BANK OF CANADA
BANQUE DU CANADA

Staff Discussion Paper/Document d’analyse du personnel—2024-17

Last updated: November 18, 2024

Seasonal Adjustment of Weekly Data

by Jeffrey Mollins and Rachit Lumb

Canadian Economic Analysis Department
Bank of Canada
jmollins@bankofcanada.ca

DOI: https://doi.org/10.34989/sdp-2024-17 | **ISSN** 1914-0568

Acknowledgements

We would like to thank Tatjana Dahlhaus, Pierre-Yves Yanni, Joshua Slive, Virginie Traclet and Daniel de Munnik for their helpful comments and suggestions. We would also like to thank Jordan Press and Nicole van de Wolfshaar for their editorial assistance.

Abstract

This paper summarizes and assesses several of the most popular methods to seasonally adjust weekly data. The industry standard approach, known as X-13ARIMA-SEATS, is suitable only for monthly or quarterly data. Given the increased availability and promise of non-traditional data at higher frequencies, alternative approaches are required to extract relevant signals for monitoring and analysis. This paper reviews four such methods for high-frequency seasonal adjustment. We find that tuning the parameters of each method helps deliver a properly adjusted series. We optimize using a grid search and test for residual seasonality in each series. While no method works perfectly for every series, some methods are generally effective at removing seasonality in weekly data, despite the increased difficulty of accounting for the shock of the COVID-19 pandemic. Because seasonally adjusting high-frequency data is typically a difficult task, we recommend closely inspecting each series and comparing results from multiple methods whenever possible.

Topics: Econometric and statistical methods
JEL codes: C1, C4, C52, C8, E01, E21

Résumé

Dans cette étude, nous présentons et évaluons les méthodes les plus répandues pour désaisonnaliser des données hebdomadaires. L'approche standard de l'industrie, connue sous le nom de X-13ARIMA-SEATS, se prête seulement aux données mensuelles et trimestrielles. Compte tenu de la disponibilité accrue et du grand potentiel des données non traditionnelles de plus haute fréquence, d'autres approches doivent être envisagées pour extraire des signaux utiles au suivi de l'économie et à l'analyse de l'information. Cette étude examine donc quatre méthodes de désaisonnalisation des données de haute fréquence. Nous constatons que le calibrage des paramètres de chaque méthode permet d'obtenir une série correctement ajustée. Nous optimisons le processus en utilisant une méthode de recherche par quadrillage et procédons à des tests destinés à détecter la présence d'une saisonnalité résiduelle dans chaque série. Bien qu'aucune méthode ne convienne parfaitement à toutes les séries, certaines s'avèrent généralement efficaces pour supprimer les variations saisonnières des données hebdomadaires – malgré la difficulté accrue qu'impose le fait de devoir tenir compte du choc de la pandémie de COVID-19. Puisqu'il est souvent difficile de désaisonnaliser des données de haute fréquence, nous recommandons d'examiner de près chaque série et, si possible, de comparer les résultats de diverses méthodes.

Sujet : Méthodes économétriques et statistiques
Codes JEL : C1, C4, C52, C8, E01, E21

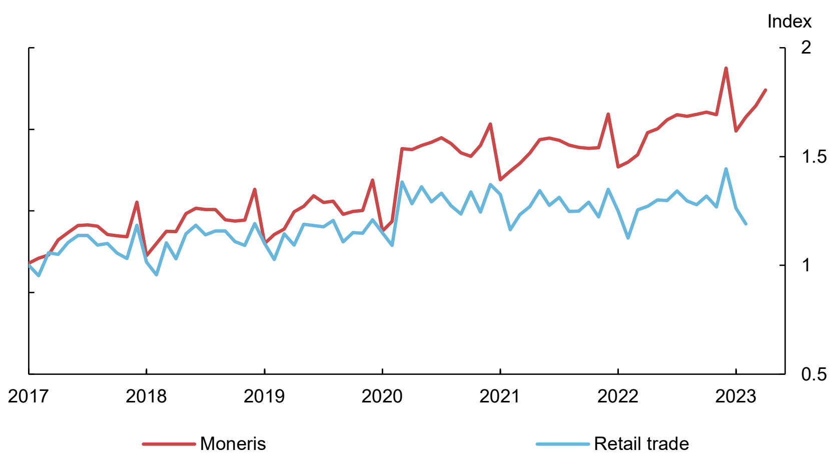
1. Introduction

Weekly data can be a very powerful source of timely information, particularly for nowcasting when official data are released with a lag. Recent literature, such as Aastveit et al. (2020), Fenz and Stix (2021), Lewis et al. (2021) and Monteforte and Raponi (2019), highlight the power of high-frequency information to accurately nowcast growth in gross domestic product (GDP). Using up-to-date information about the economy can be highly beneficial for central banks and inform policy-making in a timely manner.

We use Moneris transaction data to illustrate this benefit. Moneris, a payment processing company, aggregates debit and credit card transactions, and the data are broken down by merchant category. We find strong signals when we aggregate the transaction data to a monthly frequency and compare the results with their closest retail trade counterpart (nominal, non-seasonally adjusted) (**Chart 1**).

Chart 1: Spending on groceries

Monthly data, index: January 2017 = 1



Sources: Moneris and Statistics Canada

Last observations: Moneris, April 2023; Retail trade, February 2023

However, interpreting weekly data can be difficult because of the high degree of volatility, outliers and breaks (Ladiray et al. 2018). Moreover, seasonal components are difficult to isolate and remove even though doing so is critical to properly use the data. For example, deciding whether an increase in volume of transactions in mid-December is a sign of strength—or simply related to increased spending ahead of Christmas—can be problematic without some form of adjustment.

Given this interpretability challenge, weekly data should be seasonally adjusted. However, weekly seasonal adjustment is challenging, and no consensus exists in the literature on the most reliable method to do so. The main reasons that weekly seasonal adjustment is difficult are the following:

- A year does not contain precisely 52 weeks, but 52 weeks and one day, or two days during a leap year (i.e., non-isochronicity).
- Holidays do not always fall on the same day or week each year.
- Weekly data are noisy, making extracting the seasonal component challenging.
- Large or long-lasting outliers (such as the COVID-19 pandemic) are difficult to control for, especially for data with a short sample size.

Ollech (2021) categorizes the various issues with seasonality in high-frequency data that are different from—and more difficult to handle than—lower frequency data. These include, for example, uncommon periodic effects that result in multiple seasonalities in a single series (such as week-of-the-month effects).

The industry standard for seasonally adjusting lower frequency data is the X-13ARIMA-SEATS method. It uses a weighted moving average of the output of a seasonal ARIMA model. These weighted moving averages use constant periodicities of 12 (for monthly data) or 4 (for quarterly data). Thus, the X-13 is incompatible with weekly or daily data because of the first two challenges identified in the list above. Other procedures have been developed for high-frequency data, although many are relatively new and untested.

This paper is structured as follows. **Section 2** provides an overview of the five most popular and promising methods for seasonally adjusting high-frequency data. **Section 3** explains how we apply these methods to two datasets: US initial claims from the BLS and Moneris weekly transactions. **Section 4** outlines how we tune the methods and **section 5** compares the results for both datasets. **Section 6** provides concluding remarks.

2. Methods for seasonal adjustment

We consider five methods for seasonal adjustment of high-frequency data:

- year-over-year growth
- MoveReg
- fractional airline decomposition model (FAM)
- Prophet
- Multiple Seasonal-Trend decomposition using locally weighted polynomial regression, or LOESS (MSTL)

While this list is not exhaustive, these are the most popular and promising methods.

The most commonly used approach is to transform the time series into year-over-year growth rates. The aforementioned challenges imply that this approach is statistically naïve, but it represents a useful benchmark considering its ease of use. Statistical agencies are typically the authority on matters related to seasonal adjustment, however, their regular publications do not contain a lot of high-frequency data. The BLS uses MoveReg for some of its high-frequency output, but has expressed interest in migrating to new approaches, such as FAM (Evans et al. 2021). We test both methods as well as two others: MSTL and Prophet.

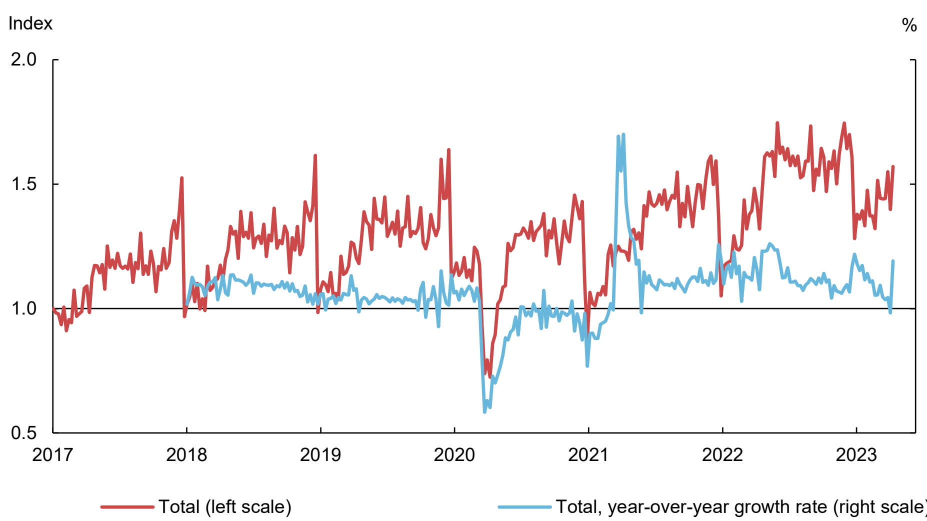
Year-over-year growth

Many articles that use high-frequency data for forecasting simply use seasonally unadjusted series or a year-over-year growth rate with manual adjustments.[1] The latter approach reduces noise and eliminates some seasonality. However, it ignores two key issues that still skew weekly data. The first is shifting holidays, which is especially difficult for holidays such as Easter, that fall on different days or months from year to year. The second issue is base-year effects, which refers to the impacts that movements in the data have on data from the corresponding period 12 months later. For example, a sharp spike appears in year-over-year growth rates in the Moneris data for 2021 (**Chart 2**). The result is more reflective of the downturn at the start of the pandemic in 2020 than with the economic circumstances of 2021. While this is an extreme example, base-year effects consistently apply to year-over-year adjustments and are often difficult to isolate.

These issues—which Statistics Canada mentions when discussing monthly data—imply that making year-over-year comparisons significantly reduces the ability to identify turning points and level shifts (Fortier and Gellatly 2024). Additionally, they further reduce sample sizes when data are available over a short sample period.

Chart 2 : Moneris spending

Weekly data, index: week of January 2, 2017 = 1



Sources: Moneris and Bank of Canada calculations

Last observation: April 8, 2023

[1] For example, Fenz and Stix (2021) apply hand-made adjustments for moving holidays and beginning-of-the-month effects.

MoveReg

MoveReg uses locally weighted regressions on sine and cosine terms, which are linked to days of the year.[2] Including more terms thus increases the frequency of cycles to capture any reoccurring seasonal patterns.

First, we conduct a regression to estimate holiday and outlier effects. Once these effects are removed, we perform a weighted regression for each year using the full sample. Weights for each year correspond to the proximity of the year being estimated. For a non-seasonally adjusted series Y, the model for the seasonal component in year *t* becomes:

$$\hat{Y}_t = X_t[X'X]^{-1}\sum_i w_i X'_i Y_i,$$

where X_t are sine and cosine variables at a daily frequency (for weekly data, we use the value on the day at the end of each week), and w_i is the weight for year *i*.

Fractional airline model

FAM is a method developed by the National Bank of Belgium as part of the JDemetra+ program (Grudkowska 2017).[3] The periodicity for FAM is not restricted to integers, which is a very helpful feature for weekly data.

This technique uses a variant of the seasonal ARIMA model of order $(0,1,1)(0,1,1)_S$. The seasonal ARIMA uses both non-seasonal and seasonal factors, where the first parameters are non-seasonal and the second are seasonal with periodicity *s*. In other words, with monthly data (and *s*=12), the seasonal ARIMA would incorporate terms from the 12th lag. The forecasting equation is as follows:

$$\hat{Y}_t = Y_{t-12} + Y_{t-1} - Y_{t-13} - \theta_1 e_{t-1} - \Theta_1 e_{t-12} + \theta_1\Theta_1 e_{t-13},$$

where $\hat{Y}_t$ is the seasonally adjusted series, Θ_1 is the MA(1) coefficient and Θ_1 is the seasonal MA(1) coefficient.

The main additional feature of FAM is that s is not required to be an integer. Instead, when using weekly data, for $s^* = s + \alpha$, where $\alpha \in [0,1]$, the 52nd and 53rd lagged terms are both included and weighted by $(1-\alpha)$ and α, respectively.

[2] We access this program using the EViews 12 application. The original program was written in FORTRAN, and more documentation can be found in Cleveland, Evans and Scott (2014).

[3] Documentation about FAM is available on Github.

Prophet

Prophet is a forecasting tool that Taylor and Letham (2017) developed for Meta, the company that owns Facebook.[4] It is a decomposable time series model with three main components: trend, seasonality and holidays. It takes the form:[5]

$$\hat{Y} = Trend * (1 + Holidays + Seasonality) + X.$$

The model can also support extra regressors as a function of time (X). By default, Prophet uses a piecewise linear regression to calculate the trend, with the user determining the number of breakpoints. Holidays and other series are modelled as independent regressors.

Seasonality is modelled through a partial Fourier sum where the order represents the number of trigonometric terms to include and thus the degree of fit for seasonality. Increasing the number of terms can quickly lead to overfitting, which is a problem even when seasonally adjusting (as opposed to Prophet's original purpose of forecasting).

Prophet contains many hyperparameters that can be tuned, such as the frequency of changes in the trend, the number of Fourier terms and additive versus multiplicative seasonality. While this flexibility helps customize the approach to any series, it leads to a tuning procedure that is computationally expensive. Further discussion on the tuning procedure is in **section 4**, and details on the hyperparameters can be found in the **Appendix**.

Multiple Seasonal-Trend decomposition using LOESS

MSTL is an additive decomposition created by Bandara, Hyndman and Bergmeir (2021).[6] The method extracts seasonal, trend and residual components from a time series. To do this, it applies local regression (LOESS) recursively.

In the LOESS algorithm, for a dependent variable *Y* at all values of an independent variable *X*, a smoothing parameter is chosen and represents the number of observations (*k*) to include in the local regression. Weights are then assigned to the nearest neighbours of *k* according to their distance to a given value of *x*. The local regression is then performed at each value of *x*. This is performed in both an outer loop to calculate robustness weights and remove outliers and an inner loop that iteratively calculates trend and seasonal terms.

MSTL extends the seasonal-trend decomposition using LOESS (STL) procedure created by Cleveland et al. (1990). The STL program is run iteratively on multiple periodicities. This allows for multiple seasonalities (yearly and within-month, for example) to be detected and estimated.

---

[4] Prophet is available in both Python and R. We use the former. For more documentation on Prophet, see the Prophet home page.

[5] Note that the default specification is additive, but multiplicative seasonality is accomplished through log transforms.

[6] See the statsmodels website for information on MSTL and STL.

The MSTL package has an option to increase robustness to outliers. The outer loop of the procedure already weighs observations based on their distance from the smoothed trend, but this additional option removes the effect of unusual outliers from the LOESS algorithm.

Literature review

While little consensus exists on the best approach, there are a few evaluations for daily and weekly data that can be used as a starting point. The creators of the MSTL program, Bandara, Hyndman and Bergmeir (2021), compare their procedure with Prophet, a seasonal-trend decomposition by regression (STR) and a trigonometric exponential smoothing state space model. The authors find that MSTL significantly outperforms the other models when used on simulated cycle and trend data. However, the degree to which each method was tuned is unclear, which we find can significantly impact the results.

Evans et al. (2021) compare MoveReg, FAM and structural time series programs. Using US initial claims data from 2009 to 2021, they find that each method gives similar results with some exceptions due to the handling of outliers. However, the authors note shortcomings with MoveReg, such as inflexibility and the fact that it does not explicitly model the time series. They also highlight that FAM is still under development but shows great promise.

3. An application: data description

We assess the effectiveness of the various seasonal adjustment methods by applying them to two datasets: initial claims data from the BLS for the United States and Moneris transaction data for Canada.

The data on US initial claims—both the non-seasonally adjusted and seasonally adjusted series—are useful because the seasonally adjusted data have been extensively reviewed.[7] And while the true value of seasonal adjustment is very difficult to assess, the adjusted series represents a useful benchmark. Initial claims represent a claim filed by an unemployed individual after separating from an employer that determines whether an individual is eligible for the US Unemployment Insurance program. Both series are at a weekly frequency between January 2010 and April 2023.

Moneris is Canada's largest provider of solutions for mobile, online and in-store payments. The Moneris dataset represents consumer spending toward different merchant groups—such as grocery retailers, financial services and hotels—and includes credit and debit payments. We use 32 series available for Canada at the national level, with each level representing a different merchant group based on merchant category codes. For example, one of the series measures the transactional dollar volume toward grocery merchants in Canada, and another series

[7] The BLS produces the seasonally adjusted claims series using the MoveReg program. The BLS employs several custom treatments, particularly during the pandemic, which is why the series serves as a benchmark even against our use of MoveReg on the same data.

measures the transactional dollar volume toward restaurant merchants. The data used for the purposes of this paper range from January 2017 to April 2023.

To be used for forecasting, the data typically require both seasonal adjustment and deflating. While the optimal order of these transformations is not obvious, seasonally adjusting first provides the benefit of interpretability.[8] In other words, the seasonality of both inflation and the Moneris series would be lumped together if we deflated the data first. However, we can more clearly break down series dynamics if we first seasonally adjust the weekly data and the consumer price index individually.

We find that both methods of computing the real seasonally adjusted series result in similar values. Given the benefit of interpretability, the rest of this paper will focus on seasonally adjusting nominal values.

4. Evaluation metrics and hyperparameter tuning

We use two metrics to evaluate the resulting seasonally adjusted series from each method and to tune the hyperparameters specifying the sensitivity and flexibility of each method:

- the QS test
- the mean squared error (MSE) compared with the X-13 approach

We also visually inspect each series because these tests sometimes overlook important dynamics that are best discovered though a manual inspection.

The QS test for residual seasonality that we use is a modified Box-Ljung test that checks for autocorrelation at a yearly lag. The intuition behind this test is that seasonality shows up in the autocorrelation of a series at some seasonal frequency (the 52nd and 53rd lag in the case of weekly data). It takes the form:

$$QS = n(n+2)\{\sum_{i=1}^{k} \frac{[max(0,\rho_{i*l})]^2}{n-i\cdot l} + \sum_{i=1}^{k} \frac{[max(0,\rho_{i*j})]^2}{n-i\cdot j}\},$$

where n is the number of observations, k is the number of lags being tested for each frequency and ρ_{i*l} and ρ_{i*j} are the autocorrelations at lags l and j, respectively.[9] We assume that the test follows an $\chi^2(2k)$ distribution.

For the MSE test, we convert the seasonally adjusted weekly values to a monthly frequency and compare it with a monthly X-13 seasonally adjusted series. Given that the X-13 approach is the industry standard for seasonally adjusting lower frequencies, we treat it as the true value. This

[8] The order of operations would not matter if the seasonal transformation and the additional data transformations (in this case, deflating the series by the overall consumer price index) are linear. However, we optimize to capture the individual seasonality of the unadjusted series, resulting in different outcomes than seasonally adjusting the deflated series. This is especially tricky given consumer price index data is released only monthly.

[9] This QS test has been further modified from the test specified in Ladiray et al. (2018) to account for both lags 52 and 53.

approach is used to compare across methods but does not identify residual seasonality. One should therefore be cautious about drawing conclusions from this test because it provides results that are ordinal but does not provide information about the degree of seasonality remaining in the adjusted series. The method with the lowest MSE compared with X-13 is considered the best according to this metric.

These evaluation metrics help determine the optimal seasonal adjustment method and the optimal hyperparameters. Prophet needed to be tuned because all but two series contained residual seasonality when using the default specifications, while MSTL and MoveReg benefitted from tuning. For each method, we calibrate the hyperparameters using a grid optimization, meaning we specify a range of values for the hyperparameters and test for all possible combinations. We then select the set of hyperparameters that perform optimally according to the two metrics. Our results show that while both metrics are informative for tuning the parameters, using the QS statistic typically results in smoother adjusted values. **Section 5** shows only the results based on tuned hyperparameters using the QS test.[10]

We do not tune the hyperparameters for the FAM given the limited options in the version we use. While some adjustments can be made, we determined they are not impactful enough to warrant tuning procedures. Instead, we use the default specifications.

Using the same metrics for tuning the hyperparameters of three of the methods and for their evaluation could lead to some confirmation bias. However, the goal of this exercise is to eliminate seasonality, which the QS metric identifies. As long as we use only the information that analysts would have in real-time, the tuning approaches are valid.

COVID-19 pandemic

One of the challenges with modelling time series data is dealing with the shock of the COVID-19 pandemic. Including pandemic data in the sample can significantly skew the results. The types of adjustments made to account for the COVID-19 shock vary depending on the seasonal adjustment method.

For example, Evans et al. (2021) identify each week during the pandemic as an outlier in the MoveReg program. The EViews application of the MoveReg program, however, limits the number of outliers that one can include. Therefore, the MoveReg results presented below include outliers for only the year 2020.

The FAM method uses a holiday matrix, inside of which are dummies that span several months at the beginning of the pandemic. After computing the seasonally adjusted series, the effect of the pandemic is added back into the series. This method is also consistent with Evans et al. (2021).

[10] We caution that the tuning approach we use is specific to the history of each series and overfitting the seasonality term of the model can present its own risks, especially as new data are received.

The MSTL has no option to manually specify outliers. The outer loop, which uses LOESS and weighting schemas to automatically smooth outliers, is meant to deal with this issue. However, pandemic effects are difficult to detect in an automated way, especially because the length of the pandemic leads to significant increases in the variance of the series, which in turn expands the tails of the distribution. Therefore, outliers need to be quite large to be detected.

To correct for the COVID-19 shock in Prophet, we use two approaches:

- COVID-19 regressor: uses the Bank of Canada's stringency index as an extra regressor
- Ex-COVID-19: excludes the COVID-19 period (March 2020 to March 2022)

The Bank's stringency index assesses the strictness of measures to slow the spread of COVID-19 and public information campaigns across provinces. Therefore, it can be used as a proxy to model the COVID-19 shock by including it as an extra regressor in Prophet.

Unlike most methods, Prophet approaches seasonality modelling as a curve-fitting exercise and does not explicitly account for the temporal structure of the data. This means that we can reliably drop observations from the series and use forecast values for the missing period. We consider the COVID-19 period that distorts seasonal effects ranges from March 2020 to March 2022 and exclude this period when training the model. We then predict the original series (including the COVID-19 period) using the forecasted seasonality. We find that the best approach for Prophet depends on the series, which is why we include both options in the tuning procedures.

Additionally, the large swings in the series caused by COVID-19 may also skew the QS test results. As a robustness check, we evaluate the optimal hyperparameters with a version of the QS test that excludes the COVID-19 period. In other words, we ignored the autocorrelation for periods compared with March 2020 to March 2022. Given some differences, we also attempt to select the set of hyperparameters based on the QS test that ignores the COVID-19 period, but a visual inspection shows only small changes in most series. The next section will therefore show the results for tuning and evaluation using the full time series from January 2017 to April 2023.

5. Evaluation of seasonal adjustment methods

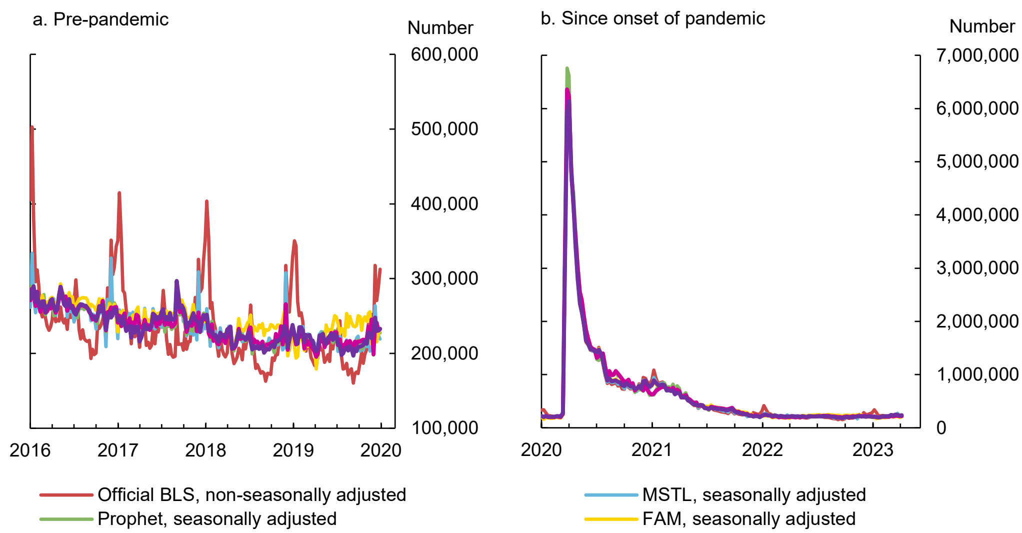
To ensure that the models are applied appropriately to the data, we perform two sets of evaluations. First, we conduct a test on weekly initial claims data, which have been seasonally adjusted by the BLS and reviewed extensively.[11] Next, we apply the methods to the Moneris data and use the described metrics to determine the success of the adjustments.

[11] For more, see the BLS website and Evans, Monsell and Sverchkov (2021).

Seasonal adjustment with initial claims data

We tune the parameters for both MSTL and Prophet over the period from 2010 to 2023 with the non-seasonally adjusted initial claims data and compare the results with the seasonally adjusted series from the BLS (**Chart 3**). We also apply the FAM and the MoveReg approaches.

Chart 3: Initial claims



Sources: US Bureau of Labor Statistics and Bank of Canada calculations

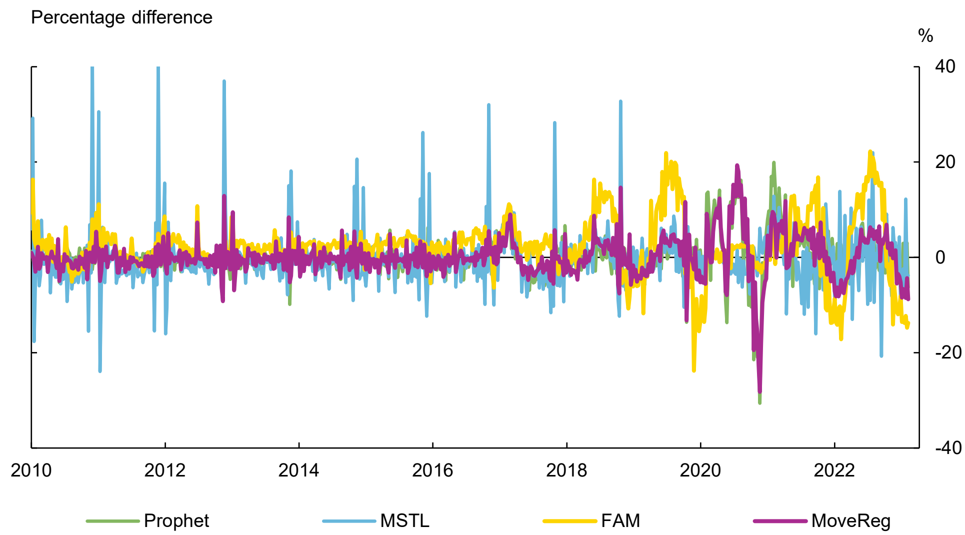
Last observation: April 8, 2023

Note: MSTL is Multiple Seasonal-Trend decomposition using locally weighted polynomial regression. FAM is fractional airline decomposition model. *Official BLS* data are from the US Bureau of Labor Statistics.

For these data, the percentage difference between Prophet and the official series is small and stable, but the two approaches handle the COVID-19 period very differently (**Chart 4**). The MSTL performs poorly and has large periodic spikes at the end of each year, indicating improper holiday treatment.[12] The FAM also deviates greatly from the official series, especially during and after the COVID-19 period.

[12] For both the MSTL and Prophet, we do not take the result that minimized the QS statistic. We use some judgment to minimize the root mean square error compared with the official seasonally adjusted series for initial claims. The QS test P-value was essentially 1 in both cases.

Chart 4: **Differences between adjustment methods and official, seasonally adjusted US initial claims data**



Sources: US Bureau of Labor Statistics and Bank of Canada calculations

Last observation: April 8, 2023

Note: MSTL is Multiple Seasonal-Trend decomposition using locally weighted polynomial regression. FAM is fractional airline decomposition model. US initial claims data is from the US Bureau of Labor Statistics.

We also compare our MoveReg specifications against the BLS results. For both Prophet and our MoveReg values, the root mean square error is much lower when the COVID-19 period is excluded (**Table 1**).

Table 1: Root mean square error of seasonal adjustment methods versus official data

	MSTL	Prophet	MoveReg	FAM
Full sample	22376	55148	37583	16849
Excluding COVID-19 period	21541	7426	8360	15514

Note: MSTL is Multiple Seasonal-Trend decomposition using locally weighted polynomial regression. FAM is fractional airline decomposition model.

The difference between MoveReg and the official series highlights the custom treatment the BLS uses for the COVID-19 shock.[13] The BLS expresses views about the treatment of the pandemic combined with additive outliers, temporary changes and level shifts that were applied in the official, adjusted series. The MoveReg program in EViews limits outliers for only 52 observations. We therefore apply extra outliers for every week during 2020.

[13] Here, we use the default values of MoveReg without any tuning procedure for consistency with the official data.

The results also highlight that Prophet aligns fairly well with the MoveReg procedure, but only when controlling for the COVID-19 period. This exercise highlights that custom treatment during the pandemic may be required for each series that is seasonally adjusted.

Lastly, the hyperparameters that minimize the QS statistic in the tuning procedure for Prophet and MSTL are less aligned with the official seasonally adjusted series than the ones we display here. We apply some judgment about the hyperparameters we choose because multiple combinations give almost identical QS results. **Table 1**, therefore, represents an upper bound on the consistency of these methods and emphasizes the need for analysts' judgment.

Seasonal adjustment of Moneris series

For the Moneris data, we apply a uniform approach of either dropping the COVID-19 period from the estimation or using an extra regressor. The extra regressor, which uses the Bank's COVID-19 stringency index, applies to varying degrees depending on the industry.[14]

For Prophet, the QS test fails to detect residual seasonality in all but one series.[15] The MoveReg procedure rejects the null hypothesis of no residual seasonality for four series, the FAM rejects the null for six series, and the MSTL rejects the null in seven series (**Table 2**).[16, 17]

Table 2: Count of sectors with residual seasonality

Confidence level	Prophet, regressor	MSTL	MoveReg	FAM
90%–95%	0	4	0	1
95%–99%	0	1	3	1
≥ 99%	1	2	1	4

Note: MSTL is Multiple Seasonal-Trend decomposition using locally weighted polynomial regression. FAM is fractional airline decomposition model. *Regressor* uses the Bank of Canada's COVID-19 stringency index.

For some series such as *professional services*, the MoveReg program does well upon visual inspection but fails the residual seasonality test (**Chart 6**). Moreover, when comparing results between MoveReg and Prophet (the latter did not fail the residual seasonality test), the superior method is difficult to discern. The MoveReg program eliminates obvious seasonal peaks and troughs but has a sharp spike in March 2021. This adjustment indicates that MoveReg expects

[14] For example, public health restrictions directly and significantly impacted travel spending, but almost certainly affected professional services to a lesser degree.

[15] For Prophet, we show only the QS test results using the COVID-19 stringency extra regressor because this option typically results in better test scores. Later in this section we discuss some instances when excluding the COVID-19 period entirely is more appropriate.

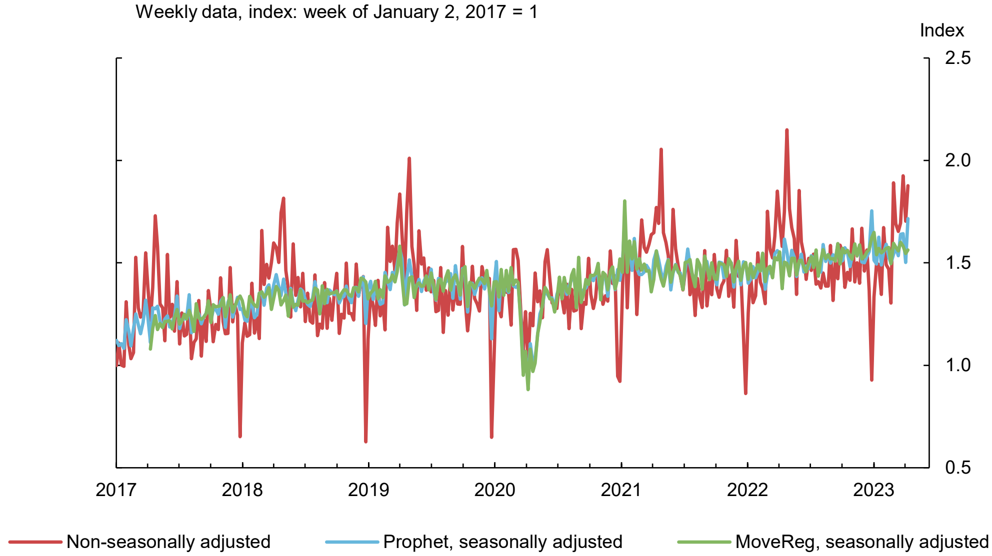
[16] We also test using year-over-year differences. The QS test detects seasonality in five series, along with the base-year effects discussed in **section 2**.

[17] No particular series is difficult to seasonally adjust across all methods, with the exception of *government, government*, where residual seasonality was found in all four procedures.

to see dynamics similar to those in the previous year, and thus incorporates the COVID-19 shock into its estimate of seasonal patterns.

Chart 5: Professional services

Weekly data, index: week of January 2, 2017 = 1



Sources: Moneris and Bank of Canada calculations

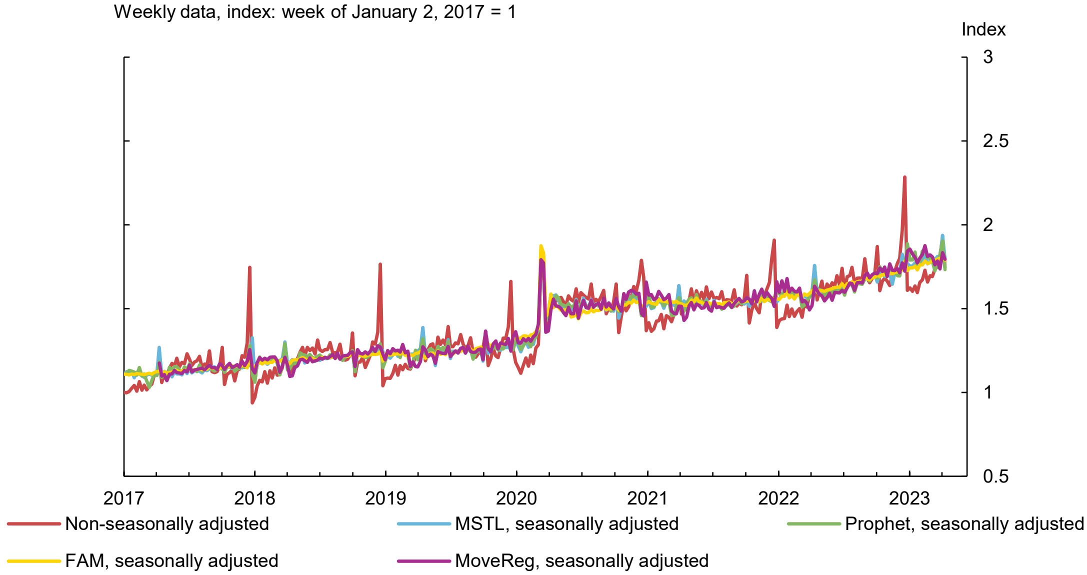
Last observation: April 8, 2023

The MSE test indicates that the MoveReg program performs the best by a large margin, with the lowest MSE for 15 series (**Table A-1**). When combining the two methods of controlling for the COVID-19 shock, Prophet outperforms on the MSE metric for 9 series, while the FAM is the best method in 6 of the 32 series. These mixed results suggest that there is likely not one clearly superior method.

It is important to note that the MSE test sometimes overlooks weekly seasonality. For example, **Table A-1** shows that the MSTL has the lowest MSE for the *retail, grocery* category, but also failed the QS test for residual seasonality. Indeed, a visual inspection of the results shows some periodic spikes in the MSTL-adjusted series at the weekly frequency (**Chart 7**).

Chart 6: Retail, grocery

Weekly data, index: week of January 2, 2017 = 1



Sources: Moneris and Bank of Canada calculations

Last observation: April 8, 2023

Note: MSTL is Multiple Seasonal-Trend decomposition using locally weighted polynomial regression. FAM is fractional airline decomposition model.

In addition to the statistical tests discussed earlier, we also visually inspect the series to assess the performance of the seasonal adjustment methods. MoveReg is consistently stable across most adjusted series, with few unexplained spikes or troughs. A stable series is not necessarily desirable, of course, especially when considering that the program failed the QS test in several instances. Additionally, the MoveReg program overcorrects for seasonality in March and April of non-pandemic years in some series, indicating that it did not always properly account for the COVID-19 outlier (**Chart 7**).

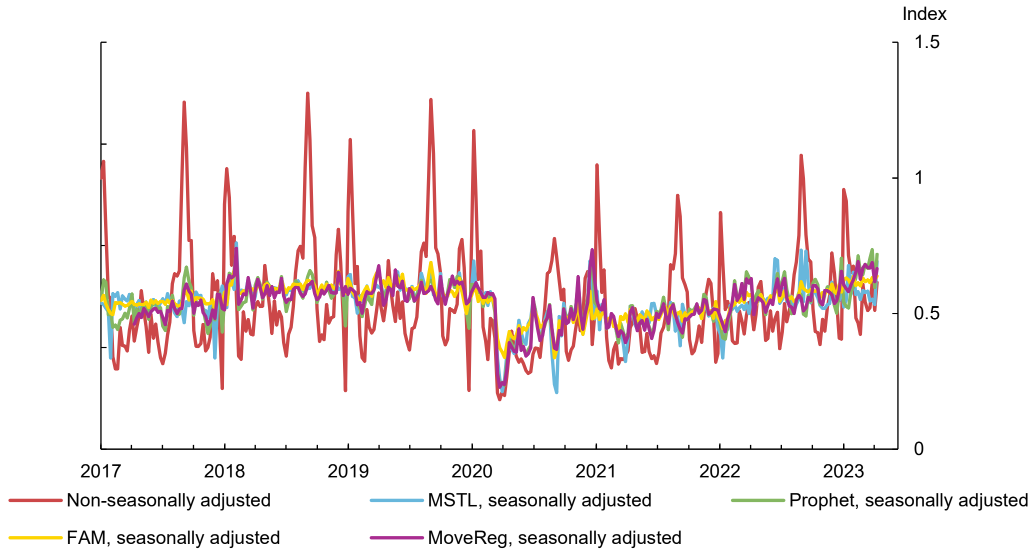
The COVID-19 regressor correction for Prophet performs best based on the QS test metric. Visually, Prophet does not seem to fully capture seasonality in multiple instances, regardless of the hyperparameters and options selected. The ex-COVID-19 correction more appropriately handles seasonality in several of these series, such as *retail, department stores* (**Chart A-1**). Even in these cases, however, there appears to be some residual seasonality in the periods before the start or after the end of the pandemic.

In most cases, the MSTL appears to have large spikes, often around March. The inability to add custom outliers in this model prevents us from effectively controlling for the COVID-19 shock, forcing us to rely on the automated outlier detection procedure. This procedure typically does well for temporary outliers, but MSTL struggles to effectively capture seasonality for large, persistent shifts like the COVID-19 pandemic and a short five-year sample.

For many series, the FAM procedure cuts through much of the seasonality and, like MoveReg, is not volatile. However, we find several instances of large unexplained spikes or troughs like with MSTL. Additionally, FAM does not appear to sufficiently adjust for the COVID-19 outlier in some cases, despite our outlier specifications. For example, **Chart 7** clearly shows that FAM adjusts in *government, schools* so that it does not decline nearly as much as with the other methods during the COVID-19 shock.

Chart 7: Government, schools

Weekly data, index: week of January 2, 2017 = 1



Sources: Moneris and Bank of Canada calculations

Last observation: April 8, 2023

Note: MSTL is Multiple Seasonal-Trend decomposition using locally weighted polynomial regression. FAM is fractional airline decomposition model.

6. Conclusion

High-frequency data have become increasingly available to economists and are potentially very useful for nowcasting the economy. But seasonally adjusting these data is a difficult problem, and standard techniques such as the X-13ARIMA-SEATS method are not suitable for high-frequency data. Non-isochronicity, moving holidays and noisy data make it difficult to develop universally successful techniques.

Of the four methods we evaluate in this paper, the MoveReg program often appears the most reliable. The default hyperparameters also perform relatively well, implying that computationally expensive tuning may not always be necessary. However, applying this approach involves two major downsides:

- MoveReg fails the QS test for residual seasonality on several series. Although MoveReg may be visually appealing, seasonality seems to remain in several cases.
- The program is currently only available in proprietary software, which can be cumbersome when creating data pipelines.

FAM often performs quite well and results in stable series that cut through much of the seasonality. However, FAM fails residual seasonality tests in several instances and is difficult to apply when not using the custom interface (JDemetra+) the creators have developed.[18]

In contrast, Prophet is available in open-source languages and successfully eliminates seasonality in all but a single series. However, finding reasonable adjustments requires an optimization procedure to tune the hyperparameters. This computationally expensive tuning procedure also results in a large set of hyperparameter combinations that pass the QS test, requiring a user to manually select optimal values. Furthermore, visually inspecting the results indicates that some seasonality may remain in a few series that the QS test does not capture.

Ultimately, no approach is perfect. The flexibility of Prophet allows analysts to further tweak individual series but ultimately relies heavily on their judgment. Prophet can be automated and applied to data pipelines, which is highly advantageous for short-term forecasting. MoveReg is a method that agencies such as the BLS use, and results in consistently smooth adjusted series. However, depending on the evaluation method, no high-frequency seasonal adjustment technique removes the presence of seasonality in all series. Further advances in methods such as FAM will be highly anticipated as high-frequency data become more readily available. For now, using any of these methods should be done with care, and comparisons and testing are advised.

[18] We applied this method in R, rather than the developed proprietary software.

Appendix

Table A-1: Mean squared error and QS test results

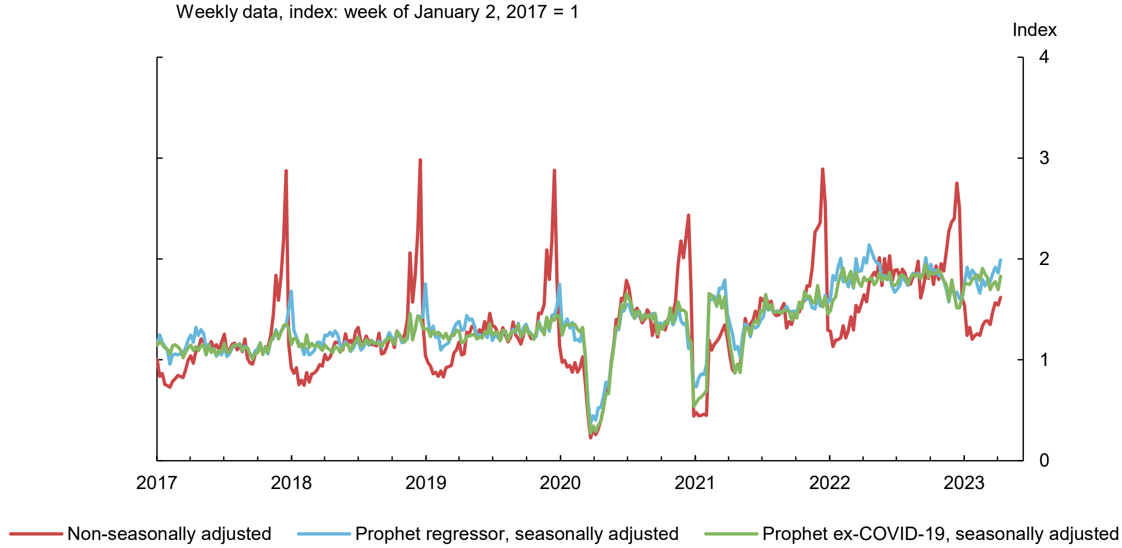
Series	Lowest mean squared error	Highest mean squared error	QS test—residual seasonality
Total	FAM	MSTL	
Government, government	MoveReg	Prophet ex-COVID-19	All methods
Government, other	MoveReg	MSTL	
Government, schools	MoveReg	MSTL	MSTL, MoveReg
Retail, alcohol	Prophet, regressor	MSTL	
Retail, apparel	MoveReg	MSTL	
Retail, business-to-business	Prophet, regressor	MSTL	Prophet ex-COVID-19, MoveReg
Retail, department stores	FAM	MSTL	
Retail, direct marketing	MoveReg	MSTL	Prophet ex-COVID-19
Retail, drug stores	MoveReg	Prophet regressor	
Retail, entertainment	FAM	MSTL	MSTL
Retail, gas and convenience	FAM	MSTL	MSTL
Retail, grocery	Prophet, regressor	MoveReg	
Retail, household	MSTL	Prophet regressor	
Retail, mass merchandisers	Prophet, ex-COVID-19	FAM	Prophet ex-COVID-19, FAM
Retail, other	MoveReg	MSTL	FAM
Retail, specialty	Prophet, ex-COVID-19	MSTL	
Retail, vehicle	MoveReg	MSTL	
Services, financial	MoveReg	MSTL	Prophet ex-COVID-19
Services, health	MSTL	FAM	
Services, insurance	Prophet, regressor	MSTL	Prophet ex-COVID-19, MoveReg, FAM
Services, other	Prophet, ex-COVID-19	MSTL	
Services, professional service[s]	MoveReg	MSTL	MoveReg
Services, trades	MoveReg	Prophet regressor	Prophet ex-COVID-19
Services, Utilities	MoveReg	MSTL	
Travel and Entertainment, airlines	MoveReg	MSTL	FAM
Travel and Entertainment, car rentals	MoveReg	MSTL	
Travel and Entertainment, hotels	FAM	MoveReg	MSTL

Travel and Entertainment, restaurant	Prophet, regressor	Prophet ex-COVID-19	MSTL
Travel and Entertainment, trains	FAM	MSTL	
Travel and Entertainment, transportation	Prophet, regressor	MSTL	FAM
Travel and Entertainment, travel	MoveReg	MSTL	MSTL

Note: MSTL is Multiple Seasonal-Trend decomposition using locally weighted polynomial regression. FAM is fractional airline decomposition model. *Regressor* uses the Bank of Canada's COVID-19 stringency index.

Chart A-1: Retail, department stores

Weekly data, index: week of January 2, 2017 = 1



Sources: Moneris and Bank of Canada calculations

Last observation: April 8, 2023

Prophet hyperparameters

Prophet contains many hyperparameters that can be fine-tuned, some of which we describe below.

- Seasonality in Prophet is estimated using a partial Fourier sum. The Fourier order refers to the number of terms in the partial sum. The *yearly seasonality* hyperparameter sets the Fourier order of the curve and **by default is assigned a value of 10**. By increasing the Fourier order, one can achieve a better fit (more flexibility) but at the risk of overfitting.
- *Seasonality_mode* refers to modelling the seasonality either as multiplicative or additive (default). Multiplicative seasonality results in a better fit when a growing pattern is observed over time in the data. In contrast, modelling the seasonality as additive would be ideal if the spread of the data is constant.
- *Seasonality_prior_scale* is a parameter that focuses on regularization, and by **default is assigned a value of 10**. Reducing the value of this parameter reduces the amount of variation in the seasonality curve. Whereas, reducing Fourier order controls the seasonality curve by reducing the number of bends that it is allowed to take.

- *Changepoint_prior_scale* is a parameter that focuses on regularizing the changepoints and **by default is assigned a value of 0.05**. Increasing this parameter will make the model more flexible and decreasing it will make the model less flexible.
- *Holidays_prior_scale* like all the other prior scales focuses on regularization but with respect to holidays. This parameter is **by default assigned a value of 10** and is responsible for controlling the flexibility of holidays.

MSTL hyperparameters

Some of the hyperparameters for the MSTL algorithm are:

- *Period* specifies the periodicity of the data. The MSTL allows for multiple seasonalities.
- *Windows* sets the length of the seasonal smoothers for each corresponding period. The number of observations is included in the LOESS for seasonality.
- *Trend* sets the length of the trend smoother. The number of observations is included in the LOESS for the trend.
- *Low pass* sets the length of the low pass filter and cuts off signals with a frequency lower than what is specified.

References

Aastveit, K. A., T. M. Fastbø, E. Granziera, K. S. Paulsen and K. N. Torstensen. 2020. "Nowcasting Norwegian Household Consumption with Debit Card Transaction Data." Norges Bank Working Paper No. 17/2020. https://www.econstor.eu/handle/10419/246114

Bandara, K., R. J. Hyndman, and C. Bergmeir. 2021. "MSTL: A Seasonal-Trend Decomposition Algorithm for Time Series with Multiple Seasonal Patterns." arXiv preprint, arXiv:2107.13462. https://doi.org/10.48550/arXiv.2107.13462

Cleveland, R. B., W.S. Cleveland, J. E. McRae and I. Terpenning. 1990. "STL: A Seasonal-Trend Decomposition Procedure Based on Loess." *Journal of Official Statistics* 6 (1): 3–73. https://www.proquest.com/docview/1266805989

Cleveland, W., T. Evans and S. Scott. 2014. "Weekly Seasonal Adjustment — A Locally-weighted Regression Approach." US Bureau of Labor Statistics Working Paper No. 473. https://www.bls.gov/osmr/research-papers/2014/ec140040.htm

Evans, T., B. Monsell and M. Sverchkov. 2021. "Review of Available Programs for Seasonal Adjustment of Weekly Data." Paper prepared for the Joint Statistical Meetings, August 8–12. https://www.bls.gov/osmr/research-papers/2021/st210020.htm

Fenz, G. and H. Stix. 2021. "Monitoring the Economy In Real Time with the Weekly OeNB GDP Indicator: Background, Experience and Outlook." *Monetary Policy & The Economy* Q4/20–Q1/21: 17–40. https://www.oenb.at/dam/jcr:805f2314-6a73-4467-aebe-3267c551276f/03_mop_Q4_20-Q1_21_Monitoring-the-economy-in-real-time.pdf

Fortier, S. and G. Gellatly. 2024. "Seasonally adjusted data – Frequently asked questions." June 6. https://www.statcan.gc.ca/en/dai/btd/sad-faq

Grudkowska, S. 2017. *JDemetra+ User Guide Version 2.2*. Narodowy Bank Polski.

Ladiray, D., J. Palate, G. Mazzi and T. Proietti. 2018. "Seasonal Adjustment of Daily Data." Paper prepared for the 16th Conference of IAOS, September 19–21. https://docslib.org/doc/4439102/seasonal-adjustment-of-daily-data

Lewis, D. J., K. Mertens, J. H. Stock and M. Trivedi. 2021. "Measuring Real Activity Using a Weekly Economic Index." *Journal of Applied Econometrics* 37 (4): 667–687. https://doi.org/10.1002/jae.2873

Monteforte, L. V. Raponi. 2019. "Short-Term Forecasts of Economic Activity: Are Fortnightly Factors Useful?" *Journal of Forecasting* 38 (3): 207–221. https://doi.org/10.1002/for.2565

Ollech, D. 2021. "Economic Analysis Using Higher Frequency Time Series: Challenges for Seasonal Adjustment." Deutsche Bundesbank Discussion Paper No. 53/2021. https://doi.org/10.2139/ssrn.4025684