coincidently, these departures line up with the change points earlier described.

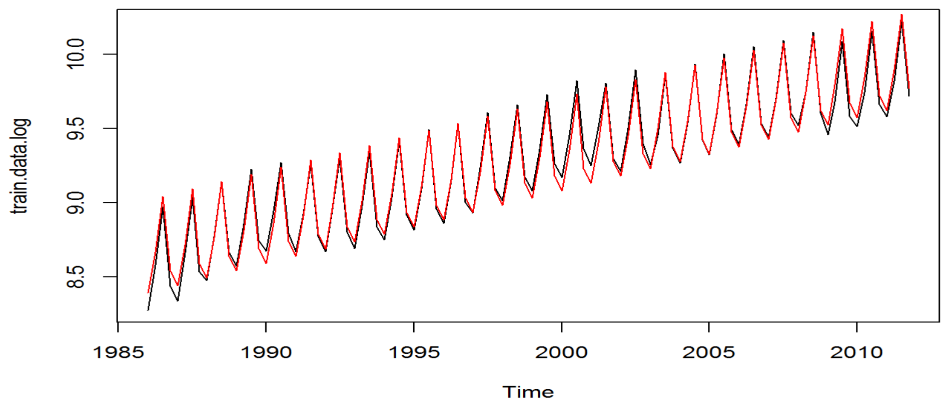

***Figure 8**. Fitted (**red**) vs. Actual values (**black**) for training set using Classical Decomposition model*

Due to the fact that it appears that the change points were not adequately taken into account in the Classical Decomposition model, adding two indicator variables as regressors to separate the data into three periods was done. This was intended to improve the fit of the model and take care of the more prominent residuals observed after 2000 in **Figure 8**.

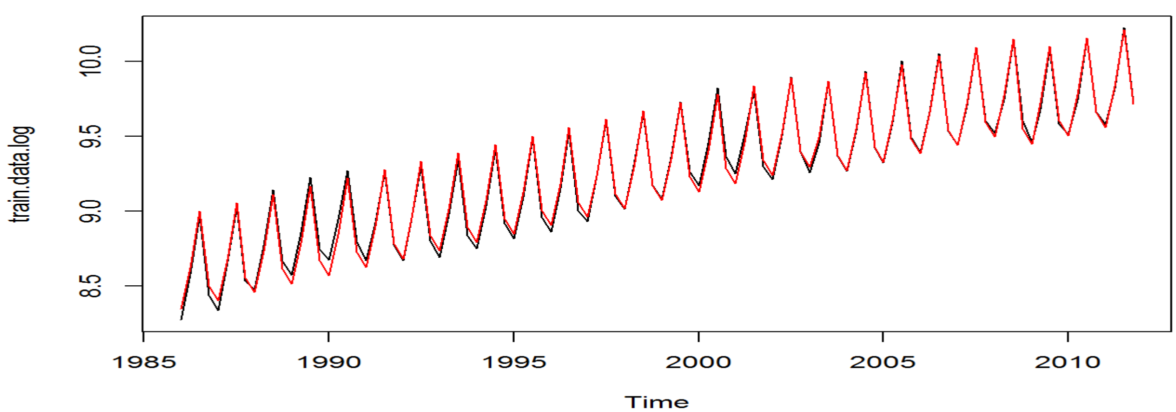

***Figure 9.** Fitted vs. Actual values for Regression model with change point regressors*

Notice that the prediction line (red) in **Figure 9** drops almost right on top of the actual values line implying this model has a better fit to the deviations within the trend.

To further improve the regression model, an ARMA model was considered for the data's residuals in order to remove correlation. By looking at the residuals in **Figure 11**, the residuals do not appear stationary since there is clearly a trend. This means that the ACF and PACF plots cannot be interpreted yet. However, one ordinary difference of the residuals was enough to make the data stationary and allowed a SARIMA model to be fitted to the residuals using the Box-Jenkins approach. Stationarity was verified using the Phillips-Perron Unit Root Test.

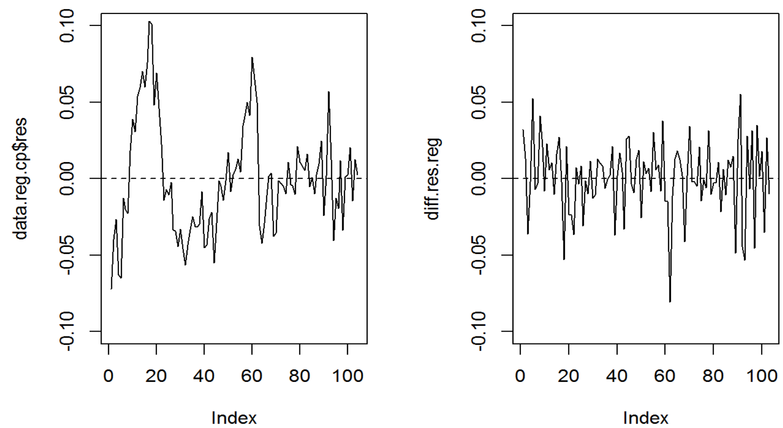

Figure 10. *Residuals of regression model before (**left**) and after (**right**) one ordinary differencing*

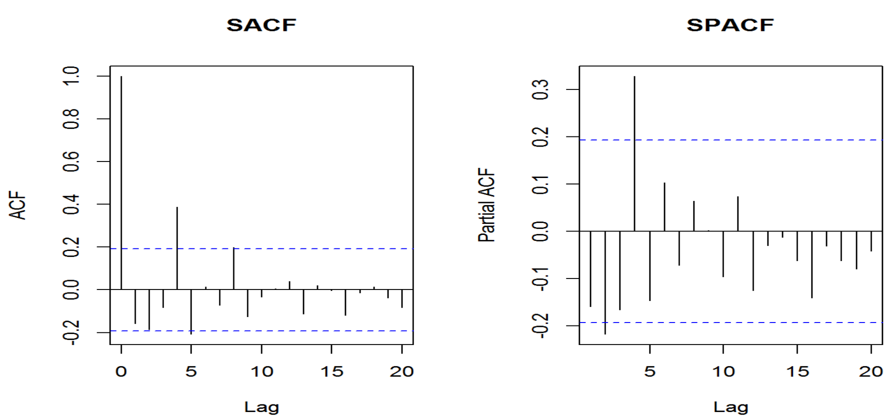

Figure 11. *Sample ACF and PACF of differenced residuals*

Plotting the sample ACF and PACF of the stationary differenced residuals (**Figure 11**), was used to determine the orders p, q, P, and Q of the SARIMA model. The seasonal lags, (I.E. lag 4,8,12, etc…) show one spike in both the ACF and PACF followed by decay, suggesting an ARMA (1,1) model, or P=1 and Q=1 for the seasonal component. For the non-seasonal component, there appears to be only two spikes in the ACF plot that are of significance and decay in the PACF. This suggests having orders of p=0 and q=2. Hence, a $(0,1,2) \times (1,0,1)_4$ SARIMA model was fitted to the residuals. Looking at **Figure 12**, there seems to be some improvement. The residual plots in **Figure 13** show that much of the correlation was removed. As well, normality is still violated with the residuals, as was the case in the Holt-Winters method. This was again verified with the Shapiro-Wilks normality test with a p-value of approximately zero. Additional improvements were taken into account including combinations of orders for the SARIMA model to the residuals, but none of these provided additional improvement in the removal of correlation or normality from this model.

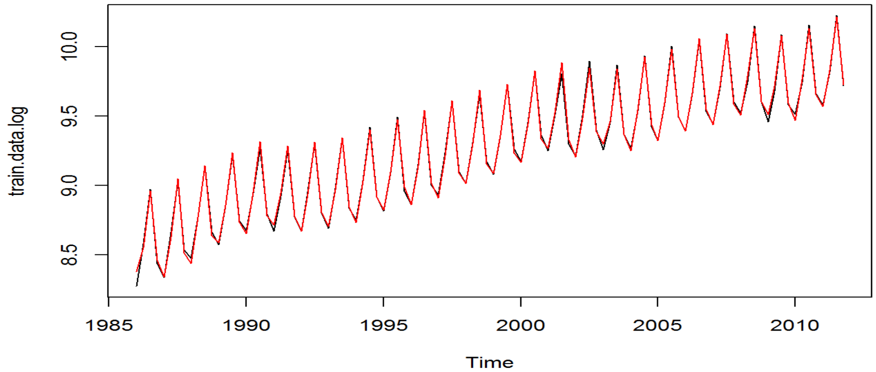

Figure 12. *Fitted (red) vs. Actual values (black) for Regression model with SARIMA residuals*

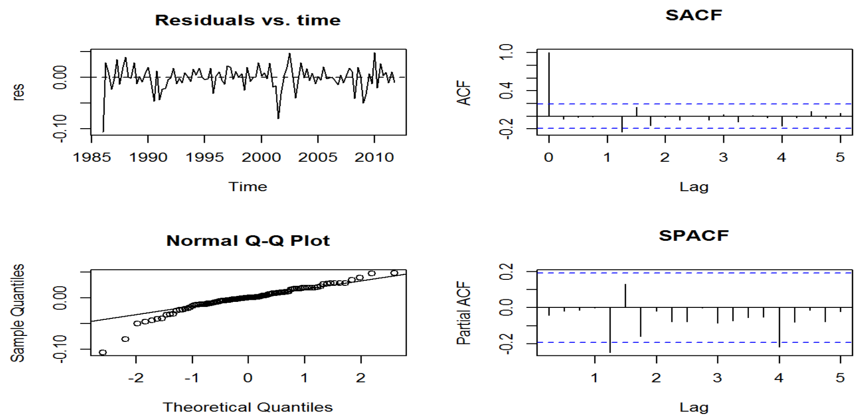

Figure 13. *Residual Diagnostic plots for Regression model with SARIMA residuals*

Again, additional improvements were considered for the regression model with SARIMA residuals. However, no further improvement was possible without introducing the issues of over fitting. The MAPE was 6.83% which was worse than the Holts-Winters method. The predictions using this model are seen in **Figure 14**. The predictions fit very well for the first few years but then appear to overestimate the actual values around 2015 until 2017. Even though no predictions fall outside the 95% prediction interval (orange lines), the intervals must be taken into consideration cautiously since the residuals violate the normality assumption.

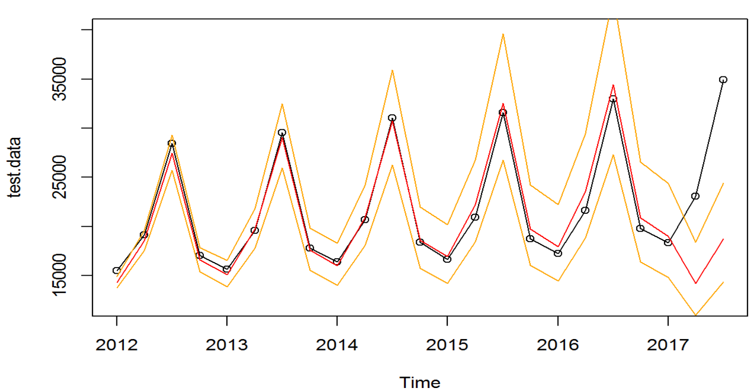

Figure 14. *Predicted (**red**) vs. Actual values (**black**) with testing set for Regression model with SARIMA residuals*

Method 3: Box-Jenkins Method

As mentioned in earlier methods, since the data has an increasing trend and multiplicative seasonality, the data required a transformation before any ARIMA models could be determined. After taking the natural logarithm to remove the increasing variance and applying a seasonal difference of period four, it was observed that the data was still not stationary. Hence, another ordinary difference was applied. The Phillips-Perron Unit Root Test verified that the series was stationary.

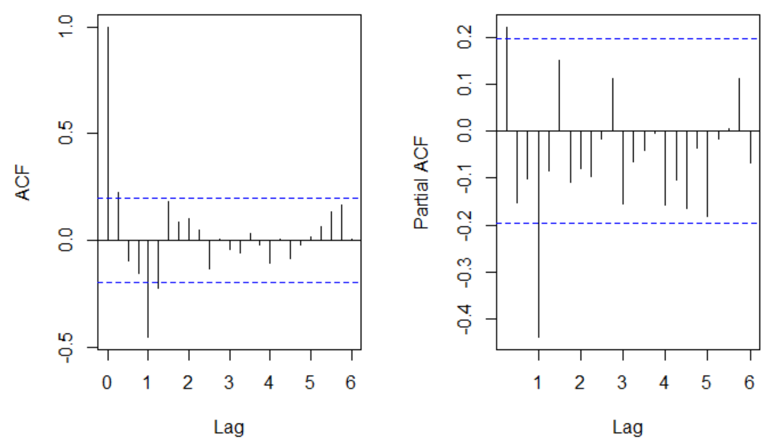

Figure 15. *Sample ACF (**left**) and PACF (**right**) of differenced time series*

By looking at the ACF and PACF in **Figure 15**, there did not seem to be any significant correlation at non-seasonal lags and hence, different combinations of p, q, P, and Q orders for the ARIMA models were compared using AIC to determine the best fit to the data.

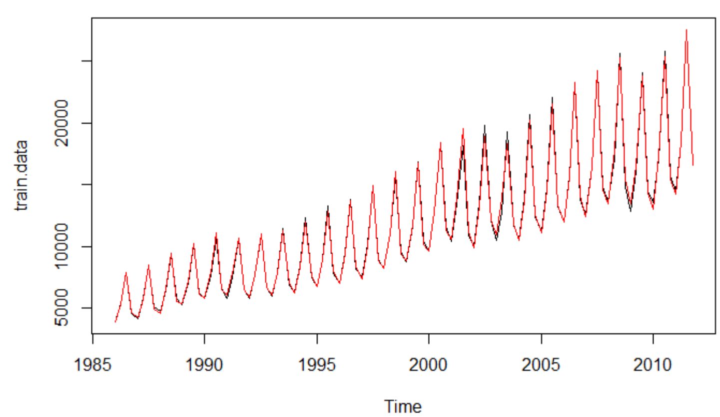

Figure 16. *Fitted vs. Actual for (0,1,0)x(0,1,1)$_4$ SARIMA model*

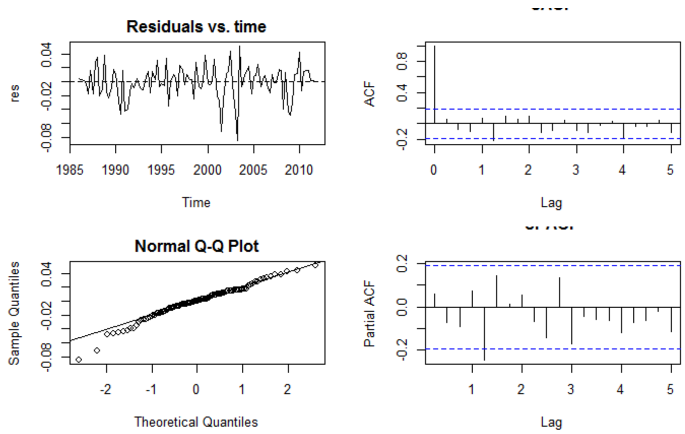

Figure 17. *Residual diagnostics plots for SARIMA model*

The SARIMA model with order $(0,1,0)\text{ x }(0,1,1)_4$ had the lowest AIC and hence, this model was selected. **Figure 16** and **17** show the model fit and the residuals of the selected model respectively. The residual plots show the same diagnostics as the Holt-Winters method. There is still slight correlation at lag 5 and there appears to be a problem with normality but seems to be better in this respect compared to the two methods prior.

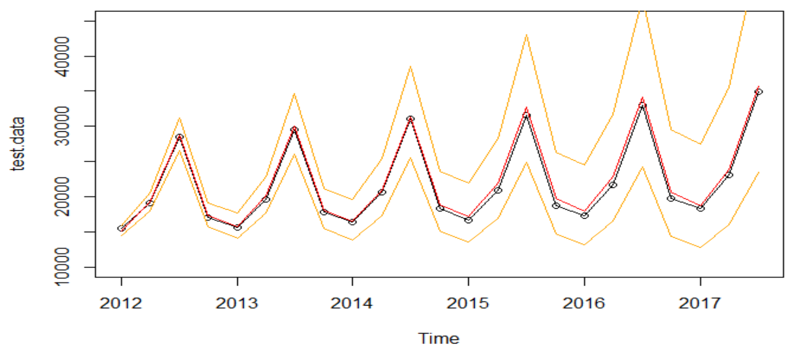

Figure 18. *Prediction (**red**) vs. Actual (**black**) values from testing set for SARIMA model*

By the Shapiro-Wilks test, normality is rejected. **Figure 18** shows the predictions of the model using the testing values. The MAPE is 2.62% which is a lot better than both of the previous methods. However, it is worth noting the larger width of the 95% prediction intervals (orange lines) due to the log-transformed data combined with the nature of the Box-Jenkins method.

Method 4: Bayesian Structural Model (with BSTS package)

Using a Bayesian Structural model is more transparent than ARIMA models and is better at dealing with uncertainty. There are no longer issues with normality assumptions. The Bayesian Structural model allows for quantifying the posterior uncertainty of individual components.[3]

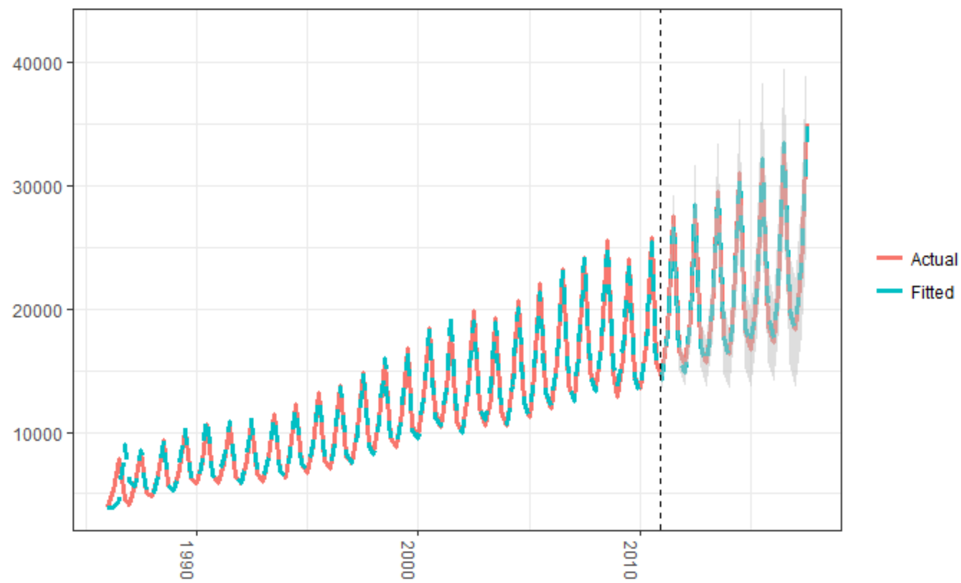

Figure 19: Actual vs. Fitted values for the Bayesian Structural model

Figure 19 shows the Bayesian Structural model that was fitted to the data set with 500 MCMC iterations using 2011-2017 as the hold out set to be consistent with the previous three methods mentioned above. A suggested burn-in was implemented to disregard the first MCMC iterations.[4]

Notice from **Figure 19**, that the Bayesian Structural model does a good job of fitting to the test set values. This means it does a good job capturing the seasonality within the data as well as the growth found from the increasing trend. The Bayesian Structural model has a holdout mean absolute percentage error (MAPE) of 1.36% which is the smallest among the four methods.

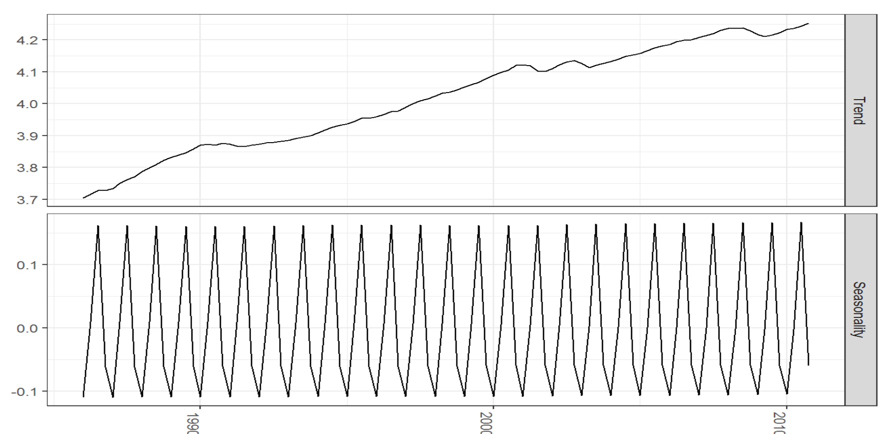

Figure 20. *Trend (**top**) and Seasonality (**bottom**) component of the Bayesian Structural Model*

Figure 20 shows the underlying components of the model.[5] This was done by averaging the MCMC draws to visualize the trend and seasonal components of the data.

Model Selection

Comparing the four methods was done by comparing the fit to the actual data, checking residual assumptions, the quality of the test set predictions, prediction interval widths, and the MAPE. Table 1 shows a summary of the four model methods used.

Table 1. Comparison factors for four methods

	Holt-Winters Method	Regression Model with SARIMA residuals	Bayesian Structural Model	Box-Jenkins Method (Selected SARIMA model)
Fit of Data	Good! There are a few noticeable residuals at peaks and troughs	Great! Fit data very well especially at change points	Great! Does a great job at capturing seasonality and growth in data	Good! However, few noticeable residuals at peaks and troughs
Residual Normality	Not normal	Not normal	Posterior looks Gaussian	Not normal but close
Prediction interval width	Narrow throughout time, starts to increase slightly in width	Starts off narrow but becomes quite wide	Starts narrow but becomes fairly wide	Starts narrow but becomes very wide
Prediction quality on testing set	Consistently underestimates	Consistently overestimates	Consistent throughout and gets better	Good but overestimates a bit at peaks
MAPE	4.11%	6.83%	1.36%	2.62%

Since ARIMA models are known to be very robust and make good "baseline" model comparisons, all models were compared to the Box-Jenkins method. From **Table 1,** notice that the Bayesian Structural model seems to perform even better than the best-chosen ARIMA model. The Bayesian model is able to capture the seasonality and trend components the best and does not have issues with its residuals. Also by taking the MAPE measurement into consideration, it is clear the Bayesian model has the least prediction error. From this, the Bayesian model was seen to be significantly better than the best ARIMA model and provided the best accuracy among the four methods. Therefore this model was selected for casting predictions.

Additional reasons for selecting the Bayesian model was due to its advantages. Bayesian structural models in general are more transparent than ARIMA models in the sense that they work pretty well with prediction uncertainty. Handling uncertainty properly is very important in this case since economies can be very unpredictable especially when both the economic state of Canada, as well as other countries, influence tourism spending. For instance, instability with the U.S's economy can have a very negative influence on Canadian tourism. Since the main goal is to forecast the next five years of tourism spending, a model that is able to adjust for uncertainty would be the ideal best choice.

Conclusion

By using the Bayesian structural model, **Figure 21** shows the 5 year forecast for tourism demand in Canada. The 95% credible intervals look like they are getting a lot larger as the prediction years increase to 2022 which is expected since the predictions are becoming less certain the further they get from 2017.

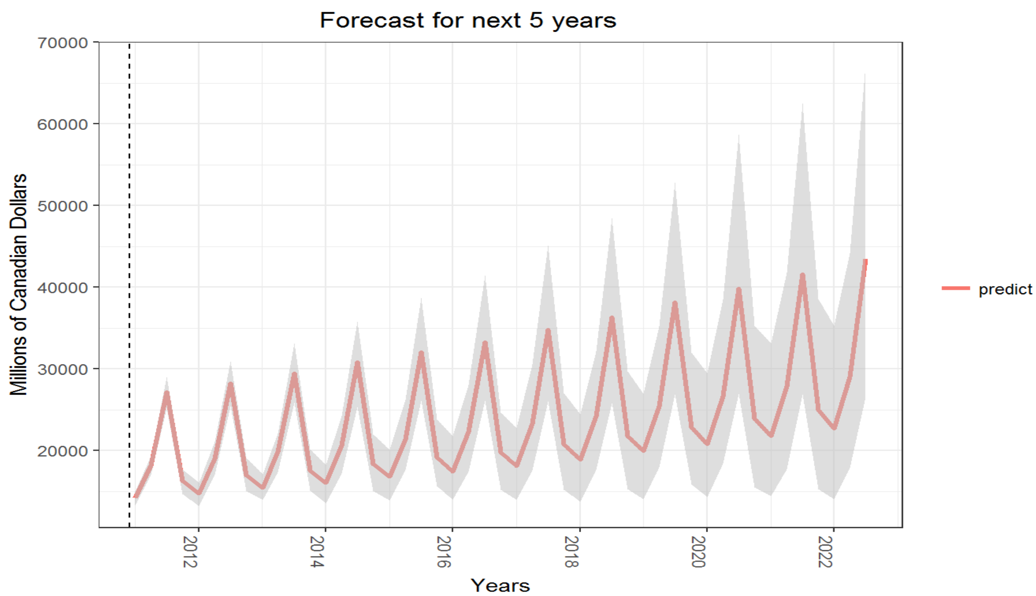

Figure 21. *A five-year prediction of tourism demand using Bayesian Structural model*

Below in Table 2 are the predictions for each quarter until the end of 2022 as well as the upper and lower bounds for the 95% credible intervals.

Table 2. *Five Year prediction values quarterly and 95% credibility interval limits*

Prediction Value	Lower Bound	Upper Bound	Date
19689.06	13776.92	24463.54	2018-01-01
25330.47	17653.62	32173.14	2018-04-01
37723.94	25859.44	48533.44	2018-07-01
22691.30	15338.26	29728.59	2018-10-01
20691.30	14122.58	27016.02	2019-01-01

26471.47	18009.47	35253.74	2019-04-01
39583.07	26867.63	52871.93	2019-07-01
23673.73	15880.92	32034.92	2019-10-01
21554.22	14340.56	29475.22	2020-01-01
27688.97	18399.02	38519.55	2020-04-01
41350.16	26942.78	58749.35	2020-07-01
24829.02	15520.38	35304.25	2020-10-01
22597.66	14450.00	33187.83	2021-01-01
28926.99	17727.98	41819.01	2021-04-01
43399.20	26936.78	62523.82	2021-07-01
25952.21	15331.15	38580.18	2021-10-01
23605.42	14113.16	35310.09	2022-01-01
30274.85	17866.51	44111.22	2022-04-01
45157.43	26616.39	67338.08	2022-07-01
26645.55	15038.99	41866.35	2022-10-01

Table 2 verifies how Canadian tourism demands are seasonal. The same trend with troughs in the first quarter and peaks in the third quarter persist throughout all five years of predictions. Notice that the model predicts that Canadian tourism demand will continue to increase over the next five years however, the increase in demand will not be as large around 2022 compared to 2019 and 2018. Below in **Table 3**, shows the growth rate year after year based on the previous year for the last ten years as well as the next projected five years from the Bayesian Structural model.

Table 3. *Projected growth from previous year for the five year predictions*

Growth from previous year	Year
5.1%	*2007*
4.5%	*2008*
-5.25%	*2009*
7.0%	*2010*
6.9%	*2011*
4.1%	*2012*
3.3%	*2013*
4.65%	*2014*
1.67 %	*2015*
4.2 %	*2016*
4.12%	*2017*

4.56%	2018
4.73%	*2019*
4.50%	*2020*
4.73%	*2021*
3.97%	*2022*

From **Table 3**, notice that the trend is increasing and continues to increase for the next five years. One thing that is interesting to note, is the negative growth rate in 2009, which appears to be followed by above average growth in 2010 and 2011. The highest growth rate in the last ten years was in 2010 at a rate of 7%. For the next five years, predictions indicate that tourism spending is going to increase on average 4% per year. A next step in this analysis would be to explore the events that cause these fluctuations in growth rates and determining if there are any major events that could have resulted in them. This could lead to discovering reasons that lead to quick recoveries and very high growth rates in tourism demand.

Several conclusions can be drawn from this analysis. Firstly, economists can view these predictions as a sign of the Canadian economy continuing to grow and showing a steady increase of spending in the country. The government may see this as justification for their spending on tourism-related infrastructure and advertising and hence, may continue to invest in this industry. For businesses in the tourism industry or businesses located in tourist areas, they can expect to see a steady increase of about 4% in traffic and spending annually.

As noticed throughout the analysis, tourism spending is very seasonal and hence, this should be taken into consideration when planning to make financial decisions. For instance, businesses should expect more visitors in the summer months (quarter 3) and fewer visitors in the winter months. This means as well that there will be more competition in the summer. As seen at the two change points, 2003 and 2008, spending can drastically decrease due to unexpected negative events and these decreases would have a much larger impact on the tourism industry in quarter 3 than any other quarters. These events should be planned for in future expenditure budgets.

Lastly, this analysis can give the government and economists insight to how resources and spending should be planned for in the upcoming years. With proper planning, the potential growth could thus be maximized and the government could then use this increased income to improve the overall state and well-being of the country.

Future Considerations

After conducting this analysis, the following future additions can be added: a hierarchical model can be explored given tourism spending from different provinces. Google correlate could also be explored for possible covariates. If a more detailed and larger dataset was available, these additions could be conducted and would be a great way to break down Canada tourism spending fluctuations by province. Using keyword searches about or related to Canadian tourism could be analyzed against Canadian tourism spending. This would give insight as to whether different provinces' tourism spending is increasing at the same rate. Therefore, these additions would give better insight and provide a more detailed approach to accurately depict the appropriate spending on tourism via federal and provincial funding.