Weather Adjustment of Economic Data – Beyond Seasonal Adjustment

Steve Matthews[1,2] , Zdenek Patak[1,3]

[1]Statistics Canada, 100 Tunney's Pasture Driveway, Ottawa, ON, K1A 0T6, Canada
[2]steve.matthews@canada.ca
[3]zdenek.patak@canada.ca

Abstract
Weather extremes are becoming more pronounced and more prolonged. It is conceivable that long stretches of unusually low or high temperatures, drought conditions, or extended rainfall, may impact the behaviour of both individuals and businesses. While seasonal adjustment removes repeating, equally spaced patterns along with calendar effects such as trading days and moving holidays, it does not remove the potential impact of atypical weather on economic time series. This paper explores this issue. Using X-12-ARIMA, weather related regressors are incorporated into seasonal adjustment to help explain unusual movements in various economic time series. Potential alternatives and ad hoc analyses are discussed.

Key Words: X-12-ARIMA, Seasonal Adjustment, Regression Analysis

1. Introduction

1.1 Economic time series data produced at Statistics Canada

Statistics Canada collects information on many aspects of the Canadian economy. Its breadth spans the supply and demand sides of the labour market (employment, job vacancies, hours worked), industrial input and output (manufacturing production, producer price index, international trade), and various factors related to personal consumption (retail sales, travel, price indexes). Estimates are further disaggregated by geography, industry, and demographic attributes.

Many of these estimates are subject to seasonal fluctuations and must be suitably adjusted to reveal the underlying signal, or trend. Seasonal adjustment removes repeating, equally spaced patterns that constitute a seasonal pattern (such as average seasonal weather fluctuations attributed to climate), as well as movement due to moving holidays (Easter, Labour Day, etc.) and other calendar effects (trading day, leap year, etc.).

Weather extremes appear to be more commonplace. It is conceivable that long stretches of unusually low or high temperatures, drought conditions, or extended rainfall, may impact the behaviour of both individuals and businesses. The departures from climatic norms may explain, at least in part, anomalous movements in economic series. Merely noting that abnormal weather may have played a role in adversely affecting economic data has evolved into a more systematic approach to incorporate weather into the seasonal adjustment process.

1.2 Weather data available to Statistics Canada

There are many sources of weather related data containing a multitude of variables (for example, temperature, precipitation, and many others) at various degrees of granularity (either by geography or frequency). The geographic detail takes on greater importance as a number of distinct climatic regions exist within Canada. The time dimension is less important as most weather information is available on a daily basis.

The weather data used in this paper were downloaded from the Environment and Climate Change Canada website (www.ec.gc.ca). The information is available for individual weather stations at various frequencies (hourly, daily, etc.). The collected data typically include minimum, maximum, and mean temperatures, heat degree days, cool degree days, total rain, total snow, total precipitation, snow on the ground, and speed and direction of maximum wind gusts. The availability of a particular set of variables varies from station to station, but core information such as temperature and precipitation is collected by all. One challenging aspect of constructing pertinent weather variables is ensuring continuity through time as meteorological stations open and close, and the desired span of data may not always be available. Fortunately, this is less of an issue in large urban centres where measured economic activity is concentrated and the impact of severe weather would be the most pronounced.

1.3 Approaches to estimate the effect of weather on time series data

There is a growing body of literature estimating the effect of weather on economic time series. The consensus is that abnormal weather plays a role in explaining, but does not entirely account for, anomalous behaviour in economic series.

Using linear regression, Bloesch and Gourio (2015) conclude that weather has a significant but short-term effect on economic activity in industries such as utilities, construction, hospitality, and, to a lesser extent, retail.

Davies and Elliott (2015) demonstrate that weather has a significant impact on retail sales of clothing and on road accidents. They investigate a “switching effect” where an increase or decrease in clothing sales in one month was followed by a compensatory bounce-back in the following month. Starr-McCluer (2000) shows that atypical weather is an important contributor in explaining monthly fluctuations in retail sales but its importance largely disappears when analyzing quarterly data.

The evidence provided in the literature warrants the use of weather variables to partly explain anomalous movements in economic data series. The format of the paper is as follows. Section 2 presents an approach (preferred by many in the literature) for analysing the impact of weather on time series and describes seasonal adjustment using X-12-ARIMA. Section 3 discusses the integration of weather data into the seasonal adjustment process, and provides an example with results. Alternative analyses and possible future directions are explored in Section 4. Concluding remarks follow in Section 5.

2. Preferred approach for analysis

There are a number of options available to model the impact of abnormal weather on the movement in economic series in a regression setting. The ultimate method chosen must take account of the correlation structure that is often present in time series. A well-established technique, regARIMA, allows the errors in the regression model to follow an ARIMA process.

The general form of a regARIMA model is

$$\phi(B)\Phi(B^s)(1-B)^d(1-B^s)^D(y_t - \sum_i \beta_i x_{it}) = \theta(B)\Theta(B^s)a_t, \quad [1]$$

where y_t is the value of the time series at time t, x_{it} is the value of the explanatory variable i at time t, β_i is the regression coefficient associated with variable i, B is the backshift operator (where $By_t = y_{t-1}$), d and D denote the non-seasonal and seasonal differences, $a_t \sim N(0, \sigma^2)$, $Cov(a_t, a_s) = \begin{cases} 0, & \text{for } t \neq s \\ \sigma^2, & \text{for } t = s \end{cases}$, ϕ and Φ specify the non-seasonal and seasonal autoregressive components of the model, and θ and Θ denote the non-seasonal and seasonal moving average components of the model.

The $\sum_i \beta_i x_{it}$ component of the model captures various types of outliers, calendar effects, and other user defined regressors. Here, one or more regressors can be included to assess the impact of deviations in weather conditions from historical averages. Average weather conditions are assumed to be part of the repeating, equally spaced pattern that constitutes seasonal effects that are removed by default as part of seasonal adjustment. It is the unusual fluctuations in weather conditions that are of interest that would typically manifest themselves as spurious movements in the underlying time series.

Equation [1] may also be expressed in a more compact notation as

$$y_t^* = y_t - \sum_i \beta_i x_{it} . \quad [2]$$

In this instance, the x-variable enters the model as a linear quantity. More sophisticated models could include various economic concepts (supply and demand, price effects, international influences, etc.) to refine the impact of weather by delineating it from other influences. Note that these are not discussed in this paper.

The regARIMA methodology is implemented in the X-12-ARIMA software that is generally used for seasonal adjustment. The method is well understood and is used by many national statistical agencies. X-12-ARIMA computes seasonal factors that are based on successive applications of moving averages. These seasonal factors are subsequently removed from a time series resulting in seasonally adjusted data.

The variable y_t^* can also be expressed as a sum of trend-cycle (C_t – combination of long term trend and shorter term business cycle), seasonal movement (S_t)and irregular component (I_t) as follows:

$$y_t^* = C_t + S_t + I_t , \quad [3]$$

which is a well-known additive decomposition model. This is the starting point of the famous X-11 (note that X-12 is X-11 with a regression pre-processor to handle outliers and a number of calendar effects) algorithm that applies moving average filters to y_t^* in an iterative fashion to extract the individual components. If the trend-cycle and seasonal components explain most of the variability in y_t^*, then the irregular component is assumed to be randomly distributed around zero with a small variance. If a discernible pattern is still present, then one can turn to weather as a potential source.

Modelling weather within the existing framework of regARIMA and X-12-ARIMA has a number of advantages. The framework automatically detects outliers, models a number of calendar effects such as various moving holidays, trading day factors, and other predefined regression variables, correctly accounts for the serial correlation structure that is typically present in time series, and identifies a seasonal pattern. The framework can examine the irregular component for the presence of a pattern that could be partially or fully explained by incorporating an additional regressor such as seasonal fluctuations in weather conditions.

There are various degrees (options) of integration of weather regressors into the seasonal adjustment framework. Option I would be to maintain the status quo (no integration) and assist the analyst on the side and on a case-by-case basis. Option II would be to *partially* include a weather regressor associated to the irregular component where its impact would be removed from the computation of the seasonal factors, but would remain in the seasonally adjusted data.

And finally, Option III would be to *fully* include a weather regressor associated to the seasonal component where the corresponding effect would be removed from the seasonally adjusted series. This approach would be suitable for a production environment once it has been confirmed that a particular weather variable is tied to the movement of an economic series. The level of integration of weather in the analysis of time series data would be a joint decision of the data and time series analysts.

3. Example

There are several considerations when selecting and using weather variables. First, temperature deviations from historical averages represent the aspect of climate in its most basic form that is not removed by seasonal adjustment. It is easily interpretable and communicable. Second, the weather variable can be weighted by the size of the population to which it applies. This is important in countries like Canada where population density is very large in urban centers in the south and then drops off quickly as one moves north. As one moves away from densely populated areas, the impact of weather on economic activity diminishes. Third, the selected weather descriptor should have the greatest explanatory power and an apriori justification for its inclusion. Finally, the variables can be modelled as continuous linear variables (temperature departures from historical average, total precipitation), as step function variables (temperature categorized into cold, seasonal, and warm), as binary variables (presence or absence of extreme temperatures, occurrence of extreme snowfall), or even as non-linear variables (temperature being non-linear beyond a certain threshold).

The example presented in this paper uses readily available weather data from Environment Canada. The focus here is on maximum daily temperatures. For each location and combination of month and year, we compute a 5% trimmed mean of maximum daily temperatures and then take the average of the trimmed means to represent the monthly historical averages. Locations within the desired geography that have continuous weather information spanning at least 20 years form the basis for the computation of the historical averages. The monthly deviations from these values form the input for the population weighted temperature regressors. The weather files also contain other useful weather descriptors such as rainfall, snowfall, and wind speed that may be exploited in the future.

The focus is on two large urban centres in Alberta, Calgary and Edmonton, which are assumed to represent their provincial population. The economic time series used in this example is consumer gas sales in Alberta, a commodity whose consumption is highly negatively correlated with temperature. Separate monthly regressors were used to allow the effect to vary between months. This approach was taken to explain, at least in part, any non-systematic residual seasonality in the data. The example illustrates a simple application of a weather regressor to model an economic data series that shows the range of its explanatory power. The graph below shows a strong linear relationship between departures from the average monthly temperatures and the irregular component (residual term) that remained after accounting for seasonality and trend-cycle. There is ample evidence that temperature can further explain the lack of randomness in the X-12-ARIMA irregular component.

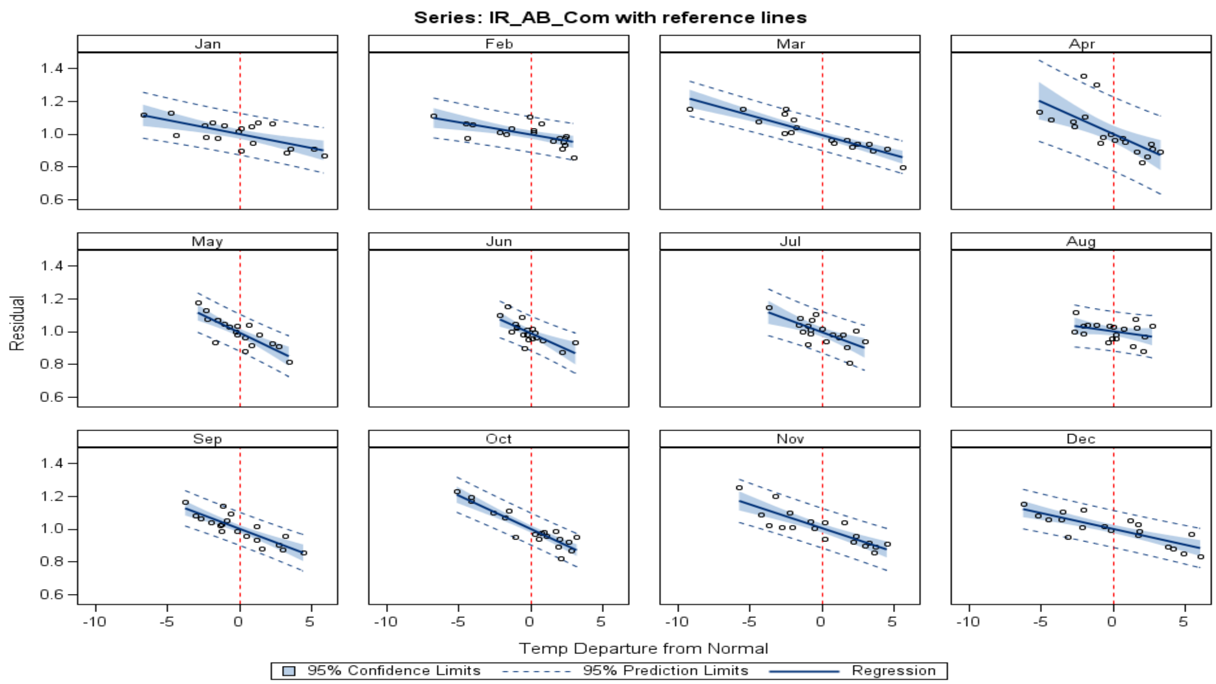

Figure 1. X-12-ARIMA irregular component vs. temperature departures from historical averages.

The next figure portrays the impact of specifying temperature as regressor associated to either the irregular or the seasonal components. The first is used to clean up the data prior to computing the seasonal factors. The analyst can then assess the importance of temperature departures from historical averages by studying the irregular component. Any highly anomalous values may coincide with highly abnormal temperatures.

The seasonal type regressor is incorporated directly into the seasonal factors and is removed from the seasonally adjusted data. This is a more direct way of using a weather regressor. It is the role of the analyst to decide which type of regressor is more appropriate. The corresponding results can be quite different as evidenced by the graphs in Figure 2. The impact of the choice of the type of regressor on the irregular component is shown in Figure 3. This is a good example of how much more of the variability in the series can be explained by a pertinent correlate.

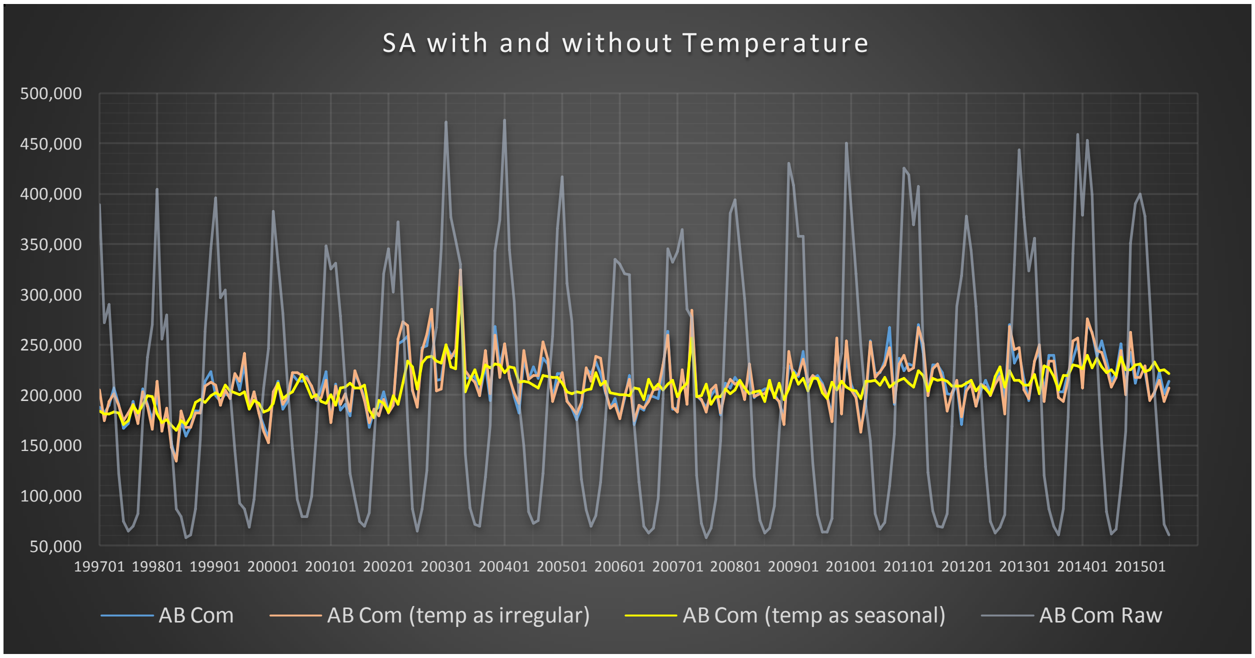

Figure 2. Comparison of seasonally adjusted series with and without temperature regressor.

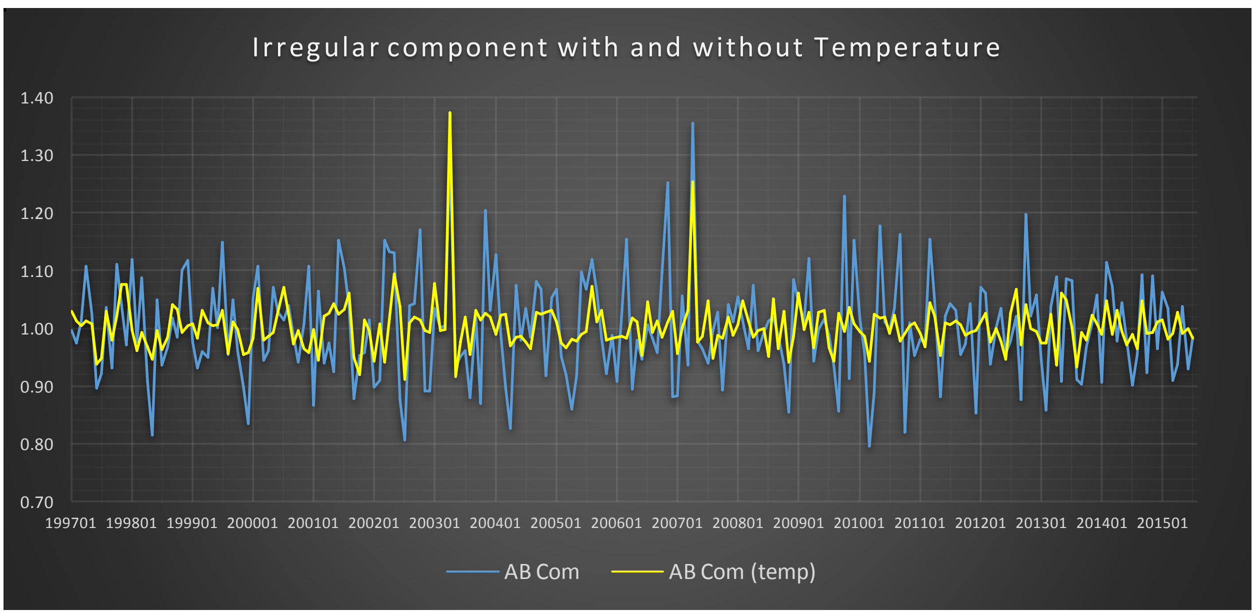

Figure 3. X-12-ARIMA irregular component with and without temperature regressor.

4. Alternative Analyses and Possible Future Directions

To estimate the effect of atypical weather on time series, regARIMA models were used with a temperature regressor to represent the current month's departures from a historical average. Some authors suggest that when economic activity is impacted by weather, there is a "switching effect" or "bounce back" where the displaced economic activity contributes in its entirety to the subsequent period(s). If a single period is involved, this may be modelled by a contrast where the regressor is defined as

$$x_{i,t} = \begin{cases} \Delta_{temp,t}, & \text{if } month(t) = i \\ -\Delta_{temp,t-1}, & \text{if } month(t-1) = i \\ 0, & \text{otherwise} \end{cases}$$
$$i = 1,...,12$$

Complicating matters is the fact that the displaced economic activity may not always bounce back during the subsequent period and it may not be linear if it extends over multiple periods. Initially, linear decay followed by varying rates of exponential decay can be used to determine which provides the best fit. The analysts responsible for the economic time series under investigation could provide valuable input regarding the most appropriate models. The assumption of linearity can be tested using classical statistical theory. It is possible that beyond a certain threshold the weather regressor may be non-linear, in which case more complicated regressors such as step functions or higher-order polynomials could be considered.

5. Concluding Remarks

Weather, in particular temperature deviations from historical averages, was found to explain unexpected movement in many economic time series. In industries such as utilities where consumers react to weather anomalies in real time, the increases and decreases in demand closely track temperature fluctuations. In other industries such as retail, the impact due to weather was confined to certain times during the year. For example, the onset of buying spring merchandise may be delayed by a long winter.

Modelling weather is yet another tool that analysts can use when looking at movement associated with a particular economic series. In some instances, it seems very reasonable to incorporate weather anomalies into the combined seasonal factors thus reducing the unexplained variability in the data. In other cases, the impact of weather may be more subtle.

At this point, incorporating information on atypical weather into the general seasonal adjustment process is premature. More research is necessary before this methodology can be implemented in a full production environment. Until such time, weather analysis will be conducted on a case by case basis.

6. References

Bloesch, J. and Gourio, F. (2015). The effect of weather on U.S. economic activity. *Economic Perspectives*, Vol.39, 1st, 2015.

Davies, J. and Elliott, D. (2015). Analysing the impact of weather and climate on official statistics time series. *Survey Methodology Bulletin*, Vol.74, Autumn 2015.

Starr-McCluer (2000). The Effects of Weather on Retail Sales. *Volume 8 of Finance and economic discussion series*. Divisions of Research & Statistics and Monetary Affairs, Federal Reserve Board, 2000.