Challenges and Recent Developments in the Seasonal Adjustment of Daily Time Series

Karsten Webel*

Abstract

Daily time series have increasingly appeared on the radar of official statistics in recent years, mostly as a consequence of the exploration of new digital data sources for information that could be used to augment established forecasting models for headline indicators, such as quarterly GDP. Many of these daily series are seasonal and thus in need for seasonal adjustment. However, traditional methods in official statistics often fail to model and seasonally adjust them appropriately. The main reason is that granular daily data typically exhibit features that are not observable in monthly and quarterly data. Prime examples include irregular spacing, coexistence of multiple seasonal patterns with integer versus non-integer seasonal periods and potential cross-dependencies as well as small sample issues, such as missing data. We provide an overview of recent modeling and seasonal adjustment approaches that are capable of handling these distinctive feature, or at least some of them, and illustrate selected methods using daily realized electricity consumption in Germany. Special attention is paid to the extended X-11 and ARIMA model-based approaches and to structural time series models as implemented in a preliminary version of JDemetra+ 3.0 that is accessible via R.

Key Words: High-frequency data, JDemetra+, seasonality, signal extraction, time series decomposition

1. Introduction

Official statistics have been concerned with the collection, seasonal adjustment and dissemination of monthly and quarterly time series for a long time. However, the recent emergence of new information technologies, digital data sources and automated data collection methods has led to an increased availability of daily and other types of high-frequency time series observed at intra-monthly intervals. Although these time series typically have a rather short history and sometimes a rather experimental character, several feasibility studies have explored ways of using them to complement established low-frequency surveys in Germany by augmenting the calculation of early estimates and/or improving timeliness. For example, Askitas and Zimmermann (2011) condense real-time truck toll data recorded by the German GPS-based toll collection system into a monthly index and show that this series is a good early indicator of the monthly German production index, which itself is a well-known leading indicator of quarterly gross domestic product (GDP). More recently, Dickopf et al. (2019) integrate daily energy production in Germany into a GDP nowcasting model that provides first estimates already 10 days after the end of a reference quarter.

The recent outbreak of the COVID-19 pandemic additionally fueled the demand of many data users for more timely, more frequent and more granular data to monitor the disruptions to virtually all economic sectors, especially public health and traffic as well as the labor market. In Germany, a weekly activity index (WAI) has been established that essentially merges information from selected daily and weekly indicators as well as monthly industrial production and quarterly GDP (Deutsche Bundesbank, 2020). In its current composition, the WAI contains six daily indicators (air pollution as measured by concentration of nitrogen dioxide, consumer confidence, number of pedestrians in selected shopping

*Deutsche Bundesbank, Central Office, Directorate General Statistics and Research Centre, Wilhelm-Epstein-Strasse 14, 60431 Frankfurt am Main, Germany. The views expressed in this paper are those of the author and do not necessarily represent those of the Deutsche Bundesbank or the Eurosystem.

streets in major German cities, realized electricity consumption, truck toll mileage index, worldwide numbers of passenger and cargo flights) and three weekly Google trends series (search terms "short-time work", "state support" and "unemployment").[1] Most of these high-frequency time series contain seasonal fluctuations that need to be removed prior to the WAI calculation.

In general, however, many daily and weekly economic time series exhibit stylized facts that are often also inherent in but usually not directly measurable from monthly and quarterly data, making their modeling and seasonal adjustment much more challenging. The main objective of this paper is thus to provide a brief overview of common features of daily data (Section 2) and a review of some recent developments in the modeling and seasonal adjustment of such data (Section 3), being somewhat similar in spirit to Findley (2005). The approaches currently implemented in the JDemetra+ (JD+) software package are then applied to daily realized electricity consumption in Germany for the purpose of illustration (Section 4), followed by some final remarks and conclusions (Section 5).

2. Data Peculiarities

The stylized facts of daily economic time series are first exemplified (Section 2.1) and then discussed from a more general perspective (Section 2.2). Afterwards, the key challenges in modeling daily data are summarized (Section 2.3).

2.1 A Motivating WAI Example

Figure 1 reveals three facets of the seasonal profile of daily electricity consumption in Germany, measured at terawatt hours (TWh). Panel (a) shows that electricity consumption basically is higher in the winter and lower in the summer. Given usual temperature curves and daylight hours in Germany, this pronounced U-shaped day-of-the-year (DOY) pattern does not come as a surprise. However, it is interrupted every year by a deep spike trough between Christmas and New Year and clearly affected by the start of the COVID-19 lockdown on 23 March 2020, which also breaks the slight upward trend in the series.

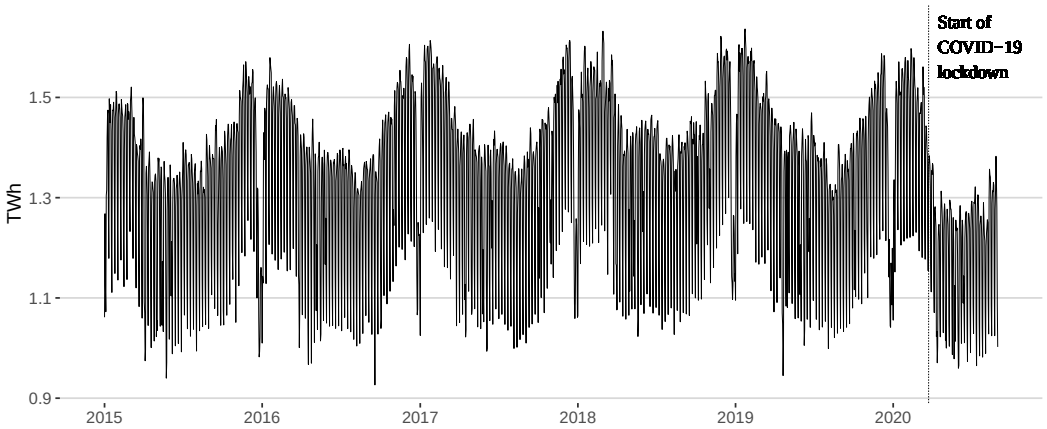
Panel (b) depicts the intra-monthly dynamics with the aid of day-of-the-month (DOM) boxplots. Although tending to be somewhat lower around the turn of the month and showing some signs of time-varying volatility, electricity consumption has an almost constant median level and an almost symmetric distribution for the majority of days and thus does not exhibit a very distinct DOM pattern.

Panel (c) zooms in on the most recent observations, revealing a persistent day-of-the-week (DOW) pattern. Electricity consumption is relatively high from Monday until Friday, mostly because of commercial consumers being active. Then it drops on Saturday and even more so on Sunday, where the share of private consumption is usually highest. However, this stable pattern is out of play from Christmas until New Year and also visibly affected by both fixed and moving holidays related to the German national calendar. Also, its volatility seems to decrease in the wake of the COVID-19 lockdown.

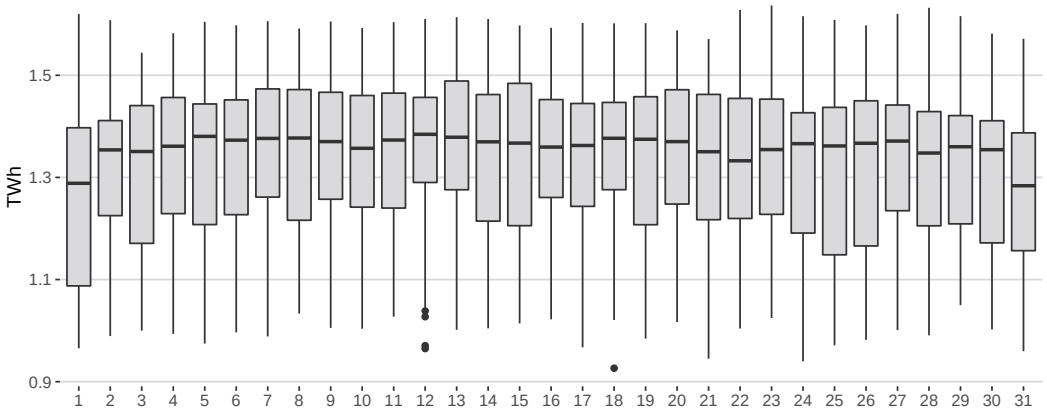
Figure 2 provides further insights into the volatility dynamics of the DOW pattern. The grouped DOW boxplots shown in Panel (a) indicate that the DOW volatility varies with the level of the DOY pattern as it is lower in the summer (group "April to September"), especially on weekends, and higher in the winter (group "October to March"), while remaining

[1]The WAI is still work in progress and therefore constantly undergoing revision and, if necessary, adjustment. Further methodological information as well as Excel data sheets can be obtained from `www.bundesbank.de/wai`.

(a) Day-of-the-year (DOY) pattern: 1 January 2015 to 30 August 2020



(b) Day-of-the-month (DOM) pattern: 1 January 2015 to 30 August 2020



(c) Day-of-the-week (DOW) pattern: 1 December 2019 to 30 August 2020

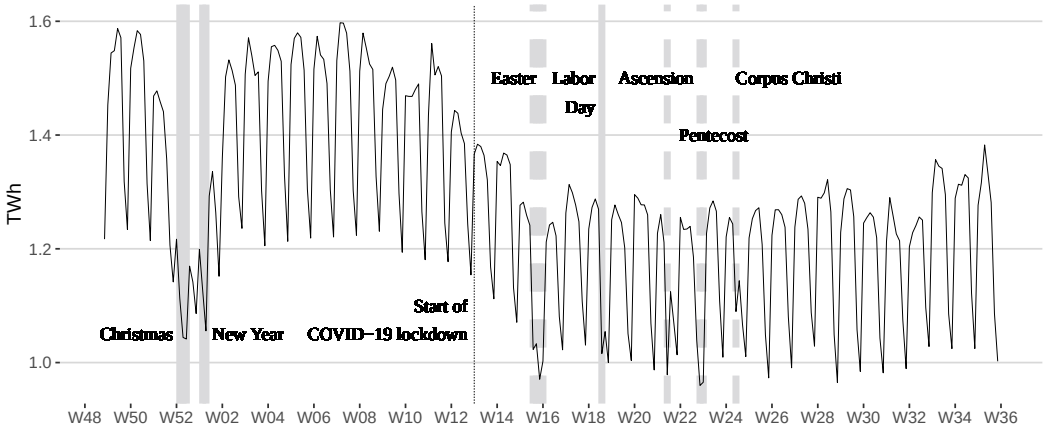


Figure 1: Seasonal profile of daily realized electricity consumption in Germany. Gray verticals in Panel (c) correspond to fixed (*solid*) and moving (*dashed*) holidays.

the pattern's characteristic shape. Note that these joint dynamics violate the orthogonality assumption about unobserved components that is often the default in the seasonal adjustment approaches applied in official statistics.

In addition, DOW volatility received a noticeable exogenous shock from the start of the COVID-19 lockdown. Panel (b) shows that the weekly ranges of electricity consumption lie consistently between 0.3 and 0.4 TWh, except for some dips around the turn of the year and some peaks in first half of the year, but sharply drop to 0.3 TWh and even below in 2020 W13. The weekly standard deviations also decrease visibly from their usual level between 0.12 and 0.16 TWh, albeit more gradually. Both measures, however, regain their normal level towards the end of the observation span.

The reason for the decline of DOW volatility is exemplified in Panel (c), using the weekly ranges. Both the weekly minima and maxima diminish as a result of the COVID-19 lockdown. The former, which usually occur on Sunday, decrease to around 1.0 TWh, a level they have reached multiple times in past years, especially during the springs of 2015 and 2016. However, the weekly maxima, which usually occur between Monday and Friday, experience a sudden drop to an unprecedented spring level of around 1.3 TWh. This is most likely a result of the national counteractions against the regional spread of the COVID-19 pandemic, such as temporary closing of production places and increased use of short-time work and teleworking, which in sum shifted a major share of consumption from the commercial to the private sector. Panel (c) also shows that the weekly means are smaller than weekly medians almost the entire time, highlighting again the persistent left-skewed nature of the DOW pattern.

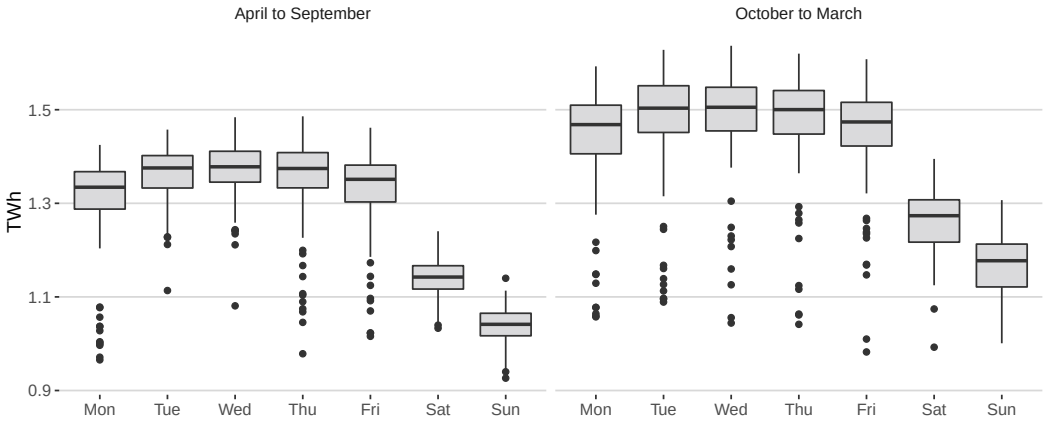
Overall, daily electricity consumption contains a complex seasonal profile that is clearly dominated by the DOW and DOY patterns, whereas the contribution of the DOM pattern is relatively small. In addition, calendar effects of both fixed and moving holidays are directly observable due to data granularity.

2.2 General Data Issues

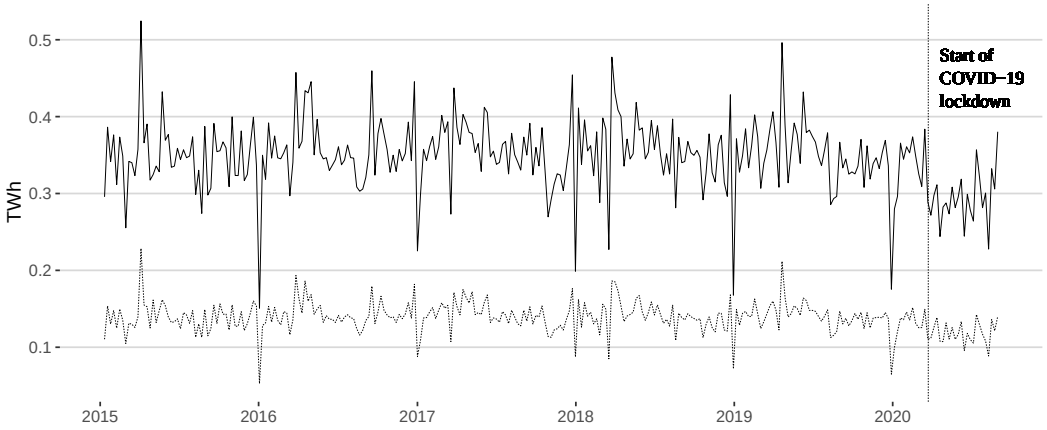
The previous example already highlighted some characteristic features of daily data with emphasis on seasonal dynamics. It also demonstrated that such data is likely to contain structures that usually are either unobservable or assumable yet immeasurable at the monthly and quarterly scales and thus may complicate modeling. This section therefore categorizes and discusses key common properties of daily data from a broader perspective, taking into account related research conducted by Cabrero et al. (2009), Cox et al. (2020), De Livera et al. (2011), Hyndman and Fan (2010), Koopman and Ooms (2003, 2006), Koopman et al. (2007), Ladiray et al. (2018), McElroy and Monsell (2017), McElroy et al. (2018), Ollech (2018) and Weinberg et al. (2007), amongst others.

Unadjusted Data The number of days naturally varies across months, quarters and years in most calendars and so does the number of observations in daily time series, leading to irregularly spaced data. This effect is not limited to equidistant time series as it also shows up quite often in monetary daily data recorded on non-equidistant banking days. Daily time series may also suffer from various small sample issues, such as missing, zero or close-to-zero observations and high sensitivity to outliers. The latter is not restricted to types well-known from monthly and quarterly data, such as short-lived additive or pulse outliers and long-lasting level breaks. Instead, daily time series are also prone to contain new and more complex outlier types with, for example, pyramidal or wavy shapes. All these issues potentially add up to irregular movements that are much more volatile and therefore at higher risk of heteroskedasticity compared to monthly and quarterly data.

(a) Time-variation: 1 January 2015 to 30 August 2020



(b) Sensitivity to exogenous shocks (part I): 2015 W02 to 2020 W35



(c) Sensitivity to exogenous shocks (part II): 2015 W02 to 2020 W35

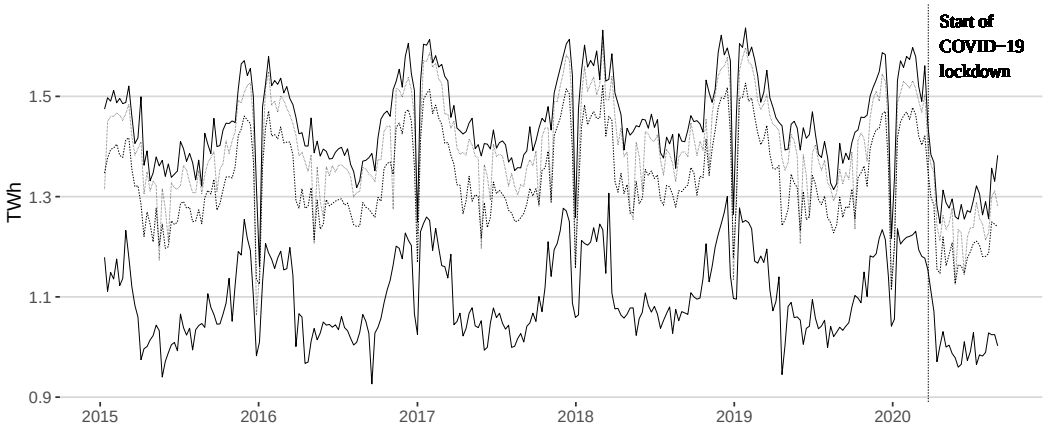


Figure 2: DOW volatility aspects of daily realized electricity consumption in Germany. Lines in Panel (b) correspond to weekly ranges (*solid*) and standard deviations (*dashed*). Lines in Panel (c) correspond to weekly maxima (*top solid*), medians (*dotted*), means (*dashed*) and minima (*bottom solid*).

Seasonal Variation The seasonal dynamics of daily time series are often generated by a set of superimposed seasonal patterns with varying degrees of complexity. Each seasonal pattern is composed of several seasonal cycles associated with the pattern's chief seasonal harmonic and its higher frequency aliases.[2] For example, the DOW pattern contains three seasonal cycles associated with movements that occur once, twice and thrice a week.

The seasonal patterns are directly affected by irregular spacing in two ways. First, each seasonal period can be an integer or non-integer. For example, the DOW pattern always consists of 7 days, whereas the DOM and DOY patterns have long-term average durations of 30.4369 and 365.2425 days in the 400-year cycle of the Gregorian calendar. Second, as a consequence, the chief harmonics of the seasonal patterns fail to be integer multiples of each other and therefore the daily patterns (or, more precisely, the seasonal cycles of different seasonal patterns) are not nested. Note that the underlying varying-length-of-period effect is also inherent in monthly and quarterly data but removed as part of the calendar variation.

Overall, the seasonal profile of daily time series is highly complex due to the coexistence of multiple seasonal patterns. This complexity also increases the chances of cross-dependencies between a seasonal pattern or cycle and any other unobserved component, including other seasonal patterns.

Calendar Variation The granularity of daily time series enables direct observation and measurement of calendar-related dynamics with pinpoint accuracy. This applies not only to the effects of fixed and moving holidays, that are typically measured indirectly in monthly and quarterly data, but also to the anticipatory and catch-up effects of related events, such as bridging days. On the flip side, the granularity of daily data may pose a potential source of interaction with the seasonal profile in general and the DOW pattern in particular. The reason is that the effects of fixed holidays may depend on the particular days of the week which the holidays fall onto. For example, Christmas effects may be noticeably different for 24 to 26 December falling onto Tuesday through Thursday versus Friday through Sunday.

The calendar-related dynamics may become even more complicated in countries where secular and religious activities mainly follow different calendars. Economic time series, especially those primarily driven by cultural and social events, such as private consumption, may then be affected by the Gregorian calendar and, for example, the Hijri calendar, which is approximately 11 days shorter. Such dual-calendar effects can also be created by the Chinese, Hindu and Jewish calendars and are likely to introduce additional non-nested seasonal patterns into the data.

2.3 Challenges In Model Building

The peculiarities highlighted in the previous sections make it extremely difficult to model daily data, especially with approaches well-established in official statistics for monthly and quarterly time series. This section therefore categorizes key steps in model building and elaborates the challenges of modeling daily data. It also mentions solutions on occasion, bearing in mind that these are usually not generic but depend on the particular problem at hand.

Data Regularization The varying number of daily observations per month and quarter raises the question if the data should be regularized prior to modeling, especially when the

[2]The reasoning behind this wording is formally explained by Equations (2) and (3). Note that it also applies to monthly and quarterly time series as the latter, for example, contain a single yearly seasonal pattern made of two seasonal cycles with two- and four-quarter durations.

modeling approach is incapable of dealing with non-integer seasonal periods. For example, Ollech (2018) uses cubic splines to stretch daily currency in circulation to 31 observations in each month, if necessary. Similarly, Koopman and Ooms (2003, 2006) incorporate a time transformation into a structural time series model for daily Dutch tax revenues and introduce artificial mid-month missing values in order to obtain 23 banking days in each month. The missing values are then imputed straightforwardly by the Kalman filter and smoother during model estimation.

If missing values occur rather occasionally – for example, as a result of some deficiency in the reporting process – simpler imputation methods may be sufficient. Examples include carrying the last observation forward and linear interpolation.

Model Calibration Besides the usual suspects regarding appropriate specification of unobserved components, the correct measurement of seasonal and especially calendar-induced dynamics may pose additional difficulties for two reasons. First, finding (and testing) appropriate regression variables for fixed and moving holidays including related special days, such as bridging days, or longer pre- and post-holiday phases of interest can quickly become a time-consuming task even when working only with dummy variables. However, more nuanced regression variables are likely to improve modeling performance and thus should be checked as well. Candidates include pyramidal variables that capture linearly increasing and decreasing holidays effects, potentially with different slopes, and non-linear versions with time-varying slopes. Note that a similar comment applies to outlier specification.

Second, calendar variation that originates from fixed holidays is closely entangled with the seasonal DOY pattern. Disentangling the two dynamics once again requires appropriate specification of either component. However, when aiming at seasonal adjustment, it could also be postulated that such calendar variations belong to those calendar-induced effects that are already captured by the DOY pattern. In this case, they will be extracted by the respective seasonal filter at a later stage and therefore tailored regression variables are not needed in the model.

Overall, the granularity of daily data enforces larger models with more parameters compared to monthly and quarterly data. This directly raises the question of how to balance model accuracy against model parsimony and flexibility. In this regard, it should also be kept in mind that larger models bear a higher risk of violating the orthogonality assumption that is often imposed on unobserved components by default. Given that models with non-orthogonal components tend to be even more complex, the willingness to sacrifice some accuracy seems inevitable.

Model Estimation Given model complexity in general and the coexistence of multiple seasonal patterns alongside other granular dynamics in particular, the benefits of sequential versus simultaneous estimation of unobserved components should be weighed up. Either way, another separability issue arises from the fact that annual seasonality can be easily confused with trend behavior in daily data since the DOY chief harmonic $\omega_{\text{DOY}} = 2\pi/365.2425 \approx 0.017$ is very close to the trend frequency $\omega = 0$. Utilizing the concept of canonical atomic models, McElroy and Monsell (2017) and McElroy et al. (2018) advocate the application of the Hodrick-Prescott filter to the combined component in order to disentangle the two dynamics, which, however, is difficult even in fairly long time series.

Data availability brings up another dilemma. On the one hand, only a short history, such as a few years, of daily observations is usually available nowadays but a long history is needed to estimate reliably all effects, especially cross-dependencies. On the other hand,

a long data history may complicate likelihood evaluation and make it even infeasible in practice for some model-based approaches.

Model Stability The estimated parameters of a given model usually undergo revisions if the model is re-estimated with updated data. These revisions can be quite large even if only a few additional daily observations have become available. To foster model stability in the long run, the inclusion of borderline significant effects should be avoided. This applies especially to outliers and calendar regression variables. For that reason, default critical values and thresholds employed, for example, in automatic outlier detection and routine significance tests should be suited to daily data, as the established defaults for monthly and quarterly time series are likely to be inappropriate.

3. Seasonal Adjustment Methods

Virtually all approaches designed for the seasonal adjustment of monthly and quarterly data and used routinely in official statistics assume orthogonality of unobserved components and the existence of a single seasonal pattern with an integer period. As a consequence, they are not "ready for action" when working with daily data. To overcome the mismatch between established default assumptions and daily data dynamics, the established approaches could be modified accordingly or new approaches could be considered.

This section provides overviews of either strand of research. A general univariate modeling framework and some basic terminology is introduced first (Section 3.1). Afterwards, current JD+ implementations of modified X-11 and ARIMA model-based (AMB) approaches as well as structural time series (STS) models are presented (Section 3.2), followed by a selection of new approaches that are probably applied less frequently in official statistics (Section 3.3).

3.1 A General Additive Model

Let $\{y_t\}$ denote a daily time series and assume that it can be decomposed additively, maybe after taking logs, into latent components according to

$$y_t = t_t + s_t + c_t + i_t. \tag{1}$$

The sequence $\{t_t\}$ covers trend behavior and periodic cyclical movements with a minimum duration of one year. The two dynamics can form either a joint component, the trend-cycle, or two separate components. The sequences $\{s_t\}$ and $\{c_t\}$ capture seasonal and calendar variation and $\{i_t\}$ is the irregular component that absorbs stationary transient intra-year fluctuations and will be white noise or some low-order MA process in most cases.

To further portray the seasonal component, we distinguish between three types of seasonal dynamics:

- *Stable seasonality* refers to seasonal behavior that can be represented by a strictly periodic function of time.
- *Moving seasonality* refers to seasonal behavior that is characterized by rather gradual changes in seasonal amplitude and/or phase over time.
- *Constrained seasonality* refers to rather abrupt and potentially short-lived changes in stable and/or moving seasonality over time.

This nomenclature allows the seasonal component to be further decomposed into a set of additive seasonal patterns according to

$$s_t = \sum_{\tau_i \in \mathcal{S}} s_t^{(\tau_i)}, \tag{2}$$

where $\mathcal{S}$ is a general pattern index set, such as $\mathcal{S} = \{7, 30.4369, 365.2425\}$, and $\{s_t^{(\tau_i)}\}$ is the seasonal pattern that captures stable, moving and constrained seasonality associated with the seasonal period $\tau_i \in \mathcal{S}$.

Each seasonal pattern in (2) is associated with a set of seasonal frequencies $\lambda_j^{(\tau_i)} = 2\pi j/\tau_i$ for $j \in \{1, \ldots, \lfloor \tau_i/2 \rfloor\}$, where $\lfloor x \rfloor$ is the largest integer not exceeding x. It can thus be further atomized according to

$$s_t^{(\tau_i)} = \sum_j s_{j,t}^{(\tau_i)}, \tag{3}$$

where $\{s_{j,t}^{(\tau_i)}\}$ is the periodic swing that is generated by $\lambda_j^{(\tau_i)}$ and therefore occurs j times during a period of τ_i days. Since these atomic swings are technically cycles, they will be referred to as seasonal cycles.

Finally, we briefly illustrate the nomenclature of seasonal dynamics. An unconstrained stable seasonal pattern can be represented by (3) with the j-th seasonal cycles given by

$$s_{j,t}^{(\tau_i)} = \alpha_j \cos(2\pi jt/\tau_i) + \beta_j \sin(2\pi jt/\tau_i). \tag{4}$$

In this case, unconstrained moving seasonality can be added by introducing time-variation into the amplitudes α_j and β_j, e.g. by specifying $\alpha_{j,t}$ and $\beta_{j,t}$ as independent random walks, or by switching to a stochastic trigonometric representation. Constrained seasonality can be incorporated by decomposing the seasonal pattern, or selected seasonal cycles, according to

$$s_t^{(\tau_i)} = s_t^{(\tau_i)} \times \mathbb{1}_{\mathcal{C}}(t) + s_t^{(\tau_i)} \times \mathbb{1}_{\bar{\mathcal{C}}}(t), \tag{5}$$

where $\mathcal{C}$ is the set of dates where the constraint is in effect, $\mathbb{1}_{\mathcal{C}}(\cdot)$ is the indicator function over this set, that is $\mathbb{1}_{\mathcal{C}}(t) = 1$ if $t \in \mathcal{C}$ and zero otherwise, and $\bar{\mathcal{C}}$ is the complement of $\mathcal{C}$. The specification of $\mathcal{C}$ may be driven by exogenous shocks or the endogenous dynamics of other unobserved components, such as seasonal cycles belonging to other seasonal patterns and the calendar component.

3.2 JD+ Approaches

JD+ implements the X-11 and AMB approaches to the seasonal adjustment of monthly and quarterly time series, which can be accessed via the graphical user interface. However, the Java source code also contains modified routines for daily (and weekly) data, including a structural time series (STS) modeling framework. These routines are accessible via the {`rjdhf`} and {`rjdssf`} R packages and some methodological details are described by Ladiray et al. (2018).[3]

[3]The R packages can be downloaded from `https://github.com/palatej/rjdhighfreq` and `https://github.com/nbbrd/rjdssf`. Alternative implementations of STS models are described by Commandeur et al. (2011).

3.2.1 Extended X-11 Approach

The modified X-11 approach sticks to the basic principle of the genuine X-11 method (Shiskin et al., 1967). Being essentially a 9-step procedure, it sequentially applies linear trend and seasonal filters to the linearized observations in order to extract estimates of the trend-cyclical, seasonal and irregular components in (1). To this end, the extended X-11 approach relies on well-known tools, such as symmetric $m \times n$ moving averages, Henderson filters, $3 \times k$ seasonal filters, asymmetric Musgrave variants and extreme value detection based on lower and upper σ-limits. However, it differs noticeably from the genuine X-11 method in three respects.

The first key difference is the addition of various trend extraction filters derived from reproducing kernel Hilbert spaces (RKHS). Basically, such a Hilbert space is characterized by a kernel that reproduces every function of the space via an inner product. The RKHS methodology thus enables the grouping of linear filters into hierarchies with respect to polynomial degree and fidelity and smoothing criteria. Dagum and Bianconcini (2008, 2015) derive RKHS representations of the Henderson and Musgrave trend filters and show that they outperform the original filters in terms of signal passing, noise suppression and size of revisions. Dagum and Bianconcini (2016) discuss this topic in the broader context of real-time trend-cycle estimation.

The second key difference is enforced by the coexistence of multiple seasonal patterns in (2). Essentially, a global seasonal-irregular component is obtained after detrending the linearized observations. This component is then split into local seasonal-irregular components, using simple moving averages of lengths dictated by the seasonal periods $\tau_i \in \mathcal{S}$. Finally, estimates of the seasonal patterns in (2) are obtained from smoothing the local seasonal-irregular components.[4] To this end, the specified $3 \times k$ seasonal filter with $k \in \{1, 3, 5, 9, 15\}$ is applied if $\tau_i \in \mathbb{N}$ and a hybrid seasonal filter is automatically selected if $\tau_i \notin \mathbb{N}$. In the latter case, the selection is essentially based on the product of τ_i and half the length of the specified filter.

The third key difference is triggered by irregular spacing. Recall that the actual number of observations per seasonal period varies over time whenever $\tau_i \notin \mathbb{N}$, leading to ragged right edges in the famous X-11 tables and thus rendering column-wise smoothing impossible. This is especially burdensome for table D8, the matrix representation of the seasonal-irregular component. For the DOM pattern, for instance, the row-wise number of missing values at the end of the month will range from 0 in months with 31 days to 3 in non-leap-year Februaries. The extended X-11 approach implements two imputation techniques for handling this issue. Missing values can be either interpolated per column, which is advantageous in case of strong DOM or DOY effects, or stretched within periods and interpolated per row, which is advantageous in case of strong end-of-month or end-of-year effects. The nature of the seasonal effect thus determines the appropriate imputation technique.

3.2.2 Extended AMB Approach

The extended AMB approach essentially translates the ARIMA model decomposition algorithm developed by Burman (1980) to the case of fractional Airline models. Its key idea can be summarized in four steps. First, an ARIMA model is fit to the linearized observations; second, the estimated model is decomposed into canonical ARIMA models for the unobserved components in (1) – or, more generally, for the signal and noise – by employing a factorization of the (stationary and non-stationary) AR polynomials and a partial fraction expansion of the MA part; third, the polynomials of the canonical ARIMA models

[4] In practice, a sequential extraction of seasonal patterns is often advocated, starting with the smallest seasonal period.

are rearranged to form the signal's Wiener-Kolmogorov (WK) filter; fourth, the estimated signal is obtained from applying the WK filter to the observations. Note that (fractional) Airline models contain only non-stationary AR polynomials, the roots of which can always be allocated to the trend-cyclical and seasonal components. As a consequence, the irregular component in (1) does not carry transitory movements and thus is white noise by construction.

The fractional Airline model also lays the base for a TRAMO-like pretreatment routine that includes estimation of user-defined regression effects and automatic outlier detection. The full regARFIMA pretreatment model is given by

$$(1-B)\prod_{\tau_i\in\mathcal{S}}(1-B^{\tau_i})\left\{y_t-\mathbf{x}_t^\top\boldsymbol{\beta}\right\}=(1-\theta_1 B)\prod_{\tau_i\in\mathcal{S}}(1-\theta_{\tau_i}B^{\tau_i})\{\varepsilon_t\}, \tag{6}$$

where $\mathbf{x}_t^\top = (\mathbf{c}_t^\top, \mathbf{o}_t^\top)$ is the vector of regression variables associated with calendar and outlier effects,[5] $\boldsymbol{\beta}^\top = (\boldsymbol{\beta}_c^\top, \boldsymbol{\beta}_o^\top)$ is the vector of unknown time-constant calendar and outlier effects and B is the backshift operator. To define fractional differencing, note that $B^{\tau_i} = B^{\lfloor\tau_i\rfloor}B^{\alpha_i}$ with $\alpha_i = \tau_i - \lfloor\tau_i\rfloor \in [0,1)$ and $B^{\alpha_i} \approx (1-\alpha_i) + \alpha_i B$ according to the first-order Taylor approximation at 1. The fractional differencing operator can thus be rewritten as

$$1 - B^{\tau_i} \approx 1 - (1-\alpha_i)B^{\lfloor\tau_i\rfloor} - \alpha_i B^{\lfloor\tau_i\rfloor+1}. \tag{7}$$

If $\tau_i \in \mathbb{N}$, then $\alpha_i = 0$ and (7) collapses to the ordinary differencing operator. This Taylor logic also applies to the seasonal MA polynomials occurring on the right-hand side of (6). For example, the fractional MA polynomial associated with the DOY pattern is approximated by

$$1 - \theta_{\text{DOY}}B^{365.2425} \approx 1 - 0.7575\,\theta_{\text{DOY}}B^{365} - 0.2425\,\theta_{\text{DOY}}B^{366}.$$

For daily data, however, the Burman-like decomposition algorithm can become unstable quite quickly. For that reason, model (6) is put into state space form (Gomez and Maravall, 1994) and estimated with a modified version of Koopman (1993)'s disturbance smoother that includes polynomial reduction and diffuse square root initialization. Given an estimator $\hat{\boldsymbol{\beta}}$ of the unknown regression effects, the estimated calendar and outlier components of model (1) are given by $\hat{c}_t = \mathbf{c}_t^\top\hat{\boldsymbol{\beta}}_c$ and $\hat{o}_t = \mathbf{o}_t^\top\hat{\boldsymbol{\beta}}_o$ in case of an additive decomposition and by $\hat{c}_t = \exp\{\mathbf{c}_t^\top\hat{\boldsymbol{\beta}}_c\}$ and $\hat{o}_t = \exp\{\mathbf{o}_t^\top\hat{\boldsymbol{\beta}}_o\}$ in case of a multiplicative decomposition.

3.2.3 Structural Time Series Models

STS and state space models have been popularized by Harvey (1989) and JD+ adopts many key concepts and routines from Durbin and Koopman (2012). The general idea is to specify a model for each unobserved component in (1) that reflects a priori beliefs about the component's dynamics. The component models then dictate the bottom-up model for the observations and its state space representation which is finally estimated with some recursive Kalman filtering and smoothing technique.

The trend-cyclical component is often given by

$$t_t \equiv \mu_t + \psi_t,$$

where $\{\mu_t\}$ is a time-varying level and $\{\psi_t\}$ denotes a (rather optional) cycle associated with some non-seasonal frequency $\lambda_c \in (0, \pi/\tau^\star)$ for some threshold $\tau^\star = \tau^\star(\mathcal{S})$ that

[5]The current implementation enables automatic detection of additive outliers, level shifts and wave-like outliers, which are defined by the sequence $(0, 1, -1, 0)$ around the outlier date.

depends on the seasonal periods actually occurring in (1). The level component is usually specified in more detail. A prime model is the local linear trend

$$\begin{aligned} \mu_{t+1} &= \mu_t + \beta_t + \eta_t, \\ \beta_{t+1} &= \beta_t + \zeta_t, \end{aligned} \tag{8}$$

where $\{\eta_t\}$ and $\{\zeta_t\}$ are independent zero-mean Gaussian white noise processes with finite variances $\sigma_\eta^2 > 0$ and $\sigma_\zeta^2 > 0$. A local linear trend thus follows a random walk with a time-varying drift, or slope, which itself is modeled as a random walk. Other popular level models in the vein of (8) include the local level (absence of slope $\{\beta_t\}$), the local level with drift ($\sigma_\eta^2 > 0$ and $\sigma_\zeta^2 = 0$) and the smooth trend ($\sigma_\eta^2 = 0$ and $\sigma_\zeta^2 > 0$).

The cycle component is usually specified in autoregressive or trigonometric form. For example, the latter is given by

$$\begin{pmatrix} \psi_t \\ \psi_t^\star \end{pmatrix} = \rho \begin{pmatrix} \cos\lambda_c & \sin\lambda_c \\ -\sin\lambda_c & \cos\lambda_c \end{pmatrix} \begin{pmatrix} \psi_{t-1} \\ \psi_{t-1}^\star \end{pmatrix} + \begin{pmatrix} \kappa_t \\ \kappa_t^\star \end{pmatrix}, \tag{9}$$

where $\rho \in (0, 1)$ is a damping factor that controls the cycle's persistence and $\{\kappa_t\}$ and $\{\kappa_t^\star\}$ are two mutually uncorrelated Gaussian white noise processes with common finite variance $\sigma_\kappa^2 > 0$.

Assuming $\tau_i \in \mathbb{N}$ for all $\tau_i \in \mathcal{S}$ and switching to standard STS notation, each seasonal pattern in (2) is modeled as $s_t^{(\tau_i)} \equiv \gamma_t^{(\tau_i)}$ with

$$\gamma_t^{(\tau_i)} = \mathbf{e}_{1,\tau_i}^\top \boldsymbol{\Phi}_t^{(\tau_i)}, \tag{10}$$

where $\mathbf{e}_{k,n}$ is the k-th unit vector of length n and $\boldsymbol{\Phi}_t^{(\tau_i)}$ is the column vector of the τ_i seasonal effects. The latter is assumed to follow the model

$$\boldsymbol{\Phi}_t^{(\tau_i)} = \mathbf{P}\boldsymbol{\Phi}_{t-1}^{(\tau_i)} + \boldsymbol{\omega}_t^{(\tau_i)}, \quad \boldsymbol{\omega}_t^{(\tau_i)} \overset{\text{iid}}{\sim} \mathcal{N}\left(\mathbf{0}, \sigma^2_{\omega^{(\tau_i)}} \times \boldsymbol{\Omega}^{(\tau_i)}\right), \tag{11}$$

where $\mathbf{P}$ is a known permutation matrix of appropriate dimensions that shuffles the first element of $\boldsymbol{\Phi}_{t-1}^{(\tau_i)}$ to the bottom of $\boldsymbol{\Phi}_t^{(\tau_i)}$. Furthermore, $\sigma^2_{\omega^{(\tau_i)}} > 0$ for all $\tau_i \in \mathcal{S}$ and each matrix $\boldsymbol{\Omega}^{(\tau_i)}$ is constructed to assure that the seasonal effects will add up to zero over any consecutive τ_i values. The representation (10) and (11) was introduced by West and Harrison (1997). JD+ implements the crude, dummy, Harrison-Stevens and trigonometric seasonal models. An in-depth description of these models is given by Proietti (2000), including the exact specification of $\boldsymbol{\Omega}^{(\tau_i)}$ in (11).

Calendar variation can again be taken into account with the aid of appropriate regression variables. This topic, however, is not discussed here as (6) will provide a common pretreatment model in our applications. The trend and seasonal dynamics specified in (8) to (11) are now put into a univariate linear Gaussian state space model. This model usually reads

$$y_t = \mathbf{Z}_t^\top \boldsymbol{\alpha}_t + \varepsilon_t, \tag{12}$$

$$\boldsymbol{\alpha}_{t+1} = \mathbf{T}_t \boldsymbol{\alpha}_t + \mathbf{R}_t \boldsymbol{\eta}_t, \tag{13}$$

where $\mathbf{Z}_t \in \mathbb{R}^m$ is the design vector, $\boldsymbol{\alpha}_t \in \mathbb{R}^m$ is the state vector, $\mathbf{T}_t \in \mathbb{R}^{m \times m}$ is the transition matrix and $\mathbf{R}_t \in \mathbb{R}^{m \times r}$ is a selection matrix that picks the $r \leq m$ elements of the state vector with non-zero disturbances. Furthermore, the observation and state disturbances are distributed as $\varepsilon_t \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma_\varepsilon^2)$ and $\boldsymbol{\eta}_t \overset{\text{id}}{\sim} \mathcal{N}(\mathbf{0}, \mathbf{Q}_t)$ with $\mathbf{Q}_t \in \mathbb{R}^{r \times r}$. Note that the former disturbances replace the irregular component in (1).

However, JD+ actually implements a state space framework which puts the observation disturbances in the state vector. The standard model (12) and (13) thus becomes

$$y_t = \tilde{\mathbf{Z}}_t^\top \tilde{\boldsymbol{\alpha}}_t, \tag{14}$$
$$\tilde{\boldsymbol{\alpha}}_{t+1} = \tilde{\mathbf{T}}_t \tilde{\boldsymbol{\alpha}}_t + \tilde{\mathbf{R}}_t \tilde{\boldsymbol{\eta}}_t, \tag{15}$$

with $\tilde{\mathbf{Z}}_t = (\mathbf{Z}_t^\top, 1)^\top$, $\tilde{\boldsymbol{\alpha}}_t = (\boldsymbol{\alpha}_t^\top, \varepsilon_t)^\top$, $\tilde{\boldsymbol{\eta}}_t = (\boldsymbol{\eta}_t^\top, \varepsilon_{t+1})^\top \overset{\text{id}}{\sim} \mathcal{N}(\mathbf{0}, \tilde{\mathbf{Q}}_t)$ and

$$\tilde{\mathbf{T}}_t = \begin{pmatrix} \mathbf{T}_t & \mathbf{0}_m \\ \mathbf{0}_m^\top & 0 \end{pmatrix}, \quad \tilde{\mathbf{R}}_t = \begin{pmatrix} \mathbf{R}_t & \mathbf{0}_m \\ \mathbf{0}_r^\top & 1 \end{pmatrix}, \quad \tilde{\mathbf{Q}}_t = \begin{pmatrix} \mathbf{Q}_t & \mathbf{0}_r \\ \mathbf{0}_r^\top & \sigma_\varepsilon^2 \end{pmatrix},$$

where $\mathbf{0}_n$ is a column vector of n zeros.

To give a brief example, assume we wish to specify a model for $\{y_t\}$ where the trend follows the local level version of (8) and the seasonal component contains only a single seasonal pattern with an integer seasonal period $\tau \in \mathbb{N}$. In this case, the observation equation is given by (14) with $\tilde{\mathbf{Z}}_t = (1, \mathbf{e}_{1,\tau-1}^\top, 1)^\top$ and $\tilde{\boldsymbol{\alpha}}_t = (\mu_t, \gamma_t, \ldots, \gamma_{t+\tau-2}, \varepsilon_t)^\top$ and the state equation is given by (15) with

$$\tilde{\mathbf{T}}_t = \begin{pmatrix} 1 & \mathbf{0}_{\tau-1}^\top & 0 \\ \mathbf{0}_{\tau-1} & \mathbf{T}_\gamma & \mathbf{0}_{\tau-1} \\ 0 & \mathbf{0}_{\tau-1}^\top & 0 \end{pmatrix}, \quad \tilde{\mathbf{R}}_t = \begin{pmatrix} \mathbf{e}_{1,3}^\top \\ \mathbf{1}_{\tau-2} \otimes \mathbf{0}_3^\top \\ \mathbf{e}_{2,3}^\top \\ \mathbf{e}_{3,3}^\top \end{pmatrix},$$

$$\tilde{\boldsymbol{\eta}}_t = \begin{pmatrix} \eta_t \\ \omega_t \\ \varepsilon_{t+1} \end{pmatrix}, \qquad \tilde{\mathbf{Q}}_t = \begin{pmatrix} \sigma_\eta^2 & 0 & 0 \\ 0 & \sigma_\omega^2 & 0 \\ 0 & 0 & \sigma_\varepsilon^2 \end{pmatrix},$$

where $\mathbf{1}_n$ is a column vector of n ones, $\otimes$ denotes the Kronecker product and

$$\mathbf{T}_\gamma = \begin{pmatrix} \mathbf{0}_{\tau-2} & \mathbf{I}_{\tau-2} \\ -\mathbf{1}_{\tau-1}^\top & \end{pmatrix},$$

with $\mathbf{I}_n$ being the n-dimensional identity matrix. Note that the system matrices are now invariant over time and that the inner block matrix $\mathbf{T}_\gamma$ essentially shuffles the τ seasonal effects similar to the permutation matrix $\mathbf{P}$ in (11). Also, its bottom row guarantees that $(1 + B + \cdots + B^{\tau-1})\,\gamma_t = \omega_t$ holds for all t. In the case of multiple seasonal patterns, all system vectors and matrices are extended appropriately. For example, the transition matrix $\tilde{\mathbf{T}}_t$ then contains multiple inner block matrices $\mathbf{T}_{\gamma^{(\tau_i)}}$ arranged in block-diagonal fashion.

Kalman filtering and smoothing techniques are used to estimate the state space model (14) and (15). JD+ implements the standard algorithms (Durbin and Koopman, 2012) as well as faster variants, such as Chandrasekhar-type recursions for time-invariant models and array filters (Morf et al., 1974; Morf and Kailath, 1975), alongside diffuse (Ansley and Kohn, 1990; Koopman and Durbin, 2003) and augmented initializations (de Jong, 1991; de Jong and Chu-Chun-Lin, 2003) and square root versions thereof.

3.3 Other Approaches

Besides the JD+ approaches, several other methods are qualified to model daily and other types of high-frequency data, even if they have not been designed primarily for this particular purpose. Focusing on the specification of seasonal dynamics, this section discusses the key ideas and concepts of selected alternative approaches.

3.3.1 Unobserved Component Models

The approaches presented in this section follow the same bottom-up modeling strategy as the STS approach: the latent component models are specified first and then aggregated to form the model for the observed time series. Conceptual overlaps with STS models are thus inevitable.

Exponential Smoothing Exponential smoothing is a forecasting technique based on moving averages with exponentially decaying weights. Examples include the Holt-Winters method (Holt, 1957; Winters, 1960) and the double and triple seasonal methods (Taylor, 2003, 2010). As demonstrated by Ord et al. (1997) and Hyndman et al. (2002), a common theoretical foundation for exponential smoothing methods is provided by the innovations state space framework, which also admits the general representation (12) and (13) but has the key difference that the observations and all unobserved components are driven by the exact same disturbances. For that reason, such models are sometimes referred to as single-source-of-error models.

Some efforts have been made recently to increase model flexibility in general and to incorporate a wider variety of seasonal patterns in particular. Gould et al. (2008) and Hyndman et al. (2008)[6] extend the double seasonal model to the case of $|\mathcal{S}| > 2$ under the restriction that the seasonal patterns are nested and associated with integer seasonal periods. De Livera et al. (2011) further generalize this approach by developing the BATS and TBATS models that can handle multiple nested and non-nested seasonal patterns with integer and non-integer seasonal periods. Here, TBATS is an acronym for the model's key features: trigonometric seasonal representation, Box-Cox transformation, ARMA disturbances, trend and seasonal components. Being similar to (9), the trigonometric representation of each seasonal cycle in (3) is given by

$$\begin{pmatrix} s_{j,t}^{(\tau_i)} \\ s_{j,t}^{(\tau_i),\star} \end{pmatrix} = \begin{pmatrix} \cos\lambda_j^{(\tau_i)} & \sin\lambda_j^{(\tau_i)} \\ -\sin\lambda_j^{(\tau_i)} & \cos\lambda_j^{(\tau_i)} \end{pmatrix} \begin{pmatrix} s_{j,t-1}^{(\tau_i)} \\ s_{j,t-1}^{(\tau_i),\star} \end{pmatrix} + \begin{pmatrix} \gamma_1^{(\tau_i)} \\ \gamma_2^{(\tau_i)} \end{pmatrix} \omega_t, \tag{16}$$

where $\gamma_1^{(\tau_i)}$ and $\gamma_2^{(\tau_i)}$ are pattern-specific smoothing parameters and $\{\omega_t\}$ is the common single-source-of-error ARMA disturbance driven by Gaussian white noise.

The TBATS model is extendable in a variety of ways. For example, the Box-Cox transformation could be replaced with the inverse hyperbolic sine transformation to fit time series with zero or negative observations. Non-Gaussian disturbances and modifications for multivariate time series could also be implemented within the general framework.

Atomic Seasonal Models Note that (16) specifies the same model for seasonal cycles that belong the same seasonal pattern. Atomic seasonal models suggested by McElroy (2017)[7] are similar in spirit but differ from the single-source-of-error approach in two respects: first, latent components are driven by individual uncorrelated innovations, as in traditional STS models; second, each seasonal cycle within a given pattern follows a separate model. Adopting the multivariate framework of McElroy (2017), the atomic model for each seasonal cycle in (3) is given by

$$\delta^{\lambda_j^{(\tau_i)}}(B)\,\mathbf{s}_{j,t}^{(\tau_i)} = \boldsymbol{\varepsilon}_{j,t}^{(\tau_i)}, \tag{17}$$

[6] Ox code is available under URL `https://robjhyndman.com/publications/multiple-seasonal-patterns`. The {`forecast`} R package also implements some exponential smoothing methods.

[7] The {`sigex`} R package can be downloaded from `https://github.com/tuckermcelroy/sigex`.

where $\delta^{\omega}(B) = 1 - 2\cos(\omega)B + B^2$ is a differencing operator at frequency $\omega \in [0, \pi]$, $\{\mathbf{s}_{j,t}^{(\tau_i)}\}$ is a vector of seasonal cycles and $\{\boldsymbol{\varepsilon}_{j,t}^{(\tau_i)}\}$ is multivariate Gaussian white noise with covariance matrix $\boldsymbol{\Sigma}_{j}^{(\tau_i)}$. Note that $\delta^{\omega}(B)$ yields a complete factorization of seasonal unit root differencing polynomials. For example, $1 - B^7 = \delta^{2\pi/7}(B)\,\delta^{4\pi/7}(B)\,\delta^{6\pi/7}(B)$.

The atomic nature of (17) increases model flexibility as it allows the widths and heights of spectral peaks to be controlled by multiple parameters and thus to be different within each seasonal pattern. This especially facilitates coverage of rapidly moving seasonality, which can be advantageous in times of strong economic changes. The multivariate nature of the entire framework adds further flexibility as the latent components can be common, related or unrelated across original series, depending on whether their white noise is collinear or, if not, has a non-diagonal versus diagonal covariance matrix. The generalized Cholesky decomposition of the potentially rank-reduced covariance matrices also allows the latent components to be interpreted as dynamic factor processes.

Model estimation can be done by OLS, method of moments (MOM) or maximum likelihood, which is evaluated with the Durbin-Levinson algorithm. Multivariate signal extraction is done with generalized Wiener-Kolmogorov (WK) matrix formulas that have been derived by McElroy and Trimbur (2015) as a generalization of the univariate case studied by McElroy (2008).

However, some refinements and modifications are needed to make atomic models and the direct matrix formulas for signal extraction (McElroy and Trimbur, 2015) feasible for daily data, or large data sets in general. McElroy and Monsell (2017) introduce canonical atomic models for which it is impossible to extract any additional (additive) white noise. As a consequence, the differenced canonical or stabilized latent components have MA-like structures instead of being white noise. Utilizing the generalized Cholesky decomposition, McElroy and Monsell (2017) also derive invertible MOM estimators that ensure positive definiteness of the estimated covariance matrices. Finally, they advocate the application of the generalized WK filter to the forecast-extended original series in combination with the Hodrick-Prescott filter that eventually separates trend and DOY seasonality. The forecasts are obtained from the invertible MOM estimates, using recursive one-step ahead predictions provided by the Durbin-Levinson algorithm. An additional application of canonical atomic models to daily data is given by McElroy et al. (2018).

3.3.2 Extended STL Approach

Cleveland et al. (1990) have propagated a seasonal-trend decomposition routine based on LOESS regressions (STL)[8] as an alternative to the X-11 method. Their procedure follows the same iterative philosophy but can handle any single integer seasonal period greater than one and missing values. STL essentially consists of an inner and outer loop, with the former being responsible for sequential estimation of latent components and the latter achieving extreme value correction, if necessary. Two variants have been developed recently and implemented in R, each of which facilitates sequential extraction of multiple seasonal patterns, an idea already alluded to by Cleveland et al. (1990).

The {`rjdhf`} **package** This package basically provides access to the Java translation of the original STL procedure. The regARFIMA pretreatment model (6) can be used for data linearization as well as the outer loop as a robust alternative for outlier treatment. The user can also choose between an additive or multiplicative form of the latent component

[8]The original Fortran code has been translated to R via the `stl()` function of the {`stats`} package. An enhanced version is available in the {`stlplus`} package.

model (1) and – calling the STL routine $|\mathcal{S}|$ times – specify the number of neighboring observations to be considered in the inner-loop local regressions for trend and seasonal smoothing. Non-integer seasonal periods are automatically rounded down to the nearest integer.

The {dsa} **package** Ollech (2018) embeds classical regARIMA pretreatment in a 4-step STL-based procedure that utilizes (2) with $\mathcal{S} = \{7, 31, 365\}$.[9] In the first step, the DOW pattern is extracted from the original series. In a second step, outlier and calendar effects are estimated from the DOW-adjusted series. To this end, a non-seasonal regARIMA model with deterministic trigonometric seasonal terms or a seasonal regARIMA model can be used. Either way, built-in routines for automatic outlier detection and generation of calendar regression variables and forecasts are readily available. In a third step, the DOM pattern is extracted from the linearized DOW-adjusted series. Cubic splines are employed to temporarily stretch shorter months to 31 days. In a fourth step, the DOY pattern is estimated from the linearized DOM and DOW-adjusted series. Leap-year Februaries are temporarily shortened by skipping the last day so that each year has 365 days. Empty spots in the seasonally adjusted series related to skipped observations are finally filled with spline interpolations.

3.3.3 STR Approach

Motivated by the STL philosophy, Dokumentov and Hyndman (2015) discuss a related procedure in which the LOESS regression is essentially replaced with a more general ridge- or LASSO-type regression model (STR).[10] The trend component in (1) is assumed to evolve smoothly over time, such as an $I(2)$ process. The key difference, however, to all approaches discussed so far is that the seasonal component is thought of as a two-dimensional array that carries both visible and invisible seasonal dynamics. To elaborate this idea, assume that the observed time series has a finite length T in (1), written as $\mathbf{y}_T$, and all seasonal periods in (2) are integers. Then, each seasonal pattern can be expressed in matrix form as

$$\mathbf{S}^{(\tau_i)} = [s_{k,l}]_{k,l}^{(\tau_i)} \in \mathbb{R}^{\tau_i \times T}, \tag{18}$$

where each column stores τ_i elements of the seasonal pattern at time t. Only one of those elements is actually in play and the others are hidden. Thus, (18) connects to (2) via

$$s_t^{(\tau_i)} = \mathbf{S}_{\tau_i^\star(t),t}^{(\tau_i)}$$

for some mapping $\tau_i^\star : \mathbb{N} \mapsto \{1, \dots, \tau_i\}$ that picks the pattern's correct season at time t.

The STR approach also enables a more nuanced specification of the calendar component as it explicitly distinguishes between three types of calendar regression variables $\mathbf{c}_t$. The first type, called static predictors, has constant effects, as in (6); the second and third types, called flexible and seasonal predictors, have time-varying effects that evolve in a non-seasonal and seasonal manner, respectively.

The entire STR equivalent of model (1) can be recast as a linear regression model that has the same likelihood function. The latter takes the form $\mathbf{y}_{\text{ext}} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, where $\mathbf{y}_{\text{ext}} = (\mathbf{y}_T^\top, \mathbf{0}_n^\top)^\top$ for some proper choice of n and $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \sigma_{\boldsymbol{\varepsilon}}^2 \boldsymbol{\Sigma})$ with $\boldsymbol{\Sigma}$ being a block matrix that carries $\mathbf{I}_T$ in the upper left block and zeros everywhere else. The STR model can thus be estimated by some maximum likelihood technique. The underlying optimization problem is expressed in terms of discrete second-order derivatives of the latent components,

[9] The {dsa} package is available for download on CRAN.
[10] The {stR} package is available for download on CRAN.

which has two consequences. First, smoothness restrictions can be imposed on the seasonal component in three dimensions: the time, time-season and season dimensions. Second, fractional seasonal periods can be incorporated quite naturally. To see this, note that the seasonal patterns (and all other latent components) can again be rewritten as functional components, that is as linear combinations of smooth basis functions, such as Fourier terms, splines or wavelets. This results in a valid reparametrization of the second-order derivative matrices and ultimately a substantial reduction of computation time.

Depending on the optimization problem at hand, OLS-based techniques or numerical solutions apply alongside an evaluation of cross-validated residuals to obtain estimates of the smoothing parameters contained in the design matrix $\mathbf{X}$ and of σ^2_{ε}. A robust STR method is also available. In this case, the residuals ε are assumed to be Laplacian which leads to an optimization problem of a quantile regression that can be solved only numerically.

3.3.4 Regularized Singular Value Decomposition

Lin et al. (2020) suggest a novel approach based on regularized singular value decomposition (RSVD) that has been developed by Huang et al. (2008, 2009).[11] Like the STR approach, the RSVD method reorganizes the observations and latent components in matrix form. The seasonal matrix, however, stores only the visible elements of (18). Assume that the observations are again given by a finite sequence of length T and contain only a single seasonal pattern with an integer seasonal period τ such that $n = T/\tau \in \mathbb{N}$. Ignoring temporarily the calendar component, we may then rewrite (1) as $\mathbf{Y} = \mathbf{S} + \mathbf{N}$, where

$$\mathbf{S} = [s_{k,l}]_{k,l} \in \mathbb{R}^{n\times\tau} \tag{19}$$

is the seasonal matrix that stores the seasonal pattern related to the i-th subspan (week, year, etc.) in its i-th row, and $\mathbf{Y}$ and $\mathbf{N}$ denote the observation and non-seasonal matrices defined analogously. Accordingly, (19) is linked to (2) via

$$s_t^{(\tau_i)} \equiv s_t = \mathbf{S}_{\lceil t/\tau\rceil, t-\lfloor t/\tau\rfloor\tau},$$

where $\lceil x\rceil$ is the smallest integer not less than x. The seasonal matrix is further decomposed according to

$$\mathbf{S} = \mathbf{1}_n\mathbf{f}^\top + \mathbf{U}\mathbf{V}^\top, \tag{20}$$

where $\mathbf{f} = (f_1, \ldots, f_\tau)^\top$ is the vector of τ fixed seasonal effects and $\mathbf{U} \in \mathbb{R}^{n\times r}$ and $\mathbf{V} \in \mathbb{R}^{\tau\times r}$ are the matrices of left and right singular vectors of $\mathbf{S}$ that satisfy $\mathbf{U}^\top\mathbf{1}_n = \mathbf{0}_n$ and $\mathbf{V}^\top\mathbf{V} = \mathbf{I}_r$ for some $r \leq \tau$ to ensure identifiability. Thus, (20) provides an explicit modeling of stable and moving seasonality where the latter is actually represented as a linear combination of r moving patterns, or layers, stored in $\mathbf{V}$ with corresponding span-varying coefficients, or magnitudes, stored in $\mathbf{U}$.

After selecting r, which can be achieved via the Bayesian Information Criterion, model estimation proceeds in three steps. The first step aims at obtaining an estimate $\hat{\mathbf{U}}$ via a loop-like sequential application of the RSVD method under a zero-sum restriction on the layers in $\mathbf{V}$. RSVD is applied to the column-wise demeaned observation matrix $\mathbf{Y}$ in the first loop, which eliminates $\mathbf{f}$, and to residual matrices in the next $r-1$ loops. If $\mathbf{N}$ contains non-stationary dynamics, such as $I(1)$ trends, then first-order column-differenced matrices are considered and thus the zero-sum restriction on $\mathbf{V}$ is not needed. The second step aims at jointly obtaining estimates $\hat{\mathbf{f}}$ and $\hat{\mathbf{V}}$. This is achieved by solving a least squares problem

[11] R code is available for download as part of the paper's online appendix.

that uses $\hat{\mathbf{U}}$ as a plug-in estimator and is subject to the zero-sum constraints $\mathbf{f}^\top \mathbf{1}_\tau = 0$ and $\mathbf{V}^\top \mathbf{1}_\tau = \mathbf{0}_r$. The estimates $\hat{\mathbf{f}}$, $\hat{\mathbf{U}}$ and $\hat{\mathbf{V}}$ then yield $\hat{\mathbf{S}}$ via (20). The third (rather optional) step aims at obtaining parameter estimates of the non-seasonal component's model.

The first step of this procedure can be slightly modified to incorporate constrained moving seasonality. The RSVD method imposes a roughness penalty on each column of $\mathbf{U}$ that includes a smoothing parameter determined by generalized cross-validation. Abrupt changes in the seasonal layer coefficients can thus be taken into account by using different smoothing parameters for different parts of the corresponding columns of $\mathbf{U}$. An automatic break detection routine has been implemented, which is based on minimizing the squared differences between the first-order differenced observations and estimated seasonal component.

Calendar effects have been ignored so far but could be easily integrated into the RSVD approach as well by including regression variables in the constrained least squares problem that is solved in the second step. Multiple seasonal patterns could also be dealt with by reorganizing the observations into a $|\mathcal{S}|$-dimensional array in place of $\mathbf{Y}$.

The RSVD approach performs particularly well in the presence of strong seasonality and/or seasonal breaks according to a simulation study for monthly time series. This study, however, also revealed that the X-11 and AMB approaches, as implemented in X-13ARIMA-SEATS, tend to be superior in cases of weak seasonality.

3.3.5 Prophet

Prophet is an open-source Bayesian forecasting procedure implemented in Python and R and released by Facebook's Core Data Science team.[12] Claiming that several popular time series models, including ARIMA and TBATS, generally struggle to produce high-quality forecasts of business time series, its primary aim is to provide a flexible and reliable forecasting tool based on component models that can be configured, interpreted and evaluated by subject-matter experts and analysts without expertise in time series modeling.

Taylor and Letham (2018) describe the Prophet approach and its latent component model, which coincides with (1), in detail. The trend can be driven by a saturating piecewise logistic or a piecewise constant growth model. Either way, changepoints in the trend can be specified by the user or detected automatically using a Laplacian prior for the changes in growth rate at the changepoints.

The seasonal component is composed of stable patterns in which a Gaussian prior is imposed on the amplitudes of the seasonal cycles in (4). Let

$$\mathbf{s}_t^{(\tau_i)} = \left[\cos\left(\frac{2\pi t}{\tau_i}\right), \sin\left(\frac{2\pi t}{\tau_i}\right), \cdots, \cos\left(\frac{2\pi J t}{\tau_i}\right), \sin\left(\frac{2\pi J t}{\tau_i}\right)\right]$$

and $\boldsymbol{\gamma} = (\alpha_1, \beta_1, \ldots, \alpha_J, \beta_J)^\top$, where $J = J(\tau_i)$ is the pattern-specific number of seasonal cycles. Then, each stable seasonal pattern is modeled as

$$s_t^{(\tau_i)} = \mathbf{s}_t^{(\tau_i)} \times \boldsymbol{\gamma}, \quad \boldsymbol{\gamma} \sim \mathcal{N}(\mathbf{0}_{2J}, \sigma_\gamma^2 \times \mathbf{I}_{2J}).$$

Prophet can thus deal with non-nested and non-integer seasonal periods. Furthermore, it also allows for the inclusion of constraints according to (5).

The calendar component is essentially modeled in the same ways as in (6), that is $c_t = \mathbf{c}_t^\top \boldsymbol{\beta}_c$. The elements of $\mathbf{c}_t$ are dummy indicators that can be defined via built-in lists of national and international holiday dates or specified by the user. The regression

[12] The {fbprophet} and {prophet} packages are available for download on PyPI and CRAN, respectively. An online documentation is available under URL https://facebook.github.io/prophet.

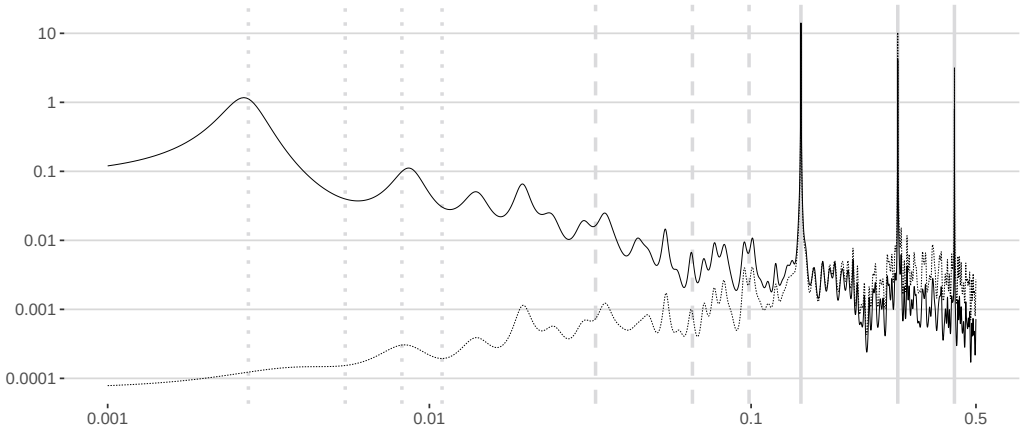


Figure 3: Estimated AR spectrum of logged (*solid*) and differenced logged (*dashed*) electricity consumption, scaled by 2π. Gray verticals correspond to selected seasonal harmonics of the DOY (*dotted*), DOM (*dashed*) and DOW (*solid*) patterns.

effects β_c are then assumed to have a Gaussian prior, which is also the case for the irregular component. Finally, model estimation is carried out in Stan using the L-BFGS optimization algorithm to find the maximum a posteriori estimates.

4. Daily Electricity Consumption in Germany

This section illustrates the capabilities of selected seasonal adjustment approaches, using daily realized electricity consumption in Germany as of 1 January 2015 up to 30 August 2020.[13] The linearization of electricity consumption is described first (Section 4.1). The unconstrained stable and moving seasonal patterns are then extracted from the linearized series with the three JD+ approaches (Section 4.2). Finally, a stable DOW pattern is estimated with Prophet, imposing a constraint on each year's period between Christmas and New Year (Section 4.3).

4.1 Pretreatment

The seasonal and calendar profiles of daily electricity consumption have already been shown in Figure 1 and described in Section 2.1. In particular, the DOW and DOY patterns were shown to clearly dominate the seasonal profile, whereas the DOM pattern turned out to be less pronounced. The estimated autoregressive spectra displayed in Figure 3 immediately confirm this first impression.

We therefore dismiss the DOM pattern and fit the regARFIMA pretreatment model (6) to logged electricity consumption with $\mathcal{S} = \{7, 365.2425\}$. Dummy regression variables are used to account for standard fixed and moving holidays in Germany, excluding Sundays. Dummies for regional holidays are weighted based on the regional employee structure and dummies for moving holidays are centered around their daily means. We also run automatic detection of additive outliers, level shifts and wave outliers and eventually keep all calendar and outlier regression variables with absolute t-values larger than 4.5.[14]

[13]This series as well as other daily data on forecasted and realized electricity consumption and production in Germany is freely available from the "Bundesnetzagentur | SMARD.de" provider under URL `https://www.smard.de/en`.

[14]Being slightly higher than the defaults used typically for monthly and quarterly data, this threshold suc-

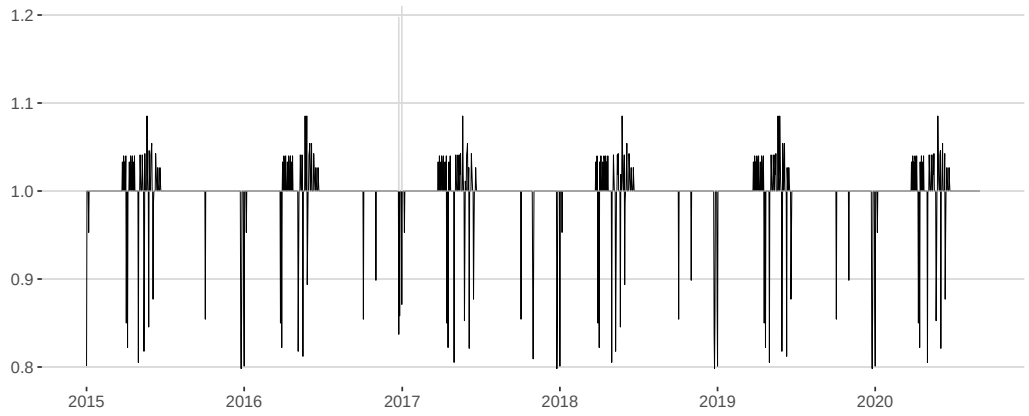


Figure 4: Estimated calendar (*black*) and outlier (*gray*) components of daily electricity consumption.

Table 1 summarizes the estimated regARFIMA pretreatment model (6). The estimated seasonal MA parameters exceed 0.85 and thus reflect nicely the stability of the DOW and DOY patterns. Some bridging days effects also turned out to be significantly different from zero. Note that this granular information also shows up in monthly and quarterly data but is not removed in official statistics, in line with international recommendations. Two additive outliers have been detected by the automatic search routine, the dates of which coincide with the Saturdays of Christmas and New Year's Eve 2016. This underlines the difficulty of estimating reliably time-constant regression effects of rare events in general and may also indicate the presence of cross-dependencies between the DOW pattern and fixed holidays. The latter issues could be solved in theory with regression variables accounting for such cross-effects. In practice, however, this is currently infeasible as the sample size is shorter than six years. Figure 4 shows the estimated calendar and outliers components.

4.2 Seasonal Adjustment

The DOW and DOY patterns are now extracted sequentially from linearized electricity consumption with the (extended) AMB, STS and X-11 approaches implemented in JD+. More precisely, the DOW pattern is estimated in the first step and the DOY pattern is extracted from the DOW-adjusted linearized series in a second step. Table 2 summarizes the key model specifications for each of the three approaches.

The extended AMB approach decomposes the pretreatment model (6) into the trend, seasonal and irregular component models according to a set of predefined rules and therefore only the seasonal periods in $\mathcal{S}$ need to be specified.

The STS approach demands integer seasonal periods and further specification of the component models. In the first step, the trend is modeled as a local level with drift and the trigonometric seasonal representation is used for the DOW pattern. In the second step, the seasonal period is rounded down to 365 days, the trend is modeled as a local level and the DOY pattern is represented by a persistent trigonometric cycle swinging at the (non-

cessfully prevents borderline significant effects from showing up in the model according to test calculations based on quasi real-time data vintages. It also accounts nicely for the number of observations during automatic outlier detection as it is very close to 4.4 which is the X-13 default critical value at the 5% level of significance according to the implemented modified formula of Ljung (1993).

Table 1: Estimated regARFIMA pretreatment model (6) for daily electricity consumption.

Event	Date	Weight	Estimate	SD	t-value
FIXED HOLIDAYS					
New Year's Day	1 Jan	1.0	−0.222	0.011	−20.701
Epiphany	6 Jan	0.4	−0.120	0.024	−4.968
Labor Day	1 May	1.0	−0.217	0.010	−22.580
German Unification Day	3 Oct	1.0	−0.157	0.010	−15.791
500th Reformation Day	31 Oct 2017	1.0	−0.212	0.021	−10.096
All Saints' Day	1 Nov	0.7	−0.153	0.016	−9.391
Christmas Eve	24 Dec	1.0	−0.178	0.013	−13.752
Christmas Day	25 Dec	1.0	−0.225	0.012	−18.370
Boxing Day	26 Dec	1.0	−0.153	0.011	−14.237
New Year's Eve	31 Dec	1.0	−0.138	0.013	−10.886
MOVING HOLIDAYS					
Good Friday		1.0	−0.195	0.009	−22.564
Easter Monday		1.0	−0.235	0.009	−27.193
Ascension		1.0	−0.241	0.009	−26.885
Pentecost Monday		1.0	−0.250	0.009	−28.861
Corpus Christi		0.7	−0.225	0.013	−17.484
BRIDGING DAYS					
Ascension Friday		1.0	−0.112	0.009	−12.423
Corpus Christi Friday		0.7	−0.095	0.013	−7.409
500th Reformation Day	30 Oct 2017	1.0	−0.118	0.021	−5.687
ADDITIVE OUTLIERS					
Christmas Eve 2016	24 Dec 2016		0.180	0.023	7.971
New Year's Eve 2016	31 Dec 2016		0.191	0.023	8.462
MA PARAMETERS					
θ_1			0.417	0.037	11.334
θ_7			0.875	0.014	61.041
$\theta_{365.2425}$			0.864	0.017	50.253

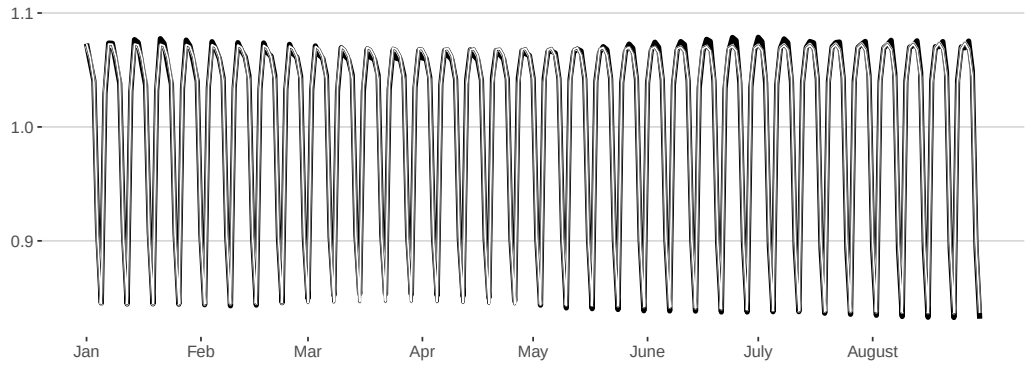
Remarks: National events are given a weight of 1.0. Regional events are weighted according to the approximate share of employees working in the federal states where the event is actually celebrated. Regression variables for moving holidays and associated bridging days have been centered by removing daily means calculated over the entire sample.

Table 2: Key specifications for seasonal adjustment of daily electricity consumption.

Pattern	Element	AMB	STS	X-11
DOW	τ_i	7	7	7
	TREND	Derived	(8) with $\sigma_\zeta^2 = 0$	9-term Henderson
	SEASONAL	from (6)	(11) with trigonometric	3×9 filter
DOY	τ_i	365.2425	365	365.2425
	TREND	Derived	(8) without $\{\beta_t\}$	371-term Henderson
	SEASONAL	from (6)	(9) with $\rho = 0.95$ and $\lambda_c = 2\pi/365.2425$	3×1 filter

Remark: Options not mentioned here are used in default mode.

(a) Day-of-the-week (DOW) pattern: 1 January to 30 August 2020



(b) Day-of-the-year (DOY) pattern: 1 January 2019 to 30 August 2020

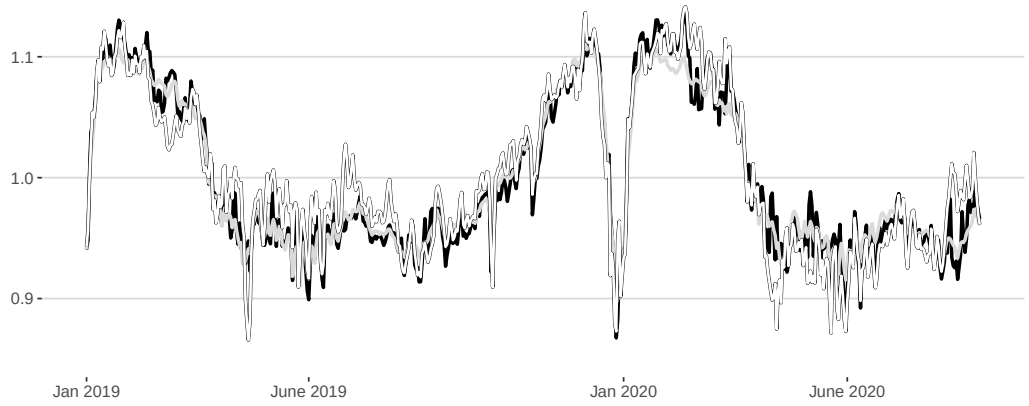


Figure 5: Seasonal patterns estimated with the extended X-11 (*black*), extended AMB (*gray*) and STS (*white*) approaches.

integer) DOY chief harmonic. The latter sparse specification[15] is chosen since any standard model of the DOY pattern turned out to be prohibitive in terms of computation time and thus infeasible in practice due to the high dimensions of the derived state space model, which is initialized with the diffuse square root algorithm in either step.

The extended X-11 approach requires further specification of trend and seasonal filters and σ-limits for detection of extremes in the detrended series. In general, we use the third-order exact Henderson kernel and default σ-limits given by $(1.5, 2.5)$. Asymmetric filters are obtained from cutting symmetric out-of-sample weights and normalizing in-sample weights, noting that the pretreatment model (6) does not provide forecasts of the unadjusted series. In the first step, the Henderson filter covers 9 days and thus is slightly longer than the DOW seasonal period to ensure that DOY dynamics remain in the trend-cycle and are passed on to the next step. Also, a longer seasonal filter is applied in line with the DOW pattern's stability. In the second step, the Henderson filter is again slightly longer than the seasonal period but an application of the short 3×1 seasonal filter is dictated by the availability of just five observations for each day in the final trimester of the calendar year in combination with the fractional seasonal period.

[15]This approach is similar to McElroy and Monsell (2017) who assign DOY dynamics to the trend-cycle, estimate the joint component and then wash out the annual seasonal cycle with the Hodrick-Prescott filter.

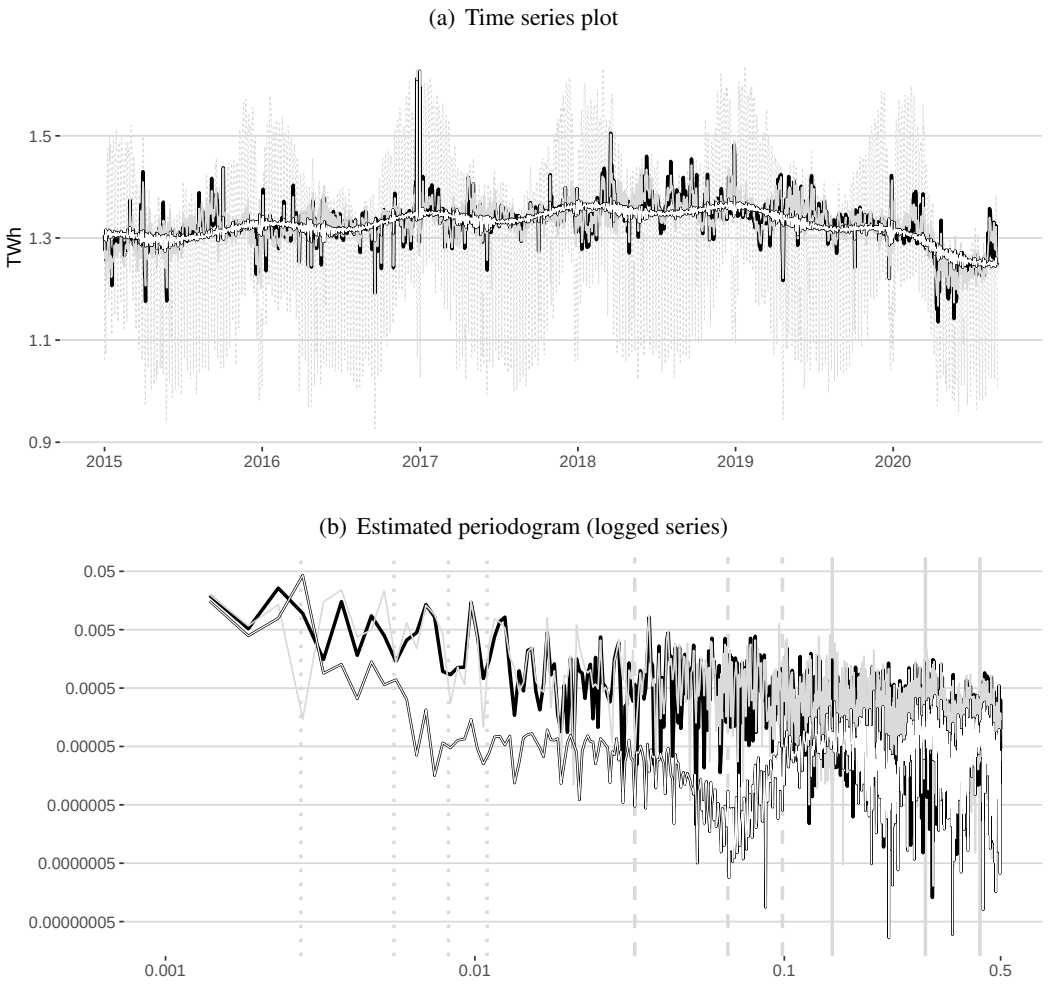


Figure 6: Seasonally adjusted electricity consumption obtained from the extended X-11 (*black*), extended AMB (*gray*) and STS (*white*) approaches. Dashed gray line in Panel (a) corresponds to unadjusted series. Gray verticals in Panel (b) correspond to selected seasonal harmonics of the DOY (*dotted*), DOM (*dashed*) and DOW (*solid*) patterns.

Figure 5 shows the seasonal patterns estimated with the three approaches. The estimated DOW patterns displayed in Panel (a) are stable and very similar in size and shape with the notable exception of more moving seasonality in the X-11 estimates, especially around the COVID-19-related phase from March to June 2020. The estimated DOY patterns shown in Panel (b) capture nicely the characteristic U-shaped intra-yearly dynamics and spike year-end troughs of electricity consumption. However, the STS estimates remain somewhat higher during the summer and deviate markedly from the AMB and X-11 estimates in and after the COVID-19-lockdown phase. The three approaches also differ visibly in terms of volatility. Whereas the AMB estimates are canonically smooth by construction, the X-11 and especially the STS estimates fluctuate much more rapidly. For the latter approach, this might be the price to be paid for specifying the sparse trigonometric cycle instead of a more traditional seasonal model.

Figure 6 displays information about the respective versions of seasonally adjusted electricity consumption. Panel (a) shows that, overall, each approach manages to remove the most distinct repetitive features of the raw series. However, the STS version seems to contain too much residual periodic annual structures and too little irregular movements. Panel

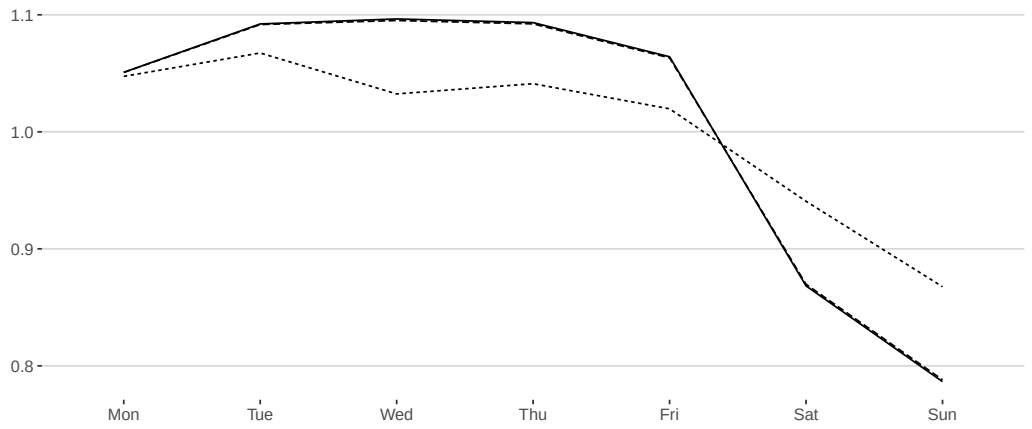


Figure 7: Prophet estimates of the stable DOW pattern without constraints (*solid*) and with constraints to W01 through W51 (*dashed*) and W52 through W53 (*dotted*).

(b) confirms this impression as the estimated periodogram of the STS version still peaks at the DOY chief harmonic, as opposed to the other approaches.

4.3 Constrained DOW Pattern

Each of the three JD+ approaches estimated DOW and DOY patterns that cover unconstrained stable and moving seasonality. However, Figure 1 (c) already provided visual evidence that the DOW pattern between Christmas and New Year (C2NY) looks noticeably different from the rest of the year. This motivates a brief comparison between unconstrained and constrained estimates. To this end, we extract stable DOW and DOY patterns from linearized electricity consumption with Prophet and use (5) with $\mathcal{C} = \{\text{W52}, \text{W53}\}$ to impose the C2NY constraint on the former pattern. Three seasonal cycles are included in the DOW pattern as before and ten in the DOY pattern. Also, a piecewise linear trend is specified alongside automatic changepoint selection.

Figure 7 shows the estimated unconstrained and constrained stable DOW pattern. Compared to the unconstrained estimates, the constrained C2NY estimates are lower from Monday to Friday and higher on Saturday and Sunday, resulting in a lower weekend drop-off and a flatter overall shape. In contrast, the estimates constrained to the rest of the year are almost indistinguishable from the unconstrained estimates.

5. Summary

Both the emergence of new digital data sources and the recent outbreak of the COVID-19 pandemic have increased the interest in and demand for more timely and more granular data. As a particular consequence, more daily macroeconomic time series have become available in official statistics. These series are often seasonal but exhibit common features that are usually not observable in monthly and quarterly time series so that traditional modeling and seasonal adjustment approaches in official statistics are not applicable. We discussed peculiarities of daily data and highlighted structural differences from monthly and quarterly time series. Key aspects included irregular spacing, coexistence of multiple seasonal patterns with integer versus non-integer seasonal periods and a high chance of cross-dependencies. We also addressed a broad range of modeling issues associated with those peculiarities and used a unified additive latent component model to give an overview

of modeling and seasonal adjustment approaches that have been developed mostly within the last ten years and are capable of handling a fair amount of them. We finally illustrated the capabilities of selected approaches, in particular of those implemented in a preliminary version of JDemetra+ 3.0, using daily realized electricity consumption in Germany.

These developments provide a solid foundation for future research that could head in a variety of directions, ranging from interactive graphical tools that facilitate visualization of granular data over time series models that balance flexibility in dynamics against sparseness in parameters to seasonal adjustment approaches that work in mass production. Each of these directions will face its own challenges. New graphical and other explorative tools, such as tailored seasonality tests, should be well-prepared for time series that contain a more complex seasonal profile and often noticeably more noise than monthly and quarterly data. Flexible and potentially multivariate modeling approaches should enable straightforward incorporation of additional daily data facets, such as correlated latent components (e.g. McElroy and Maravall, 2014) or GARCH-type heteroskedasticity (e.g. Koopman et al., 2007). Fast algorithms and parallel computing could be used to speed up model estimation in such increasingly complex cases. Modeling approaches, however, should also take into account that finding proper latent component models, especially for complex seasonal and calendar variations, can quickly turn into a time-consuming task even for a single time series. Functional component approaches seem promising in terms of flexible yet sparse models and some basis functions, such as splines, are directly applicable to irregularly spaced data (e.g. Rodríguez and Hernández, 2010). Dimension reduction techniques could also be used in general to control model complexity. Functional data analysis and continuous-time models could be another interesting strand of research, especially when studying data at higher intra-daily frequencies.

These elaborated or even new modeling approaches for daily data will naturally pave the way for new developments in the field of seasonal adjustment. Existing approaches, however, could also be developed further. For example, the extended AMB approach could be generalized to (non-Airline) ARFIMA models. Hybrid approaches that combine model-based and ad hoc filters (e.g. McElroy and Monsell, 2017) seem promising as well. Assuming that the availability of and demand for daily and other types of high-frequency time series will continue to increase, mass production of seasonally adjusted data may become another challenge at some point. In this case, automatic detection and selection procedures, such as I/C and I/S ratio-based filter selection in the extended X-11 approach, are needed more than ever given the high degree of individuality of daily time series.

Acknowledgement

This study greatly benefitted from the work of the Bundesbank's research group "Extensions JD+", which ran from January 2019 to October 2020. I therefore thank all the participants (in alphabetical order): Sindy Brakemeier, Andreas Dietrich, Nina Gonschorreck, Jan-Stephen Heller, Lea Hengen, Christiane Hofer, Julian LeCrone, Andreas Lorenz, Jörg Meier, Daniel Ollech, and Thomas Witthohn. I also thank Duncan Elliott of the UK Office for National Statistics, Steve Matthews of Statistics Canada, and Tucker McElroy of the U.S. Census Bureau for inspiring discussions on methodological issues during the Time Series Meeting held at the Royal Statistical Society in London on 28-29 March 2019, and Brian Monsell of the U.S. Bureau of Labor Statistics for sharing R code related to X-13 routines.

REFERENCES

Ansley, C. F., and Kohn, R. (1990), "Filtering and Smoothing in State Space Models with Partially Diffuse Initial Conditions," *Journal of Time Series Analysis*, 11, 275–293.

Askitas, N., and Zimmermann, K. F. (2011), "Nowcasting Business Cycles Using Toll Data," IZA Discussion Paper No. 5522.

Burman, J. P. (1980), "Seasonal Adjustment by Signal Extraction," *Journal of the Royal Statistical Society: Series A (General)*, 143, 321–337.

Cabrero, A., Camba-Mendez, G., Hirsch, A., and Nieto, F. (2009), "Modelling the Daily Banknotes in Circulation in the Context of the Liquidity Management of the European Central Bank," *Journal of Forecasting*, 28, 194–217.

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

Commandeur, J. J. F., Koopman, S. J., and Ooms, M. (2011), "Statistical Software for State Space Methods," *Journal of Statistical Software*, 41, 1–18.

Cox, M., Triebel, J., Linz, S., Fries, C., Flores, L. F., Lorenz, A., Ollech, D., Dietrich, A., LeCrone, J., and Webel, K. (2020), "Täglicher Lkw-Maut-Fahrleistungsindex aus digitalen Prozessdaten der Lkw-Mauterhebung," *Wirtschaft und Statistik*, 4/2020, 63–76 (in German).

Dagum, E. B., and Bianconcini, S. (2008), "The Henderson Smoother in Reproducing Kernel Hilbert Space," *Journal of Business and Economic Statistics*, 26, 536–545.

Dagum, E. B., and Bianconcini, S. (2015), "A New Set of Asymmetric Filters for Tracking the Short-Term Trend in Real-Time," *Annals of Applied Statistics*, 9, 1433–1458.

Dagum, E. B., and Bianconcini, S. (2016), *Seasonal Adjustment Methods and Real Time Trend-Cycle Estimation*, Berlin: Springer.

de Jong, P. (1991), "The Diffuse Kalman Filter," *The Annals of Statistics*, 19, 1073–1083.

de Jong, P., and Chu-Chun-Lin, S. (2003), "Smoothing With An Unknown Initial Condition," *Journal of Time Series Analysis*, 24, 141–148.

De Livera, A. M., Hyndman, R. J., and Snyder, R. D. (2011), "Forecasting Time Series With Complex Seasonal Patterns Using Exponential Smoothing," *Journal of the American Statistical Association*, 106, 1513–1527.

Deutsche Bundesbank (2020), "A weekly activity index for the German economy," *Monthly Report*, 72 (5), 68–70.

Dickopf, X., Janz, C., and Mucha, T. (2019), "Moving from GDP flash estimates to GDP nowcasts: first results of a feasibility study to further accelerate early GDP estimation," *Wirtschaft und Statistik*, 6/2019, 47–58 (in German, English version available under URL `https://www.destatis.de`).

Dokumentov, A., and Hyndman, R. J. (2015), "STR: A Seasonal-Trend Decomposition Procedure Based on Regression," *Working Paper*, 13/15, Monash University.

Durbin, J., and Koopman, S. J. (2012), *Time Series Analysis by State Space Methods* (2nd ed.), Oxford, UK: Oxford University Press.

Findley, D. F. (2005), "Some Recent Developments and Directions in Seasonal Adjustment," *Journal of Official Statistics*, 21, 343–365.

Gomez, V., and Maravall, A. (1994), "Estimation, Prediction, and Interpolation for Nonstationary Series with the Kalman Filter," *Journal of the American Statistical Association*, 89, 611–624.

Gould, P. G., Koehler, A. B., Ord, J. K., Snyder, R. D., Hyndman, R. J., and Vahid-Araghi, F. (2008), "Forecasting Time Series With Multiple Seasonal Patterns," *European Journal of Operational Research*, 191, 207–222.

Harvey, A. C. (1989), *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge, UK: Cambridge University Press.

Holt, C. C. (1957), "Forecasting Trends and Seasonals by Exponentially Weighted Moving Averages," *O.N.R. Memorandum*, 52/1957, Carnegie Institute of Technology.

Huang, J. Z., Shen, H., and Buja, A. (2008), "Functional Principle Components Analysis via Penalized Rank One Approximation," *Electronic Journal of Statistics*, 2, 687–695.

Huang, J. Z., Shen, H., and Buja, A. (2009), "The Analysis of Two-Way Functional Data Using Two-Way Regularized Singular Value Decompositions," *Journal of the American Statistical Association*, 104, 1609–1620.

Hyndman, R. J., and Fan, S. (2010), "Density Forecasting for Long-Term Peak Electricity Demand," *IEEE Transactions on Power Systems*, 25, 1142–1153.

Hyndman, R. J., Koehler, A. B., Ord, J. K., and Snyder, R. D. (2008), *Forecasting with Exponential Smoothing – The State Space Approach*, Berlin: Springer.

Hyndman, R. J., Koehler, A. B., Snyder, R. D., and Grose, S. (2002), "A state space framework for automatic forecasting using exponential smoothing methods," *International Journal of Forecasting*, 18, 439–454.

Koopman, S. J. (1993), "Disturbance Smoother for State Space Models," *Biometrika*, 80, 117–126.

Koopman, S. J., and Durbin, J. (2003), "Filtering and Smoothing of State Vector for Diffuse State-Space Models," *Journal of Time Series Analysis*, 24, 85–98.

Koopman, S. J., and Ooms, M. (2003), "Time Series Modelling of Daily Tax Revenues," *Statistica Neerlandica*, 57, 439–469.

Koopman, S. J., and Ooms, M. (2006), "Forecasting Daily Time Series Using Periodic Unobserved Components Time Series Models," *Computational Statistics and Data Analysis*, 51, 885–903.

Koopman, S. J., Ooms, M., and Carnero, M. A. (2007), "Periodic Seasonal Reg-ARFIMA-GARCH Models for Daily Electricity Spot Prices," *Journal of the American Statistical Association*, 102, 16–27.

Ladiray, D., Palate, J., Mazzi, G. L., and Proietti, T. (2018), "Seasonal Adjustment of Daily and Weekly Data," in *Handbook on Seasonal Adjustment*, eds. G. L. Mazzi, D. Ladiray and D. A. Rieser, Luxembourg: Publications Office of the European Union, pp. 757–783.

Lin, W., Huang, J. Z., and McElroy, T. (2020), "Time Series Seasonal Adjustment Using Regularized Singular Value Decomposition," *Journal of Business and Economic Statistics*, 38, 487–501.

Ljung, G. M. (1993), "On Outlier Detection in Time Series," *Journal of the Royal Statistical Society: Series B (Methodological)*, 55, 559–567.

McElroy, T. S. (2008), "Matrix Formulas for Nonstationary ARIMA Signal Extraction," *Econometric Theory*, 24, 1–22.

McElroy, T. S. (2017), "Multivariate Seasonal Adjustment, Economic Identities, and Seasonal Taxonomy," *Journal of Business and Economic Statistics*, 35, 611–625.

McElroy, T. S., and Maravall, A. (2014), "Optimal Signal Extraction with Correlated Components," *Journal of Time Series Econometrics*, 6, 237–273.

McElroy, T. S., and Monsell, B. C. (2017), "Issues Related to the Modeling and Adjustment of High Frequency Time Series," *Research Report*, 2017-08, U.S. Census Bureau.

McElroy, T. S., and Trimbur, T. (2015), "Signal Extraction for Non-Stationary Multivariate Time Series with Illustrations for Trend Inflation," *Journal of Time Series Analysis*, 36, 209–227.

McElroy, T. S., Monsell, B. C., and Hutchinson, R. J. (2018), "Modeling of Holiday Effects and Seasonality in Daily Time Series," *Research Report*, 2018-01, U.S. Census Bureau.

Morf, M., and Kailath, T. (1975), "Square-Root Algorithms for Least-Squares Estimation," *IEEE Transactions on Automatic Control*, 20, 487–497.

Morf, M., Sidhu, G. S., and Kailath, T. (1974), "Some New Algorithms for Recursive Estimation in Constant, Linear, Discrete-Time Systems," *IEEE Transactions on Automatic Control*, 19, 315–323.

Ollech, D. (2018), "Seasonal Adjustment of Daily Time Series," *Discussion Paper*, 41/2018, Deutsche Bundesbank.

Ord, J. K., Koehler, A. B., and Snyder, R. D. (1997), "Estimation and Prediction for a Class of Dynamic Nonlinear Statistical Models," *Journal of the American Statistical Association*, 92, 1621–1629.

Proietti, T. (2000), "Comparing seasonal components for structural time series models," *International Journal of Forecasting*, 16, 247–260.

Rodríguez, G. M., and Hernández, J. J. C. (2010), "Splines and the proportion of the seasonal period as a season index," *Economic Modelling*, 27, 83–88.

Shiskin, J., Young, A. H., and Musgrave, J. C. (1967), "The X-11 Variant of the Census Method II Seasonal Adjustment Program," *Technical Paper*, 15, U.S. Department of Commerce, Bureau of the Census.

Taylor, J. W. (2003), "Short-Term Electricity Demand Forecasting Using Double Seasonal Exponential Smoothing," *Journal of the Operational Research Society*, 54, 799–805.

Taylor, J. W. (2010), "Triple seasonal methods for short-term electricity demand forecasting," *European Journal of Operational Research*, 204, 139–152.

Taylor, S. J., and Letham, B. (2018), "Forecasting at Scale," *The American Statistician*, 72, 37–45.

Weinberg, J., Brown, L. D., and Stroud, J. R. (2007), "Bayesian Forecasting of an Inhomogeneous Poisson Process with Applications to Call Center Data," *Journal of the American Statistical Association*, 102, 1185–1198.

West, M., and Harrison, J. (1997), *Bayesian Forecasting and Dynamic Models* (2nd ed.), New York: Springer.

Winters, P. R. (1960), "Forecasting Sales by Exponentially Weighted Moving Averages," *Management Science*, 6, 324–342.