DEUTSCHE BUNDESBANK
EUROSYSTEM

Discussion Paper

Deutsche Bundesbank
No 31/2022

A review of some recent developments in the modelling and seasonal adjustment of infra-monthly time series

Karsten Webel

Deutsche Bundesbank, Wilhelm-Epstein-Straße 14, 60431 Frankfurt am Main,
Postfach 10 06 02, 60006 Frankfurt am Main

Tel +49 69 9566-0

Please address all orders in writing to: Deutsche Bundesbank,
Press and Public Relations Division, at the above address or via fax +49 69 9566-3077

Internet http://www.bundesbank.de

ISBN 978–3–95729–905–5
ISSN 2749–2958

Non-technical summary

Research Question

Thanks to the ongoing digital transformation, official statistics has started to collect more and more infra-monthly economic time series alongside traditional data types, and the recent COVID-19 pandemic has accelerated the demand of many users for such timely data. However, infra-monthly time series are likely to exhibit seasonal behaviour and other peculiarities not commonly seen in monthly and quarterly economic data, rendering official statistics' traditional modelling and seasonal adjustment approaches inapplicable and raising the question for alternatives.

Contribution

We discuss the most common peculiarities of infra-monthly economic time series and the new challenges they pose for model building and estimation. We give an extensive review of recent methodological advances in the modelling and seasonal adjustment of such data. This overview covers extensions of well-established approaches with long-standing traditions of treating monthly and quarterly data in official statistics as well as new developments that seemingly have barely been noticed outside academia so far. We illustrate key peculiarities and selected advances using daily realised electricity consumption and hourly counts of TARGET2 customer transactions in Germany.

Results

We provide a systematic overview of the multitude of new approaches to the seasonal adjustment of infra-monthly economic time series that have been developed in recent years. Our illustrations suggest that extensions of several approaches used conveniently for treating monthly and quarterly data provide solid seasonal adjustments of infra-monthly time series. However, more research is needed to achieve permanent acceptance of these and other new methods in official statistics, and we finally share some thoughts on potential areas for future developments.

Nichttechnische Zusammenfassung

Fragestellung

Im Zusammenhang mit der anhaltenden Digitalisierung wurden zuletzt in der amtlichen Statistik neben den traditionellen Datentypen auch immer mehr untermonatliche ökonomische Zeitreihen gesammelt, wobei der jüngste Ausbruch der COVID-19-Pandemie die Nachfrage vieler Nutzer nach solchen zeitnah verfügbaren Daten nochmals erhöhte. Untermonatliche Zeitreihen weisen jedoch häufig saisonale Muster und andere Charakteristika auf, die in dieser Form in monatlichen und vierteljährlichen ökonomischen Zeitreihen für gewöhnlich nicht zu beobachten sind. Traditionelle Modellierungs- und Saisonbereinigungsansätze der amtlichen Statistik sind damit auf sie nicht anwendbar und es stellt sich die Frage nach geeigneten Alternativen.

Beitrag

Wir stellen häufig zu beobachtende Charakteristika untermonatlicher ökonomischer Zeitreihen sowie die sich aus ihnen ergebenden Probleme bezüglich Modellwahl und Modellschätzung vor. Zudem geben wir einen ausführlichen Überblick über jüngste methodische Fortschritte bei der Modellierung und Saisonbereinigung solcher Daten. Dieser Überblick deckt Weiterentwicklungen der traditionell für Monats- und Quartalsdaten verwendeten Ansätze in der amtlichen Statistik sowie neue und außerhalb der akademischen Welt bisher eher wenig beachtete Ansätze ab. Zwecks Illustration sowohl der eingangs genannten Charakteristika als auch ausgesuchter methodischer Fortschritte betrachten wir den täglich realisierten Stromverbrauch und die stündliche Anzahl an TARGET2-Kundentransaktionen in Deutschland.

Ergebnisse

Wir liefern einen systematischen Überblick über die Vielzahl neuer und auf untermonatliche Zeitreihen anwendbarer Methoden zur Saisonbereinigung, die in den letzten Jahren entwickelt wurden. Unsere Illustrationen lassen den Schluß zu, dass die Weiterentwicklungen mehrerer in der amtlichen Statistik traditionell für Monats- und Quartalsdaten verwendeten Ansätze zufriedenstellende Saisonbereinigungen untermonatlicher Zeitreihen liefern. Nichtsdestotrotz ist weitere Entwicklungsarbeit erforderlich, um diese und andere neue Methoden in der amtlichen Statistik zu etablieren, und wir schließen mit einigen Gedanken zu möglichen zukünftigen Forschungsfeldern.

Bundesbank Discussion Paper No 31/2022

A review of some recent developments in the modelling and seasonal adjustment of infra-monthly time series*

Karsten Webel
Deutsche Bundesbank

Abstract

Infra-monthly time series have increasingly appeared on the radar of official statistics in recent years, mostly as a consequence of a general digital transformation process and the outbreak of the COVID-19 pandemic in 2020. Many of those 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 mainly since data of such temporal granularity exhibit stylised facts 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 periodicities and potential interactions as well as small sample issues, such as missing values and a high sensitivity to outliers. We provide an overview of recent modelling and seasonal adjustment approaches that are capable of handling these distinctive features, or at least some of them. Hourly counts of TARGET2 customer transactions and daily realised electricity consumption in Germany are discussed for illustrative purposes.

Keywords: official statistics, seasonality, signal extraction, time series decomposition, unobserved components.

JEL classification: C01, C02, C14, C18, C22, C40, C50.

*Contact address: Karsten Webel, Deutsche Bundesbank, Central Office, Statistics Department and Research Centre, Wilhelm-Epstein-Strasse 14, 60431 Frankfurt am Main, Germany. Phone: +49 69 9566 2702. E-mail: karsten.webel@bundesbank.de. This study greatly benefited from the work of the Bundesbank's research group "Extensions JD+", which ran from January 2019 to October 2020. The author therefore thanks 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. He also thanks 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, Brian Monsell of the U.S. Bureau of Labor Statistics for sharing R code related to X-13 routines, Marc Glowka, Alexander Müller and Stefan Mitzlaff for providing the TARGET2 data and further insights, and Edgar Brandt and Robert Kirchner for valuable comments.

1 Introduction

The compilation and analysis of infra-monthly economic data in official statistics can be traced back to the mid-nineteenth century. For example, Babbage (1856) studied the day-of-the-week and day-of-the-month patterns in daily monetary data and conducted what might be seen as one of the first seasonal adjustments ever as he removed respective averages from the data. In the postwar twentieth and early twenty-first centuries, however, economic measurement models in official statistics have been centred around monthly, quarterly and annual household and business sample surveys. The recent emergence of new information technologies and automated data collection methods may change this paradigm as it has drawn official statistics' attention slowly but surely to unconventional sources that provide timely access to data of unprecedented spatial and temporal granularity. Prime examples include point-of-sale and credit/debit card transactions, newspaper and social media data and various types of traffic data. Accompanied by a general discussion about official statistics' role in a digital society, statistical agencies across the world have explored ways to replace or complement traditional measurement models with new models that are capable of blending survey data with administrative and digital data sources (Jarmin, 2019; Radermacher, 2019, 2020). Despite relatively short histories and sometimes rather experimental characters, such infra-monthly data has already been used to improve early estimation of key economic indicators (Aprigliano, Ardizzi, and Monteforte, 2019; Askitas and Zimmermann, 2011; Dickopf, Janz, and Mucha, 2019), to give modelling of moving holiday effects in monthly data subtle nuances (Czaplicki and Hutchinson, 2020; Liou, Lin, and Peng, 2012), and to analyse the impact of uncertainty shocks on economic activity in real-time (Nyamela, Plakandaras, and Gupta, 2020).

The recent outbreak of the COVID-19 pandemic has accelerated this general transformation process, primarily as a result of official statistics' endeavours to satisfy the demand of many users for more timely data in order to monitor pandemic-induced disruptions and recoveries both nationwide and within specific economic sectors and geographical regions (Biancotti, Rosolia, Veronese, Kirchner, and Mouriaux, 2021; Tissot and De Beer, 2020). Infra-monthly economic data has been used, for example, to implement—often under the "experimental statistics" label—dashboards as well as early warning and sentiment indicators (Aguilar, Ghirelli, Pacce, and Urtasun, 2021; Alvarez and Lein, 2020; Coffinet, Delbos, Kaiser, Kien, Kintzler, Lestrade, Mouliom, Nicolas, Bricongne, and Meunier, 2020; Eckert and Mikosch, 2020; Eraslan and Götz, 2021; Fenz and Stix, 2021; Keane and Neal, 2021; Lewis, Mertens, Stock, and Trivedi, 2022; Lourenço and Rua, 2021; Seiler, 2020; Wegmüller, Glocker, and Guggia, 2021; Woloszko, 2020), sometimes blending economic measurements especially with Google search data. Another strand of research has been concerned with trend extraction and short-term forecasting in medical data, especially in daily infected, recovered and deceased COVID-19 cases (Doornik, Castle, and Hendry, 2022; Lee, Liao, Seo, and Shin, 2021; Li and Linton, 2021).

Although the latter medical studies employ unobserved-components decompositions that are well-known in official statistics, seasonality is not an issue there as models are often built upon short (sliding) data spans, sometimes shorter than two weeks, and thus day-of-the-week effects are typically considered negligible. However, longer—but still relatively short—histories of at least a few years are usually available in economics and therefore seasonality is no stranger to infra-monthly socio-economic data. Overall, the

latter are likely to affect established data collection methods as well as data processing and dissemination steps in the statistical production chains and thus pose an increasing number of challenges for statistical agencies. One of the key challenges in the processing step is the modelling and seasonal adjustment of such time series. Theoretical considerations often suggest that infra-monthly observations genuinely display stylised facts that are also inherent but not directly measurable from less granular data, such as official statistics' traditional headline indicators: monthly and quarterly economic time series. As a consequence, established modelling and seasonal adjustment approaches are likely to become inapplicable to infra-monthly data.

The purpose of this paper is thus to summarise common features of infra-monthly economic time series (Section 2) and to provide a thorough review of recent developments in the modelling and seasonal adjustment of such data (Section 3). This overview can also be seen as a partial update of the most recent yet slightly dated review given in Findley (2005). Selected modelling and seasonal adjustment approaches are then applied to daily realised electricity consumption in Germany and hourly transaction counts from the TARGET2 system for the purpose of illustration (Section 4), followed by some final remarks and conclusions (Section 5).

2 Data peculiarities

Let $\{y_t\}$ be a discrete time series with a seasonal periodicity of τ, that is a long-term average of τ observations per year. For example, $\tau = 12$ for monthly data and $\tau = 365.2425$ for daily data, assuming—here and throughout the rest of the paper—the 400-year cycle of the Gregorian calendar. In official statistics, $\{y_t\}$ is commonly categorised as a higher-frequency (HF) time series if data is observed at infra-monthly intervals ($\tau > 12$) and as a lower-frequency (LF) time series if data is observed at monthly or lower periodicities ($\tau \leq 12$). We first catalogue common features of HF time series (Section 2.1) and then summarise key challenges which these features pose in terms of modelling (Section 2.2).

2.1 Stylised facts

This section synthesises research on various types of HF data and attempts to categorise key stylised facts according to the underlying time series dynamics they are most likely associated with. The non-exhaustive list of related literature includes studies on:

- minute-by-minute (Taylor, 2008),
- half-hourly (Cottet and Smith, 2003; Taylor, 2003, 2010a,b; Taylor, de Menezes, and McSharry, 2006; Taylor and McSharry, 2007; Taylor and Snyder, 2012),
- hourly (Alonso, García-Martos, Rodríguez, and Sánchez, 2011; Cancelo, Espasa, and Grafe, 2008; Dordonnat, Koopman, Ooms, Dessertaine, and Collet, 2008; Gould, Koehler, Ord, Snyder, Hyndman, and Vahid-Araghi, 2008; Harvey and Koopman, 1993; Liu, Chen, Liu, and Harris, 2006; Monteiro, Menezes, and Silva, 2017; Martín-Rodríguez and Cáceres-Hernández, 2005; Mestekemper, Windmann, and Kauermann, 2010; Ramanathan, Engle, Granger, Vahid-Araghi, and Brace, 1997; Soares and Medeiros, 2008; Soares and Souza, 2006),

- 4-hourly (Pedregal and Young, 2006),
- daily (Cabrero, Camba-Mendez, Hirsch, and Nieto, 2009; Cancelo and Espasa, 1996; Cox, Triebel, Linz, Fries, Flores, Lorenz, Ollech, Dietrich, LeCrone, and Webel, 2020; De Livera, Hyndman, and Snyder, 2011; Hyndman and Fan, 2010; Koopman and Ooms, 2003, 2006; Koopman, Ooms, and Carnero, 2007; Ladiray, Palate, Mazzi, and Proietti, 2018; McElroy and Monsell, 2017; McElroy, Monsell, and Hutchinson, 2018; Ollech, 2021; Puindi and Silva, 2021; Webel, 2020; Weinberg, Brown, and Stroud, 2007), and
- weekly data (Cleveland, Evans, and Scott, 2018; Cleveland and Scott, 2007; Farley and O'Brien, 1987; Harvey, Koopman, and Riani, 1997; Pierce, Grupe, and Cleveland, 1984).

Hourly counts of customer payments (HCP) in the trans-European automated real-time gross settlement express transfer (TARGET2) system[1] in Germany are used to illustrate selected stylised facts (Figure 1).

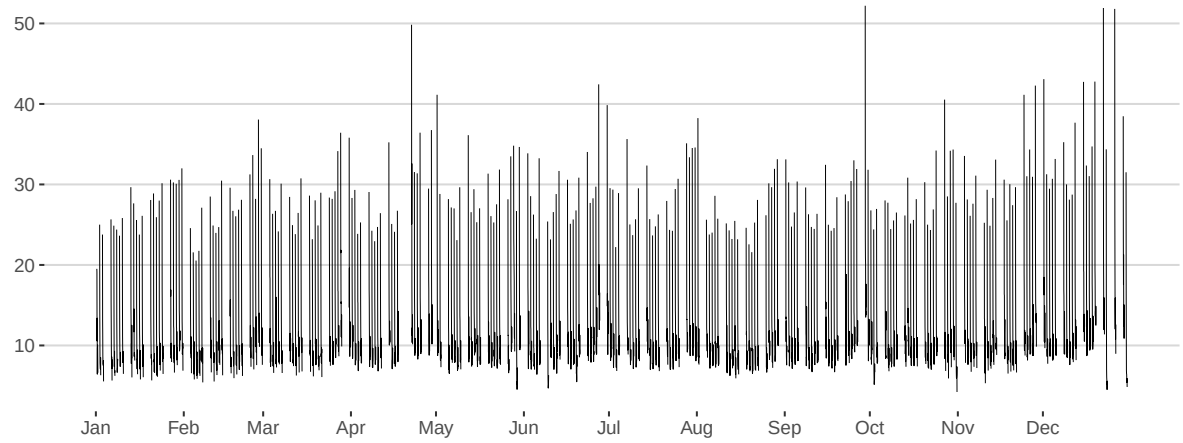
Irregular variation. Due to their granular nature, HF time series are prone to various irregularities that are well-known from small sample theory. Those include zero or close-to-zero observations and missing values, which may result from some deficiency in the reporting process. Gaps may also occur rather naturally as some data generating processes inevitably distribute observations non-uniformly across time. Technically, such gaps do not count as missing values as measurements are not expected there and thus cannot be missing. Data containing such gaps is said to be irregularly spaced, or non-equidistant. The HCP series shown in Panel (a) is a prime example as transactions are recorded only during the business hours for customer payments on TARGET2 business days, that is between 07:00 and 17:00 on Monday through Friday excluding TARGET2 holidays which are New Year's Day, Good Friday, Easter Monday, Labour Day, Christmas Day and Boxing Day.

HF data may also display aberrant behaviour in the sense of outlying observations. Such behaviour may manifest itself in a similar manner as it does in LF data: short-lived additive outliers, transitory changes, and persistent level shifts. However, it seems that HF data is also susceptible to new forms of transitional outliers with ramp-like, triangular, wavy, or otherwise smooth signatures. As a general rule of thumb, the complexity of such signatures tends to increase with the periodicity of the data. For example, a single peak in daily sales may display a U-shaped signature in the corresponding hourly data due to infra-daily consumer habits.

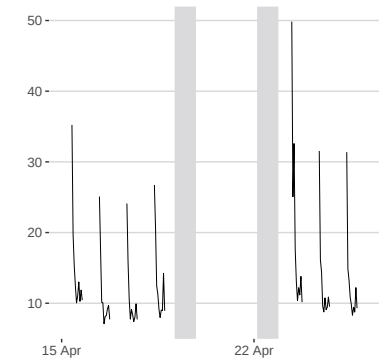
Overall, HF data tends to be subject to several irregularities that may easily add up to very volatile dynamics and therefore seems to be at higher risk of heteroskedasticity compared to LF data.

[1]TARGET2 is the real-time gross settlement system owned and operated by the Eurosystem, which comprises the European Central Bank and the national central banks of those countries that have adopted the euro and is the monetary authority of the euro area. It settles, in central bank money, customer and interbank payments as well as payments related to the Eurosystem's monetary policy operations and operations of other financial market infrastructures.

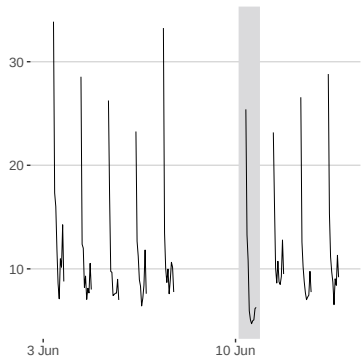
(a) Entire data span: 2 January to 31 December 2019



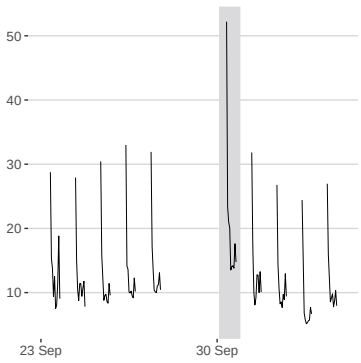
(b) Good Friday, Easter Monday



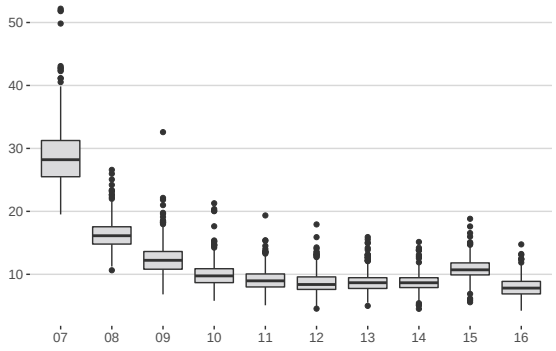
(c) Pentecost Monday



(d) End of Q3



(e) Infra-daily pattern



(f) Infra-weekly pattern: 07:00–08:00

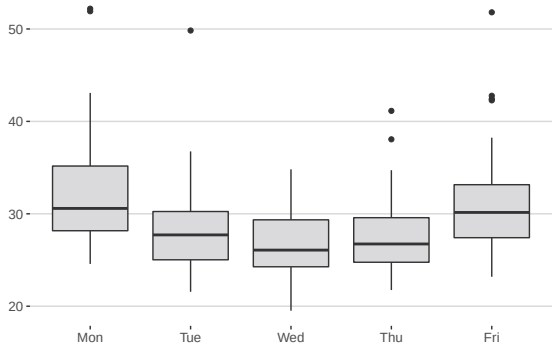


Figure 1: Hourly counts of TARGET2 customer payments in Germany (thousands). Grey verticals in Panels (b)–(d) indicate the calendar event(s) given in the subtitles.

Calendar variation. According to the above categorisation of HF and LF data, the monthly periodicity forms a natural boundary between the two types. Since the length of months is not constant over time, the number of HF observations naturally varies across LF periodicities. For example, a regularly spaced daily time series has 28 observations in non-leap-year Februaries, 29 observations in leap-year Februaries and 30 or 31 observations in the other months. This effect is referred to as the length-of-lower-frequency-periodicities (LOLFP) effect. As opposed to LF data,[2] it does not affect directly the HF observations but their seasonal periodicities (see discussion below).

In HF and especially (infra-)daily time series, calendar-related dynamics can often be measured directly with pinpoint accuracy. This applies particularly to the effects of fixed and moving holidays as well as related events, such as bridging days, including anticipatory and catch-up effects. Panels (b)–(d) exemplify different facets of such effects in the infra-daily dynamics of the HCP series: Panel (b) reveals an unusually high "early-bird" and a usual "later-bird" activity immediately after the Easter-related TARGET2 holidays, which most likely results from unprocessed orders that have accumulated over a stretch of four consecutive non-TARGET2 business days; Panel (c) shows a consistently lower activity with a noticeably smoother infra-daily profile on Pentecost Monday, which is a public holiday but still a TARGET2 business day; Panel (d) reveals a higher activity with usual infra-daily dynamics throughout the entire final day of the third quarter.

The previous examples already allude to the flip side of data granularity. In general, as more and more effects become measurable, the probability of interactions among them usually increases. This is particularly true for calendar and seasonal effects. For example, fixed-holiday and end-of-period effects may depend on the particular days of the week which the corresponding events fall onto. Christmas effects may be noticeably different for 24 to 26 December falling onto Tuesday through Thursday, as in Panel (a), versus Friday through Sunday. The same applies to the short-lived end-of-Q3 elevation in level if the final day of that quarter had not been a Monday, as in Panel (d).

Calendar-related dynamics may become even more complicated when 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 solar Gregorian calendar and, for example, the lunar Hijri calendar, which is approximately 11 days shorter.

Seasonal variation. The seasonal dynamics of HF time series are often generated by a set of superimposed seasonal patterns, each of which is composed of several seasonal cycles associated with the pattern's fundamental seasonal frequency and its harmonics. For example, Panel (e) shows that the payment counts tend to peak at the early business hours and flatten out during the rest of the day, except for a small yet visible increase between 15:00 and 16:00. This hour-of-the-day pattern contains five seasonal cycles associated with movements that occur once, twice and three to five times a business day. Panel (f) reveals that the "early-bird" effect associated with the first business hour is more pronounced on Mondays and Fridays. Similar U-type shapes can be observed for the other business hours, which in sum constitute the hour-of-the-week pattern.

[2]For example, the length-of-month effect in monthly flows can be decomposed into a constant level effect, a non-seasonal leap-year effect and a seasonal effect (Bell, 1984), which are usually modelled with the aid of preadjustment factors or regression variables.

Some of these patterns have a fractional periodicity, as opposed to LF data and mainly as a consequence of the LOLFP effect. Seasonal periodicities of infra-weekly and shorter patterns will still be integers but those of the monthly, quarterly and yearly patterns will be non-integers, except for second-by-second data which, however, rarely occur in official statistics. For example, the hour-of-the-day pattern of a regularly spaced hourly time series consist of 24 hours, whereas the hour-of-the-month pattern has a long-term average duration of 730.485 hours. In such a case, the patterns' fundamental seasonal frequencies are not integer multiples of each other and therefore the patterns or, more precisely, the seasonal cycles of the two patterns are not nested, which facilitates identification at the expense of increased model complexity. Additional non-nested patterns are likely to be present if the data is subject to dual-calendar effects.

Overall, the seasonal profile of HF time series is highly complex due to the coexistence of multiple seasonal patterns with integer versus non-integer periodicities. Interactions both among those patterns and with other HF variations, especially calendar variation, are lurking underneath.

2.2 Challenges in model building

The peculiarities of HF time series complicate modelling, especially with approaches well-established in official statistics for LF data. This section therefore elaborates on the challenges of finding adequate HF data models with respect to four key steps in model building. Selected solutions are referred to on occasion, bearing in mind that those are usually not generic but depend on the particular problem at hand.

Data regularisation. Several data irregularities, such as non-equidistant observations, missing values and the LOLFP effect, suggest that HF data should be regularised first, especially when the modelling approach is incapable of dealing with fractional periodicities. For example, Koopman and Ooms (2003, 2006) transform the time axis and introduce artificial mid-month missing values in order to obtain a structural time series model for daily Dutch tax revenues with 23 banking days in each month. The missing values are then imputed straightforwardly by the Kalman filter and smoother during model estimation. Similarly, Ollech (2021) uses cubic splines to stretch the time axis so that each month contains 31 days when modelling daily currency in circulation.

Non-artificial missing values can be imputed with simpler standard techniques, such as carrying the last observation forward and linear interpolation.

Remark 1. When regularising weekly data, the phasing of weeks must be considered as well for some modelling problems. The reason is that, for example, the first Monday through Sunday period of the year is not the same in each year and thus relates to slightly different time phases relative to the seasons in adjacent years. □

Model calibration. The identification of relevant seasonal patterns may be hampered by a shortage of established tools. Spectral diagnostics seem to perform quite reliably. But many standard seasonality tests have been derived for LF data with a single seasonal pattern and an integer-valued periodicity and thus it is not obvious that those work equally well for HF data. Finding a sparse representation of the identified patterns may pose

another issue. Seasonal dummies clearly become infeasible in many instances and thus trigonometric seasonal models seem to have the edge, not least because of their capability of incorporating fractional periodicities. However, even such models can become quite large for some patterns. For example, the full trigonometric representation of the annual seasonal pattern in daily data would require about 365 parameters and, therefore, reduced representations that contain only a small subset of seasonal cycles have been considered repeatedly in such cases (e.g. Ollech, 2021; Taylor and Letham, 2018).

The correct specification of calendar-related dynamics can be a difficult task as well since data granularity enables measurement not only of fixed- and moving-holiday effects but also of effects of related events, such as bridging days or longer pre- and post-holiday phases of interest. Finding and testing appropriate HF regression variables for such a variety of effects can quickly become time-consuming even when working only with dummy variables. However, more nuanced HF regression variables are often advisable and thus should be checked as well. Candidates include impact models that capture linearly increasing and decreasing holiday effects, potentially with different slopes, and non-linear variants with time-varying slopes.

A separability issue arises from the fact that calendar variation associated with fixed holidays is closely entangled with the annual seasonal 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 variation belongs to those calendar-induced effects that are already captured by the annual seasonal pattern and thus will be extracted by the respective seasonal filter at a later stage. Tailored HF regression variables are not needed in this case.

Overall, data granularity tends to enforce larger models in terms of both components and parameters compared to LF data. This directly raises the question of how to balance model accuracy against parsimony and flexibility, keeping in mind that larger models with potentially interacting components bear a higher risk of violating the default orthogonality assumption of many LF models. Given that models with non-orthogonal components tend to be even more complex, the willingness to sacrifice some accuracy seems inevitable.

Model estimation. In the light of model complexity and especially superimposed seasonal patterns, the benefits of sequential versus simultaneous extraction of HF dynamics should be weighed up. In either case, another separability issue arises from the fact that annual seasonality can be easily confused with trend behaviour since the fundamental annual seasonal frequency is very close to zero, except for weekly data. For example, the fundamental day-of-the-year (DOY) frequency in daily data is $\omega_{\text{DOY}} = 2\pi/365.2425 \approx 0.017$. Utilising the concept of canonical atomic models, McElroy and Monsell (2017) and McElroy et al. (2018) advocate to estimate the joint trend-DOY component first and to disentangle the two dynamics afterwards with the Hodrick-Prescott filter that targets ω_{DOY}, which, however, is difficult even in fairly long time series.

Data availability brings up another dilemma. On the one hand, only a rather short history, such as a few years, of HF observations is usually available nowadays but a long history is needed to estimate reliably all HF dynamics, including potential interaction effects. On the other hand, a long data history may complicate established estimation principles, such as likelihood evaluation, and make them even infeasible in practice for some model-based approaches.

Model stability. Many analyses in official statistics are repeated on a regular basis and therefore HF models will be re-estimated with updated data more or less frequently, depending on the revision policy. In such a case, the estimated parameters will inevitably undergo revisions and those can be quite large even if only a few additional HF observations have become available. Model and parameter stability can be fostered in the long run by avoiding the inclusion of variables with borderline significant effects, especially with respect to outliers and calendar variation. Critical values and thresholds used, for example, in automatic detection procedures and routine significance tests should thus be suited to HF data, as the defaults established for LF time series may be misleading.

3 Modelling and seasonal adjustment methods

Virtually all approaches designed for the seasonal adjustment of LF data and used routinely in official statistics assume orthogonality of unobserved components (UC) and presence of a single seasonal pattern with an integer period. However, the discussion in the previous section showed that the dynamics of HF time series rarely match those LF assumptions. After extending the standard UC decomposition from the LF to the HF case (Section 3.1), this section therefore provides an overview of new methods applicable to HF data, considering both extensions of conventional LF approaches (Section 3.2) and unconventional approaches that probably have not received much attention in official statistics so far (Section 3.3). Throughout this overview, the focus will be on the specification of the seasonal and calendar variations.

3.1 Modelling framework

Let $\{y_t\}$ denote a HF 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-cyclical behaviour, that is the long-term growth path and periodic fluctuations around it with a minimum duration of one year. The sequences $\{s_t\}$ and $\{c_t\}$ capture seasonal and calendar variation and the irregular component $\{i_t\}$ absorbs stationary transient infra-yearly movements that will be modelled as 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 behaviour that can be represented by a strictly periodic function of time; moving seasonality refers to seasonal behaviour that is characterised by gradual changes in seasonal amplitude and/or phase over time; constrained seasonality refers to abrupt yet periodically recurring changes in stable and/or moving seasonality over time. The seasonal component in model (1) is then decomposed into additive seasonal patterns according to

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

where $\mathcal{S}$ is an index set and $\{s_t^{(\tau_i)}\}$ is the seasonal pattern that captures stable, moving and constrained seasonality associated with the seasonal periodicity $\tau_i \in \mathcal{S}$. Each seasonal pattern in (2) is further atomised into additive seasonal cycles according to

$$s_t^{(\tau_i)} = \sum_{j=1}^{\lfloor \tau_i/2 \rfloor} s_{j,t}^{(\tau_i)}, \tag{3}$$

where $\lfloor x \rfloor$ is the largest integer not exceeding x and $\{s_{j,t}^{(\tau_i)}\}$ is the periodic swing generated by the seasonal frequency $\lambda_j^{(\tau_i)} = 2\pi j/\tau_i$ and therefore occurs j times during a period of τ_i HF units of time.

Overall, model (1) is the traditional low-resolution decomposition of $\{y_t\}$ developed originally for LF data, whereas model (2)–(3) zooms in on the seasonal component's higher resolutions that are characteristic of HF data.

Example 1. An unconstrained stable seasonal pattern can be represented by (3) with the j-th seasonal cycles given by

$$s_{j,t}^{(\tau_i)} = \alpha_j \cos\left(\lambda_j^{(\tau_i)} t\right) + \beta_j \sin\left(\lambda_j^{(\tau_i)} t\right). \tag{4}$$

Unconstrained moving seasonality can be introduced into (4) by allowing for smooth transitions in 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 temporally decomposing the seasonal pattern, or selected seasonal cycles, according to

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

where $\mathcal{C}$ is the set of HF time units where the constraint is in effect, $\bar{\mathcal{C}}$ is the complement of $\mathcal{C}$ and $\mathbb{1}_{\mathcal{A}}(\cdot)$ is the indicator function over $\mathcal{A}$, that is $\mathbb{1}_{\mathcal{A}}(t) = 1$ if $t \in \mathcal{A}$ and $\mathbb{1}_{\mathcal{A}}(t) = 0$ otherwise. An extension to multiple constraints is straightforward. □

Remark 2. Constrained seasonality is a flexible concept applicable to a wide range of phenomena. For example, differences between seasonal dynamics in summer versus winter could be modelled with the same structural form, such as (4), but different parameters for $\{\bar{s}_t^{(\tau_i)}\}$ and $\{\underline{s}_t^{(\tau_i)}\}$. Interactions between $\{s_t^{(\tau_i)}\}$ and other seasonal patterns or calendar variations could also be modelled within (5) by using multiple constraints. □

Remark 3. The application of (4) to weekly data is not straightforward due to the phasing of weeks. A possible solution for weekly aggregates of daily data is to assume validity of model (4) for the daily data and to aggregate the trigonometric terms over the relevant days in order to obtain a weekly seasonal cycle (Cleveland and Grupe, 1983). Week-of-the-month and some holiday effects can be modelled in a similar way, using also intervention models. □

Some of the modelling approaches to be discussed in the next subsections will impose certain restrictions on the higher-resolution seasonal dynamics (2)–(3). The following assumptions are thus introduced for convenience.

Assumption 1 ($\mathfrak{I}$)**.** The seasonal component is constituted by seasonal patterns with integer-valued seasonal periodicities, that is $\tau_i \in \mathbb{N}$ for all $\tau_i \in \mathcal{S}$ in (2). □

Assumption 2 ($\mathfrak{S}$)**.** The seasonal component is constituted by a single seasonal pattern, that is $|\mathcal{S}| = 1$ in (2). □

3.2 Conventional methods

Many statistical agencies around the world conduct seasonal adjustment of LF time series with one of the following well-established approaches: the X-11 method, the ARIMA model-based (AMB) approach, and structural time series (STS) models. JDemetra+ (JD+), a software officially recommended by Eurostat and the European Central Bank for seasonal adjustment in official statistics in the European Union,[3] provides easy access to the former two approaches via a graphical user interface. The Java source code also contains extensions and modifications applicable to HF data (Ladiray et al., 2018) and an STS framework.

3.2.1 Extended X-11 approach

The modified X-11 approach sticks to the basic 4-step principle of the genuine X-11 method (Shiskin, Young, and Musgrave, 1967) as it sequentially applies linear trend and seasonal filters to the linearised observations in order to estimate the trend-cyclical, seasonal and irregular components in (1). To this end, it relies on well-known tools, such as symmetric $m \times n$ moving averages, $3 \times k$ seasonal filters with $k \in \{1, 3, 5, 9, 15\}$ plus predefined asymmetric variants and extreme value detection based on lower and upper σ-limits.

Notwithstanding these commonalities, the extended X-11 approach incorporates two key modifications to the genuine X-11 method. The first modification is the addition of various kernel-based trend extraction filters derived from weighted least squares (WLS) estimators in local polynomial regressions (Proietti and Luati, 2008). Asymmetric variants can be constructed within this framework either through direct minimisation of the WLS objective function with the data available near the endpoints, using appropriate partitions of the involved vectors and matrices, or through minimisation of a more general objective function that allows for a bias-variance trade-off in the asymmetric filters (Grun-Rehomme, Guggemos, and Ladiray, 2018). Hence, the latter minimise the mean-squared revision error at the expense of preserving only lower-degree polynomials, subject to reproduction constraints dictated by the corresponding symmetric filter. These generalisations encompass X-11's pioneering Henderson filters and Musgrave surrogates as special cases. Another implementation is the "cut-and-normalise" approach (Gasser and Müller, 1979) according to which asymmetric variants are found through dropping the unneeded weights from the symmetric filter and dividing the remaining ones by their sum.

The second modification is triggered by the LOLFP effect and fractional seasonal periodicities in (2). Recall that the actual number of HF observations per LF periodicity often varies over time. In such a case, the famous X-11 tables have ragged right edges

[3] See the institutions' joint release letter (Ref. Ares(2015)241738–21/01/2015).

and column-wise smoothing is impossible. This is especially burdensome for table D8, the matrix representation of the seasonal-irregular component. For the day-of-the-month (DOM) pattern in daily data, for instance, the row-wise number of empty entries at the right edge will range from 0 in months with 31 days to 3 in non-leap-year Februaries. Therefore, if $\tau_i \notin \mathbb{N}$, weighted averages of the nearest integer-lagged values of the seasonal-irregular component will be computed according to (7) below, and the seasonal pattern will be extracted from those averages, using the specified $3 \times k$ seasonal filters. For example, the DOM pattern in daily data, where $\tau_i = 30.44$, can be estimated from the detrended data with a symmetric 3×3 seasonal filter according to

$$\begin{aligned}\hat{s}_t &= \frac{1}{9}\left[0.88\,(\widehat{si})_{t-61} + 0.12\,(\widehat{si})_{t-60}\right] + \frac{2}{9}\left[0.44\,(\widehat{si})_{t-31} + 0.56\,(\widehat{si})_{t-30}\right] + \frac{3}{9}\,(\widehat{si})_t \\ &+ \frac{2}{9}\left[0.56\,(\widehat{si})_{t+30} + 0.44\,(\widehat{si})_{t+31}\right] + \frac{1}{9}\left[0.12\,(\widehat{si})_{t+60} + 0.88\,(\widehat{si})_{t+61}\right].\end{aligned}$$

Weighted averages for asymmetric variants and other fractional periodicities are calculated analogously.

3.2.2 Extended AMB approach

The extended AMB approach essentially translates the classical ARIMA model decomposition algorithm (Burman, 1980) to the case of fractional Airline models. Its key idea can be summarised in four steps: first, an ARIMA model is fit to the linearised observations; second, the estimated model is decomposed into canonical ARIMA models for the UCs in (1)—or, more generally, for the signal and noise—by employing a factorisation 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 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.

The classical decomposition is founded on the assumption that each UC in (1) admits an ARIMA representation. The extended AMB approach adopts this assumption but allows for fractional periodicities in the UC models for the seasonal patterns according to

$$(1 - B^{\tau_i})\, s_t^{(\tau_i)} = \varepsilon_t^{(\tau_i)}, \tag{6}$$

where $\{\varepsilon_t^{(\tau_i)}\}$ is typically white noise that is mutually uncorrelated with any other pattern's innovations and B is the backshift operator, $B^k x_t = x_{t-k}$. Its fractional variant is written as $B^{\tau_i} = B^{\underline{\tau}_i} B^{\overline{\tau}_i}$ with $\underline{\tau}_i = \lfloor \tau_i \rfloor$, $\overline{\tau}_i = \tau_i - \underline{\tau}_i$ and $B^{\overline{\tau}_i} \approx (1 - \overline{\tau}_i) + \overline{\tau}_i B$ according to the first-order Taylor approximation at 1. Hence, B^{τ_i} is essentially replaced with a weighted average of $B^{\underline{\tau}_i}$ and $B^{\underline{\tau}_i + 1}$:

$$B^{\tau_i} \approx (1 - \overline{\tau}_i) B^{\underline{\tau}_i} + \overline{\tau}_i B^{\underline{\tau}_i + 1}, \tag{7}$$

which also defines the generalised first-order differencing operator in (6). For example, $\tau_i = 52.18$ for weekly data, so that first-order differencing becomes

$$\left(1 - B^{52.18}\right) y_t \approx y_t - (0.82\, y_{t-52} + 0.18\, y_{t-53}).$$

The UC model for $\{s_t\}$ thus follows directly from (2) and (6) using standard ARIMA theory.

The fractional Airline model also lays the base for a TRAMO-like pretreatment routine that includes estimation of user-defined regression effects and automatic detection of additive outliers, level shifts and lag-1 switch interventions, which are a special type of reallocation outliers (Wu, Hosking, and Ravishanker, 1993) defined by the sequence $(0, 1, -1, 0)$ around the outlier date. The full pretreatment model is given by

$$(1-B)\prod_{\tau_i\in\mathcal{S}}(1-B^{\tau_i})\left\{y_t-\mathbf{x}_t^\top\boldsymbol{\beta}\right\}=(1-\theta_1 B)\prod_{\tau_i\in\mathcal{S}}(1-\theta_{\tau_i}B^{\tau_i})\{\varepsilon_t\}, \tag{8}$$

where $\mathbf{x}_t$ is the vector of regression variables associated with calendar and outlier effects and $\boldsymbol{\beta}$ is the vector of unknown time-constant calendar and outlier effects. The Taylor logic utilised in (7) also applies to the seasonal MA polynomials occurring on the right-hand side of (8).

Since the fractional Airline model in (8) contains only non-stationary AR polynomials, its roots can always be allocated to the trend-cyclical and seasonal UCs. As a consequence, the irregular component in (1) does not carry transitory movements and is white noise by construction. However, the Burman-like decomposition of HF data can become unstable quite quickly. For that reason, model (8) is put into state space form (Gómez and Maravall, 1994) and estimated with a modified version of Koopman (1993)'s disturbance smoother that includes polynomial reduction and diffuse square root initialisation.

3.2.3 Structural time series models

STS and state space models have been popularised by Harvey (1989) and JD+ adopts many key concepts and standard routines described in Durbin and Koopman (2012). The general idea is to specify a model for each UC in (1) that reflects a priori beliefs about the component's dynamics. The UC models then dictate the bottom-up model for the observations and its state space representation that is finally estimated with the Kalman filter and smoother. JD+ implements classical time-varying level models, such as local linear trends, local levels with or without drift and smooth trends, as well as autoregressive and trigonometric representations of cyclical movements.

The seasonal component is modelled according to the West-Harrison (WH) representation (West and Harrison, 1997). This is a first-order vector autoregressive (VAR) model that needs validity of Assumption $\mathfrak{I}$. Letting Assumption $\mathfrak{S}$ also hold temporarily, the WH representation is either a τ-dimensional VAR model with a singular disturbance covariance matrix, which induces one cointegrating relationship and hence a stochastic zero-sum restriction in the τ seasonal effects, or a $(\tau-1)$-dimensional reparametrised model with a non-singular disturbance covariance matrix that already incorporates the above constraint explicitly. We discuss the latter form here. Let $\mathbf{s}_t^\top=(s_{1,t},\ldots,s_{\tau,t})$ be the vector of the τ seasonal effects at time t and define the permutation matrix $\mathbf{P}\in\mathbb{R}^{\tau\times\tau}$ and the dimension reduction matrix $\mathbf{D}\in\mathbb{R}^{\tau\times(\tau-1)}$ as

$$\mathbf{P}=\begin{pmatrix}\mathbf{0}_{\tau-1} & \mathbf{I}_{\tau-1}\\ 1 & \mathbf{0}_{\tau-1}^\top\end{pmatrix}\quad\text{and}\quad\mathbf{D}=\begin{pmatrix}\mathbf{I}_{\tau-1}\\ -\mathbf{1}_{\tau-1}^\top\end{pmatrix},$$

where $\mathbf{0}_n$ and $\mathbf{1}_n$ are n-dimensional column vectors of zeros and ones, respectively, and $\mathbf{I}_n$ is the n-dimensional identity matrix. The $(\tau-1)$-dimensional reduced form of the τ seasonal effects is then given by

$$\tilde{\mathbf{s}}_t = \mathbf{D}^{-}\mathbf{P}^{t-1}\mathbf{s}_t = \begin{pmatrix}\tilde{s}_{1,t} & \cdots & \tilde{s}_{\tau-1,t}\end{pmatrix}^{\top},$$

where $\mathbf{D}^{-} = \left(\mathbf{D}^{\top}\mathbf{D}\right)^{-1}\mathbf{D}^{\top} \in \mathbb{R}^{(\tau-1)\times\tau}$ is the Moore-Penrose inverse of $\mathbf{D}$, and the seasonal component is specified as

$$s_t = \mathbf{e}_{1,\tau-1}^{\top}\tilde{\mathbf{s}}_t, \tag{9}$$

where $\mathbf{e}_{k,n}$ is the k-th unit vector of length n and the reduced form of the seasonal effects is assumed to follow the VAR model

$$\tilde{\mathbf{s}}_t = \mathbf{T}\,\tilde{\mathbf{s}}_{t-1} + \boldsymbol{\omega}_t, \quad \mathbf{T} = \mathbf{D}^{-}\mathbf{P}\mathbf{D} = \begin{pmatrix}\mathbf{0}_{\tau-2} & \mathbf{I}_{\tau-2} \\ -\mathbf{1}_{\tau-1}^{\top}\end{pmatrix}, \tag{10}$$

with $\boldsymbol{\omega}_t \overset{iid}{\sim} \mathcal{N}(\mathbf{0}_{\tau-1}, \boldsymbol{\Sigma}_{\omega})$. The seasonal dynamics are thus completely specified by the choice of $\boldsymbol{\Sigma}_{\omega}$ and JD+ implements the crude, dummy, Harrison-Stevens and trigonometric models discussed in detail by Proietti (2000). Some generalisations of this approach are given in Proietti and Pedregal (2022).

Example 2. The classical form of a stochastic trigonometric seasonal model reads

$$\begin{pmatrix}s_{j,t} \\ s^{\star}_{j,t}\end{pmatrix} = \begin{pmatrix}\cos\lambda_j & \sin\lambda_j \\ -\sin\lambda_j & \cos\lambda_j\end{pmatrix}\begin{pmatrix}s_{j,t-1} \\ s^{\star}_{j,t-1}\end{pmatrix} + \begin{pmatrix}\omega_{j,t} \\ \omega^{\star}_{j,t}\end{pmatrix} \tag{11}$$

for the j-th seasonal cycle in (3), where it is usually assumed that the disturbances are distributed as $(\omega_{j,t}, \omega^{\star}_{j,t})^{\top} \overset{iid}{\sim} \mathcal{N}(\mathbf{0}_2, \sigma^2_{\omega}\,\mathbf{I}_2)$ with $\sigma^2_{\omega} > 0$ being a common variance for all seasonal cycles. The WH representation of (11) is obtained for setting $\boldsymbol{\Sigma}_{\omega} = \sigma^2_{\omega} \times \boldsymbol{\Omega}\boldsymbol{\Omega}^{\top}$ in (10), where $\boldsymbol{\Omega} = \mathbf{D}^{-}\mathbf{H}$ with $\mathbf{H}^{\top} = (\mathbf{h}_1, \ldots, \mathbf{h}_{\tau}) \in \mathbb{R}^{(\tau-1)\times\tau}$ and

$$\mathbf{h}_i^{\top} = \begin{bmatrix}\cos(\lambda_1 i) & \sin(\lambda_1 i) & \cdots & \cos\left(\lambda_{\lfloor\tau/2\rfloor} i\right) & \sin\left(\lambda_{\lfloor\tau/2\rfloor} i\right)\end{bmatrix}.$$

When τ is even, the last element is dropped from each $\mathbf{h}_i$, so that $\mathbf{h}_i \in \mathbb{R}^{\tau-1}$ holds indeed for each $i \in \{1, \ldots, \tau\}$. □

If Assumption $\mathfrak{S}$ does not hold, then each seasonal pattern in (2) is modelled according to (9)–(10) with proper adjustments in order to obtain the correct dimensions of all system vectors and matrices given $\mathcal{S}$.

The UC models of the trend-cyclical and seasonal components are finally translated into a univariate linear Gaussian state space model. This form also enables direct modelling of calendar variation by including regression variables in the observation equation. Outliers can be considered in similar fashion, and an automatic detection procedure (Grassi, Mazzi, and Proietti, 2018) has been implemented that is essentially a "forward-addition-backward-deletion" algorithm based on the point-wise maximum $\tau_t^{\star 2}$-statistics (de Jong and Penzer, 1998). Regarding model estimation, JD+ implements some non-standard features, such as the inclusion of the observation disturbances in the state vector and numerically more stable variants of the classical Kalman filtering and smoothing

techniques, such as Chandrasekhar-type recursions for time-invariant models and square root filters (Morf, Sidhu, and Kailath, 1974; Morf and Kailath, 1975), alongside diffuse (Ansley and Kohn, 1990; Koopman, 1997; Koopman and Durbin, 2003) and augmented initialisations (de Jong, 1991; de Jong and Chu-Chun-Lin, 2003).

3.3 Unconventional methods

Several alternatives to the conventional methods have been suggested over the years. Some of those have been known for quite some time but not been recognised as much in official statistics as the conventional methods, while some others are relatively new and target primarily HF data.

3.3.1 STS-type model-based approaches

The approaches presented in this section follow the same bottom-up modelling strategy as the STS approach: the UC 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, 2010b). Ord, Koehler, and Snyder (1997) and Hyndman, Koehler, Snyder, and Grose (2002) show that the innovations state space framework provides a common theoretical foundation for exponential smoothing methods. This framework also admits the general state space representation but has the key difference that both the observations and all UC models 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, Koehler, Ord, and Snyder (2008) extend the double seasonal model to the case of $|\mathcal{S}| > 2$ under validity of Assumption $\mathfrak{I}$ and the restriction that the seasonal patterns are nested. De Livera et al. (2011) further generalise 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 periodicities. TBATS is an acronym for the model's key features: trigonometric seasonal representation, Box-Cox transformation, ARMA disturbances, trend and seasonal components. The trigonometric representation of each seasonal cycle in (3) is similar to the classical STS model (11) and given by

$$\begin{pmatrix} s_{j,t}^{(\tau_i)} \\ s_{j,t}^{(\tau_i),\star} \end{pmatrix} = \begin{pmatrix} \cos\lambda_j^{(\tau_i)} & \sin\lambda_j^{(\tau_i)} \\ -\sin\lambda_j^{(\tau_i)} & \cos\lambda_j^{(\tau_i)} \end{pmatrix} \begin{pmatrix} s_{j,t-1}^{(\tau_i)} \\ s_{j,t-1}^{(\tau_i),\star} \end{pmatrix} + \begin{pmatrix} \gamma_1^{(\tau_i)} \\ \gamma_2^{(\tau_i)} \end{pmatrix} \omega_t, \tag{12}$$

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. The Box-Cox transformation could be replaced with the inverse hyperbolic sine transformation to fit data with zero

or negative observations. Non-Gaussian disturbances and modifications for multivariate time series could also be implemented within the general framework. Puindi and Silva (2021) introduce trigonometric structural models with covariates (TSCov) which integrate explanatory variables with time-invariant effects into the TBATS framework. This approach also extends well-known bootstrap techniques for improving the forecasts from the Kalman filter recursions (Cordeiro and Neves, 2009; Rodriguez and Ruiz, 2009). However, it needs to compromise the single-source-of-error concept and is currently limited to additive trends and seasonality.

Atomic seasonal models. Atomic seasonal models have been introduced by McElroy (2017) as part of a multivariate version of model (1). They are similar in spirit to the single-source-of-error approach but differ in two respects: first, the UC models are driven by individual innovations that are uncorrelated with one another and potentially collinear; second, the seasonal cycles within a seasonal pattern do not follow the same model as in (12). Instead, assuming τ_i is odd, the atomic model for the j-th vector seasonal cycle in (3) is given by

$$\delta^{\lambda_j^{(\tau_i)}}(B)\,\mathbf{s}_{j,t}^{(\tau_i)} = \boldsymbol{\varepsilon}_{j,t}^{(\tau_i)}, \tag{13}$$

where $j \in \{1, \ldots, (\tau_i - 1)/2\}$, $\delta^{\lambda}(B) = 1 - 2\cos\lambda B + B^2$ is a "Gegenbauer-type" differencing operator at frequency $\lambda \in [0, \pi]$ and $\{\boldsymbol{\varepsilon}_{j,t}^{(\tau_i)}\}$ is multivariate Gaussian white noise with covariance matrix $\boldsymbol{\Sigma}_j^{(\tau_i)}$. When τ_i is even, then (13) applies only to the first $(\tau_i/2 - 1)$ vector seasonal cycles and the model for the last cycle is given by $(1 + B)\,\mathbf{s}_{\tau_i/2,t}^{(\tau_i)} = \boldsymbol{\varepsilon}_{\tau_i/2,t}^{(\tau_i)}$. The structural similarity to (6), which utilises the fractional backshift operator, is apparent in either case. However, note that $\delta^{\lambda}(B)$ yields a complete factorisation of seasonal unit root differencing polynomials. For example, the weekly differencing operator for daily data factorises as $1 - B^7 = (1 - B)\,\delta^{2\pi/7}(B)\,\delta^{4\pi/7}(B)\,\delta^{6\pi/7}(B)$.

The atomic nature of (13) allows the widths and heights of spectral seasonal 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 may occur during times of strong economic changes. In addition, the multivariate nature of the entire framework allows each UC to be common, related or unrelated across original series, depending on whether its white noise is collinear or, if not, has a non-diagonal versus diagonal covariance matrix.

However, some refinements and modifications are needed to make atomic models and the generalised Wiener-Kolmogorov (WK) matrix formulas for multivariate signal extraction (McElroy, 2008; McElroy and Trimbur, 2015) feasible for HF data, or large data sets in general. McElroy and Monsell (2017) introduce canonical variants in the spirit of Hillmer and Tiao (1982), from which all extractable white noise has already been removed. As a consequence, the differenced canonical UCs have MA-like structures instead of being white noise. McElroy and Monsell (2017) also utilise the generalised Cholesky decomposition to derive invertible method-of-moments (MOM) estimators that ensure positive definiteness of all estimated covariance matrices. Forecasts of the observations can then be generated from the MOM-estimated data model—with fixed outlier and calendar effects being univariately removed beforehand—, using recursive one-step ahead predictions provided by the Durbin-Levinson algorithm. UC estimates are finally obtained from running the generalised WK filter on the forecast-extended observations. Due to a separability

issue, the trend and annual seasonal components are first extracted as a joint component that is eventually separated with the Hodrick-Prescott filter. Applications to daily data are discussed in McElroy and Monsell (2017) and McElroy et al. (2018).

3.3.2 Generalised Gegenbauer processes

Gegenbauer processes (Gray, Zhang, and Woodward, 1989) and their generalisations (Giraitis and Leipus, 1995; Woodward, Cheng, and Gray, 1998) provide a flexible way of modelling time series with trend and seasonal, or cyclical, long memory. Recall that the atomic seasonal models (13) already utilise the filter $\left[\delta^{\lambda}(B)\right]^{-d}$, which is the generating function of the orthogonal Gegenbauer polynomials, with $d = 1$ at each seasonal frequency and white noise innovations. Gegenbauer processes allow for fractional orders of integration, different orders of integration at different frequencies, and any short memory innovations. Ignoring calendar variation, the k-factor Gegenbauer process is given by

$$\prod_{j=1}^{k} \left(1 - 2\cos\lambda_j B + B^2\right)^{d_j} y_t = \varepsilon_t, \tag{14}$$

where $\lambda_j \in [0, \pi]$ can be related to the trend or any seasonal pattern in (1), d_j is the memory parameter associated with λ_j, and $\{\varepsilon_t\}$ is a linear short memory process (often a finite-order ARMA process). Model (14) is stationary and invertible, if $|\, d_j \,| < 1/2$ for all $\lambda_j \in (0, \pi)$ and $|\, d_j \,| < 1/4$ for all $\lambda_j \in \{0, \pi\}$, and it encompasses standard fractionally integrated processes, such as ARFIMA processes and seasonal variants, as special cases. The spectral density of a Gegenbauer process $\{y_t\}$ takes the form

$$f_y(\lambda) = f_{\varepsilon}(\lambda) \prod_{j=1}^{k} |2\left(\cos\lambda - \cos\lambda_j\right)|^{-2d_j}, \tag{15}$$

where $f_{\varepsilon}(\cdot)$ is the spectral density of the short memory process. Hence, (15) can have multiple poles/zeros at arbitrary locations and of arbitrary shapes: the locations are specified through the λ_j's and the shapes are governed by d_j's, where $d_j > 0$ corresponds to a pole (i.e. long memory behaviour) and $d_j < 0$ corresponds to a zero (i.e. intermediate, or negative, memory behaviour).

The autocovariances of (14) can be computed either from the Wold decomposition (McElroy and Holan, 2012) or from utilising the so-called "splitting method"—which essentially convolves the long memory and short memory dynamics—after an additive decomposition of (15) such that each summand has a single pole/zero (McElroy and Holan, 2016). Assuming a generalised exponential model, either approach employs a cepstral representation of the short memory dynamics, and the latter is akin to a partial fraction decomposition, which is utilised in the classical and extended AMB approaches (Section 3.2.2). For some special cases of the seasonal fractionally differenced exponential model, Holan and McElroy (2012) develop a block Metropolis-Hastings algorithm for model estimation within a fully Bayesian framework, and finite-sample MMSE signal extraction formulas (McElroy, 2008) for seasonal adjustment, which also enables the inclusion of calendar regression variables in a straightforward way.

In many applications, the problem at hand will suggest a proper choice of the fre-

quencies λ_j to be considered in (14), and hence the model order k. In some applications, however, the choice of k might be less clear a priori, and an automatic selection procedure could be warranted. Leschinski and Sibbertsen (2019) provide such a procedure, which is based on sequential tests for the maximum of the spectral density of the iteratively Gegenbauer-filtered data and illustrated for hourly electricity loads. The latter type of HF data has also been analysed by Soares and Souza (2006) with the aid of periodic regressions that contain dummy variables for weekly seasonality and holiday variation and 2-factor Gegenbauer errors that capture trend behaviour and annual seasonality. Another strand of application is the modelling of volatility in daily and infra-daily financial time series (e.g. Asai, McAleer, and Peiris, 2020; Bisaglia, Bordignon, and Lisi, 2003; Bordignon, Carporin, and Lisi, 2007). In this context, Voges and Sibbertsen (2021) introduce a bivariate extension of model (14) and seasonal multiple local Whittle estimation in order to study cyclical fractional cointegration in half-hourly trading volume and realised volatility of the component stocks of the Dow Jones Industrial Average index.

3.3.3 STL-based approaches

Cleveland, Cleveland, McRae, and Terpenning (1990) have propagated a seasonal-trend decomposition based on LOESS regressions (STL) as an alternative to the genuine X-11 method for LF data with $\tau \in \{4, 12\}$. This procedure also relies on iterative filtering under Assumption $\mathfrak{S}$ but can handle any integer $\tau > 1$ and missing values. It essentially consists of an inner loop responsible for sequential UC estimation and refinement and an outer loop responsible for extreme value correction. Two variants have been developed recently, each of which facilitates sequential extraction of multiple seasonal patterns in (2) under Assumption $\mathfrak{I}$, an idea already alluded to by Cleveland et al. (1990).

The {`rjd3highfreq`} package. This package basically provides access to the Java translation of the original FORTRAN and S implementations of STL. Thus, pretreatment model (8) can be used for adjusting outlying observations as an alternative to the outer loop. Additive and multiplicative forms of the general UC model (1) are available and—calling the STL routine $|\mathcal{S}|$ times—the numbers of neighbouring observations to be considered in the inner-loop local regressions for trend and seasonal smoothing can be specified. Non-integer seasonal periodicities will be automatically rounded down to the nearest integer.

The {`dsa`} package. This packages provides an alternative integrated framework tailored to daily data with $\mathcal{S} = \{7, 31, 365\}$, in which classical regARIMA pretreatment replaces the fractional Airline model (8) and the day-of-the-week (DOW), DOM and DOY patterns are extracted sequentially from a single call of STL (Ollech, 2021). The key difference to the {`rjd3highfreq`} package is the order of operations: in {`dsa`}, the DOW pattern is estimated first from the observed data and regARIMA pretreatment is run on the DOW-adjusted data afterwards. For that reason, applying the outer-loop robust filter weights during DOW extraction is usually beneficial. In a third step, the DOM pattern is extracted from the linearised DOW-adjusted series. Cubic splines are employed to temporarily stretch-and-interpolate or extrapolate months with less than 31 days to 31 days. Both tactics are similar to row-wise interpolation in the extended X-11

approach. In a fourth step, the DOY pattern is estimated from the linearised 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 those skipped observations are finally filled with spline interpolations.

3.3.4 STR approach

Dokumentov and Hyndman (2015) discuss a procedure inspired by the genuine STL approach in which the LOESS regression is essentially replaced with a more general ridge- or LASSO-type regression model (STR), so that validity of Assumption $\mathfrak{I}$ is still required. The key conceptual difference 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. Assume that the observed time series in (1) has a finite length T, written as $\mathbf{y}_T$. Each seasonal pattern can then be expressed in matrix form as

$$\mathbf{S}^{(\tau_i)} = [s_{k,l}]_{k,l}^{(\tau_i)} \in \mathbb{R}^{\tau_i \times T}, \tag{16}$$

where each column stores τ_i elements of the seasonal pattern at time t. Only one of those elements is actually in effect and the others are hidden. Thus, (16) connects to (2) via

$$s_t^{(\tau_i)} = \mathbf{S}_{\tau_i^\star(t),t}^{(\tau_i)} \tag{17}$$

for some mapping $\tau_i^\star : \mathbb{N} \mapsto \{1, \dots, \tau_i\}$ that picks the pattern's correct season at time t in rotational fashion. This is akin to model (9), in which the vector of seasonal effects is ceaselessly shuffled and $\mathbf{e}_{1,\tau-1}$ picks the first entry at each time t.

The STR approach also enables a more nuanced specification of calendar variation as it explicitly distinguishes between three types of calendar regression variables: static predictors have constant effects as in (8), whereas 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 of 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 $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \sigma_\varepsilon^2 \boldsymbol{\Sigma})$ with $\boldsymbol{\Sigma}$ being a block matrix that carries $\mathbf{I}_T$ in the upper left block and zeros everywhere else. This model can be estimated by some maximum likelihood technique and the underlying optimisation problem can be expressed in terms of discrete second-order derivatives of the UCs, which has two consequences: first, additional smoothness restrictions can be imposed on the seasonal component in the time, time-season and season dimensions; second, Assumption $\mathfrak{I}$ can be relaxed quite naturally. To see this, note that seasonal patterns with fractional periodicities—and all other UCs as well—can be rewritten as linear combinations of smooth basis functions, such as Fourier terms, splines or wavelets. Those functional components provide a valid reparametrisation of the second-order derivative matrices in the optimisation problem and ultimately lead to a substantial reduction of computation time.

Depending on the optimisation 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_ε^2. A robust STR method is also available, which assumes the residuals $\boldsymbol{\varepsilon}$ to be Laplacian and thus leads to an optimisation problem of a quantile regression that can be solved only numerically.

3.3.5 Regularised singular value decomposition

Lin, Huang, and McElroy (2020) suggest a signal extraction approach based on regularised singular value decomposition (RSVD) (Huang, Shen, and Buja, 2008, 2009). This method rewrites the observations and UCs as matrices in STR-like fashion but the seasonal matrix stores only the visible elements in (16). Let Assumptions $\mathfrak{I}$ and $\mathfrak{S}$ hold and assume that $n = T/\tau \in \mathbb{N}$. Ignoring temporarily calendar variation, the matrix form of model (1) is given by $\mathbf{Y} = \mathbf{S} + \mathbf{N}$, where

$$\mathbf{S} = [s_{k,l}]_{k,l} \in \mathbb{R}^{n \times \tau} \tag{18}$$

is the seasonal matrix that stores the seasonal effects related to the i-th subspan in its i-th row, and $\mathbf{Y}$ and $\mathbf{N}$ denote the observation and non-seasonal matrices that are defined analogously. Accordingly, (18) can be linked to (2) via

$$s_t = \mathbf{S}_{\lceil t/\tau \rceil, t - \lfloor t/\tau \rfloor \tau},$$

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

$$\mathbf{S} = \mathbf{1}_n \mathbf{f}^\top + \mathbf{U}\mathbf{V}^\top, \tag{19}$$

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, (19) disentangles stable from moving seasonality, which is actually represented as a linear combination of r layers stored in $\mathbf{V}$ with corresponding span-varying magnitudes stored in $\mathbf{U}$.

Once the number of layers has been chosen, model estimation proceeds in three steps: first, an estimate $\hat{\mathbf{U}}$ is obtained through loop-like sequential runs of the RSVD method under a zero-sum restriction on the layers in $\mathbf{V}$; second, estimates $\hat{\mathbf{f}}$ and $\hat{\mathbf{V}}$ are obtained jointly through solving a least-squares minimisation problem 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 $\mathbf{V}^\top \mathbf{1}_\tau = \mathbf{0}_r$; third, parameter estimates related to non-seasonal dynamics are obtained, if warranted. The RSVD method imposes a roughness penalty on each column of $\mathbf{U}$ that includes a layer-specific smoothing parameter and hence opens up the possibility of taking constrained moving seasonality into account by using different smoothing parameters in different segments of a given layer. An automatic routine for detecting such segments has also been implemented.

If Assumption $\mathfrak{S}$ does not hold, then model (1) can be rewritten in terms of $|\mathcal{S}|$-dimensional arrays in order to extract multiple seasonal patterns. In addition, calendar variation can be considered by including regression variables in the constrained least-squares minimisation problem. Overall, simulations with monthly time series indicate that RSVD-based signal extraction performs well in the presence of strong seasonality and/or seasonal breaks, whereas it tends to be outperformed by the classical X-11 and AMB approaches in cases of weak seasonality.

3.3.6 Prophet

Prophet is a Bayesian approach developed by Facebook's Core Data Science team (Taylor and Letham, 2018) with the primary aim of providing a flexible and reliable forecasting

tool that can be configured, interpreted and evaluated by subject-matter experts and analysts without great expertise in time series modelling. The general idea is to specify relatively sparse UC models in (1) and impose Gaussian or Laplacian priors on the unknown parameters.

The notion of sparsity translates into a seasonal component that is composed of stable patterns in which each seasonal cycle has the trigonometric form (4) and a Gaussian prior is imposed on the amplitudes. Let

$$\mathbf{s}_t^{(\tau_i)} = \left[\cos\left(\frac{2\pi t}{\tau_i}\right), \sin\left(\frac{2\pi t}{\tau_i}\right), \cdots, \cos\left(\frac{2\pi Jt}{\tau_i}\right), \sin\left(\frac{2\pi Jt}{\tau_i}\right)\right] \tag{20}$$

and $\boldsymbol{\gamma} = (\alpha_1, \beta_1, \ldots, \alpha_J, \beta_J)^\top$, where $J = J(\tau_i)$ is a pattern-specific number of seasonal cycles. Then, each seasonal pattern in (2) is modelled 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 both non-nested seasonal patterns and fractional periodicities. In addition, concept (5) can be used to place constraints on any stable seasonal pattern.

Calendar variation is modelled with the aid of dummy indicators that can be defined via built-in lists of national and international holiday dates or specified by the user. The regression effects are then assumed to have a Gaussian prior as well. Model estimation is finally carried out using the L-BFGS optimisation algorithm (Byrd, Lu, Nocedal, and Zhu, 1995) to find the maximum a posteriori estimates.

3.3.7 CAMPLET

The CAMPLET method decomposes an observed time series into its non-seasonal and seasonal parts without the need for data pretreatment as the method's tuning parameters automatically adapt to outliers and other irregularities (Abeln and Jacobs, 2015; Abeln, Jacobs, and Ouwehand, 2019). Both the data and decomposition is assumed to be known up to time t and a new point-wise decomposition at $t+1$ is to be found as observation y_{t+1} becomes available. The notion of seasonality is akin to the matrix representation (16) and mapping (17) employed in the STR approach. However, the vectors of seasonal effects are not estimated jointly but updated sequentially at each point in time based on past information, which is similar to Kalman filtering without smoothing. Under Assumptions $\mathfrak{I}$ and $\mathfrak{S}$, let $\mathbf{s}_t = (s_{1,t}, \ldots, s_{\tau,t})^\top$ be the t-th column of $\mathbf{S}^{(\tau)}$ in (16) with $\mathbf{s}_t^\top \mathbf{1}_\tau = 0$. Then, the visible and invisible seasonal effects are assumed to evolve through time according to

$$\mathbf{s}_{t+1} = \mathbf{s}_t + \mathbf{p} \times a_{t+1}, \tag{21}$$

where $\mathbf{p} = (p_1, \ldots, p_\tau)^\top$ is a fixed vector of positions with $p_i = (\tau + 1)/2 - i$ for $i \in \{1, \ldots, \tau\}$, so that $\mathbf{p}^\top \mathbf{1}_\tau = 0$, and a_{t+1} is a time-varying adjustment scalar. The latter is defined as the ratio between the extrapolation error and the adjustment length and hence is essentially the change in the average growth of the non-seasonal component, where the average is calculated over a recent time segment that ends in t. In other words: the evolution of the seasonal effects in (21) is driven by recent non-seasonal dynamics and thus naturally encompasses interactions especially with the trend-cycle.

3.3.8 Direct filter approach

The direct filter approach (DFA) (Wildi, 2005, 2008) provides a real-time framework for estimating any target signal that is expressible as a linear function of the observations, which is not necessarily required in UC model (1). Like generalised Gegenbauer processes, this approach can handle series with multiple spectral peaks of arbitrary width and location. Let $\Gamma(B) = \sum_k \gamma_k B^k$ be a possibly bi-infinite linear filter. A target is then defined as $z_t = \Gamma(B)\, y_t$, which could be a UC, combination of UCs, or even a forecast of $\{y_t\}$. The estimated target is given by $\hat{z}_t = \hat{\Gamma}(B)\, y_t$, where $\hat{\Gamma}(B) = \sum_{k=0}^{L-1} w_k B^k$ is a concurrent linear approximation to $\Gamma(B)$ of finite length L. The prediction error can thus be written as

$$z_t - \hat{z}_t = \left[\Gamma(B) - \hat{\Gamma}(B)\right] y_t,$$

and its mean squared error (MSE) is given by

$$\mathrm{MSE}_\Gamma(\mathbf{w}, g) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\left|\Gamma\left(e^{-i\lambda}\right) - \hat{\Gamma}_{\mathbf{w}}\left(e^{-i\lambda}\right)\right|^2}{\left|\delta\left(e^{-i\lambda}\right)\right|^2} g(\lambda)\, d\lambda, \tag{22}$$

where $\mathbf{w} = (w_0, \ldots, w_{L-1})^\top \in \mathbb{R}^L$ is the vector of weights of the concurrent filter, $\delta(B)$ is the differencing polynomial needed to render $\{y_t\}$ weakly stationary, and $g(\cdot)$ is a generic real-valued non-negative function over $[-\pi, \pi]$, usually a design-specific estimator of the spectral density of $\{\delta(B)\, y_t\}$. Let $\mathbf{w}(g)$ be a solution to minimising (22) with respect to $\mathbf{w}$ (under additional assumptions on $\Gamma(B)$ and $\hat{\Gamma}(B)$ to ensure zero-mean weakly stationary prediction errors in case $\{y_t\}$ is integrated). Then $\hat{\Gamma}_{\mathbf{w}(g)}(B)$ is the optimal concurrent linear approximation within the considered class of filters for a given function $g(\cdot)$.

Remark 4. The generic specification of an "ideal" target as the outcome of an ideal band-pass filter can account for spectral peaks of arbitrary width and location in $\{y_t\}$. The DFA thus enables replication—and further customisation—of classical filter-based and model-based signal extraction results by plugging the respective target signals and spectrum estimates in (22). However, it finds the "best" target estimator through direct optimisation over the filter weights $\mathbf{w}$, whereas classical model-based approaches optimise over some underlying parameters $\boldsymbol{\theta}$. Nevertheless, $\mathbf{w}$ will be expressible as a function of $\boldsymbol{\theta}$ in some cases. □

Example 3. Let Assumption $\mathfrak{S}$ hold and $S(B) = 1 + B + \cdots + B^{\tau-1}$ be the annual aggregation operator. A naive estimator of $\{s_t\}$ in (1) is then obtained through $\Gamma(B) = 1 - \tau^{-2} S(B) S(B^{-1})$ and the periodogram for $g(\cdot)$ in (22). □

Wildi and McElroy (2016) embed the DFA into the more general framework of linear prediction problems for time series. Wildi and McElroy (2019) provide generalised optimisation criteria based on a decomposition of the frequency response functions in (22) into their amplitude and phase components, which yields three constituent error components of the MSE associated with the accuracy, timeliness and smoothness of the involved filters. McElroy and Wildi (2020) develop a multivariate DFA for stationary and non-stationary vector time series in which the determinant of the vector prediction errors' covariance matrix is minimised in place of (22). An early application to trend-cycle extraction can be found in Buss (2016).

4 Illustration

This section demonstrates the capabilities of selected seasonal adjustment approaches discussed in Section 3, without intending a qualitative comparison or even a ranking. It considers daily electricity consumption in Germany (Section 4.1) and hourly counts of TARGET2 customer payments (Section 4.2).

4.1 Daily electricity consumption

Daily realised electricity consumption is a key component of the weekly activity index that has been developed in the wake of the COVID-19 pandemic to track the German economy (Eraslan and Götz, 2021). The series covers electricity supplied to the network for the general supply—excluding electricity supplied to the railroad network and to internal industrial and closed distribution networks as well as electricity consumed by the producers—, is freely available from the Federal Network Agency ("Bundesnetzagentur | SMARD.de") under URL `https://www.smard.de/en` and is considered here in units of terawatt hours (TWh) as of 1 January 2015 up to 31 December 2021, resulting in 2,557 daily observations. The seasonal profile is described first (Section 4.1.1), followed by the linearisation of the series (Section 4.1.2). Unconstrained stable and moving seasonal patterns are then extracted from the linearised series with the three unconventional approaches (Section 4.1.3). Finally, a stable DOW pattern is estimated with Prophet, imposing a constraint on each year's volatile period between Christmas and New Year (Section 4.1.4). The results are discussed rather briefly as an extended discussion of the series is given by Webel (2020), albeit for the slightly shorter data span up to 30 August 2020.

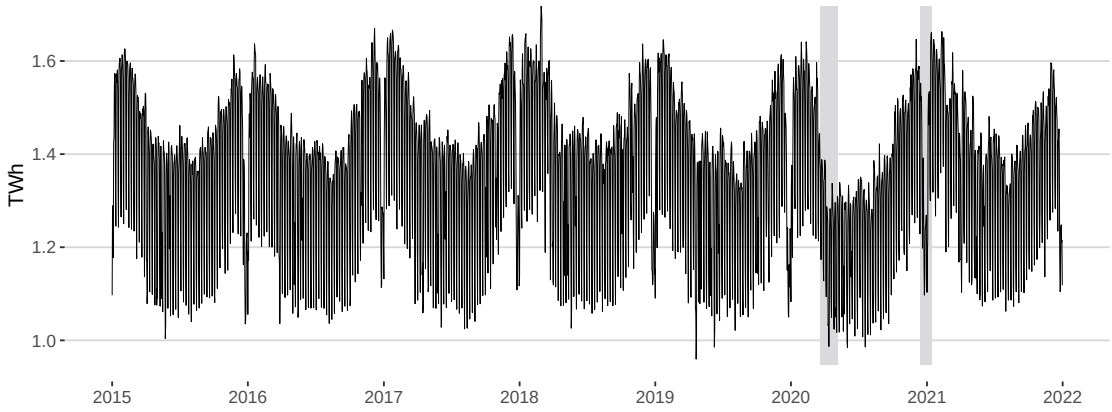
4.1.1 Seasonal profile

Figure 2 reveals key facets of the series' seasonal profile. Panel (a) shows that electricity consumption is higher in the winter and lower in the summer. Given usual temperature curves and daylight hours in Germany, this pronounced U-shaped DOY pattern may 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 first COVID-19 lockdown, which also breaks the slight upward trend in the series.

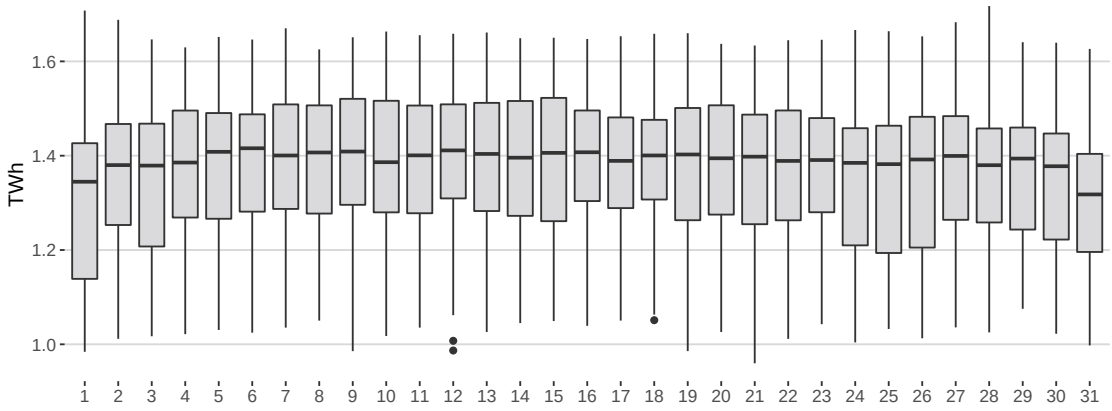
Panel (b) depicts the infra-monthly dynamics with the aid of DOM boxplots. Although electricity consumption tends to be somewhat lower around the turn of the month—partly due to the occurrence of fixed public holidays such as New Year's Day, Labour Day and All Saints' Day—and shows some signs of time-varying volatility, it 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 a recent subspan, revealing a persistent 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—regardless of the second COVID-19 lockdown—and also visibly affected by both fixed and moving holidays.

(a) Day-of-the-year (DOY) pattern: 1 January 2015 to 31 December 2021



(b) Day-of-the-month (DOM) pattern: 1 January 2015 to 31 December 2021



(c) Day-of-the-week (DOW) pattern: 1 December 2020 to 30 June 2021

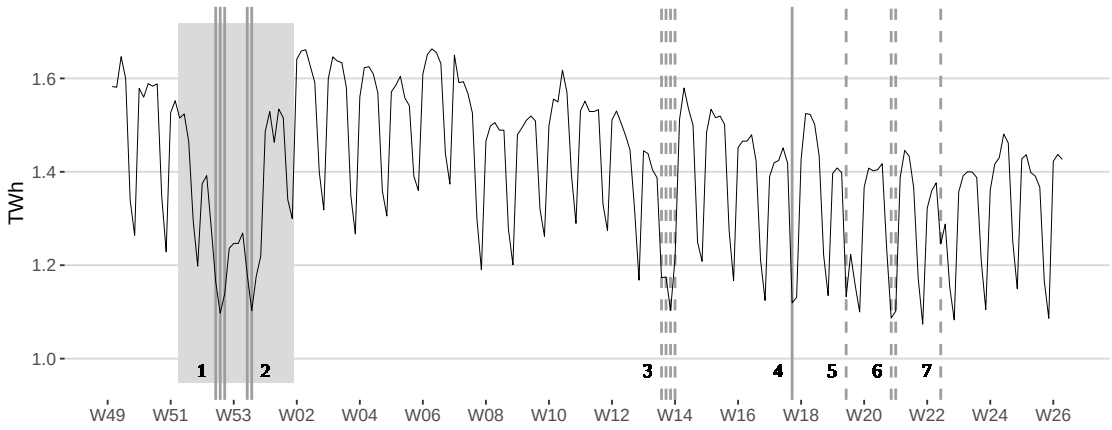


Figure 2: Seasonal profile of daily realised electricity consumption in Germany. Shaded backgrounds in Panels (a) and (c) correspond to COVID-19 lockdowns: 23 March to 3 May 2020, and 16 December 2020 to 10 January 2021. Verticals in Panel (c) correspond to fixed (*solid*) and moving (*dashed*) holidays: Christmas (1), New Year (2), Easter (3), Labour Day (4), Ascension (5), Pentecost (6), and Corpus Christi (7).

Table 1: Estimated variance ratios (23) for daily electricity consumption.

Seasonal adjustment approach	DOW pattern $\tau_i = 7$	DOM pattern $\tau_i = 30.5369$	DOY pattern $\tau_i = 365.2425$
Extended X-11	0.9431	0.1684	0.8774
Extended AMB	0.9416	0.0521	0.8775

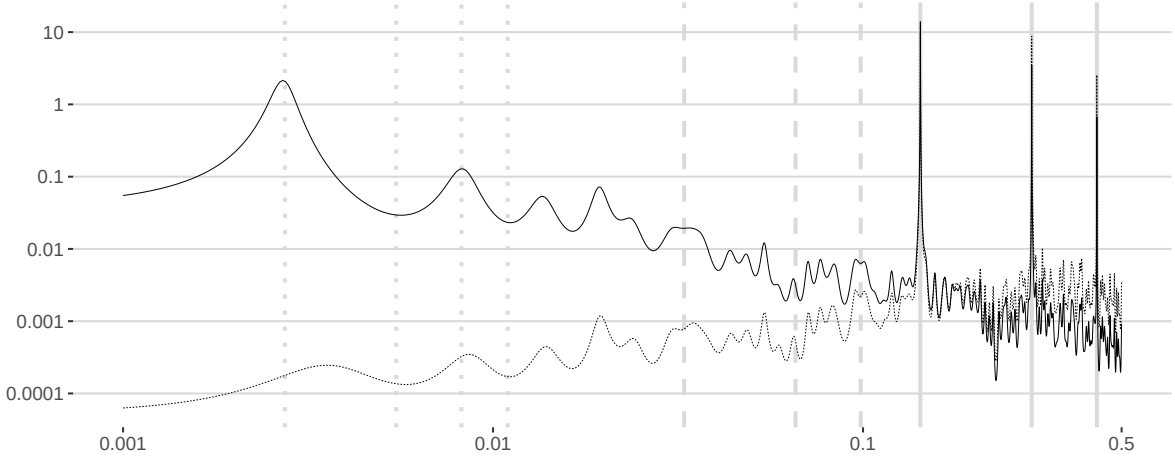


Figure 3: Estimated autoregressive spectrum of logged (*solid*) and differenced logged (*dashed*) electricity consumption, scaled by 2π. Grey verticals correspond to selected seasonal harmonics of the DOY (*dotted*), DOM (*dashed*) and DOW (*solid*) patterns.

The strength of the three seasonal patterns can be determined with the aid of the variance ratio

$$1 - \frac{\mathbb{V}(i_t)}{\mathbb{V}\left(s_t^{(\tau_i)} + i_t\right)}, \quad \tau_i \in \mathcal{S}, \tag{23}$$

which measures a pattern's portion of volatility relative to the irregular dynamics (Kang, Hyndman, and Smith-Miles, 2017; Wang, Smith, and Hyndman, 2006). A ratio close to one (zero) indicates a strong (weak) seasonal pattern. In fact, 0.5 seems to be a reasonable threshold as a lower ratio indicates that the irregular component explains more volatility in the data than the seasonal pattern. A tentative version of pretreatment model (8) is thus estimated for the full set of seasonal patterns, including the regression variables explained in the next section. Table 1 reports the estimated ratios (23) for the three seasonal patterns obtained from the extended X-11 and AMB approaches. For either approach, the DOW and DOY patterns appear strong, whereas the DOM pattern does not. Estimated autoregressive spectra confirm this dominance (Figure 3) and, hence, the DOM pattern is not considered in subsequent analyses.

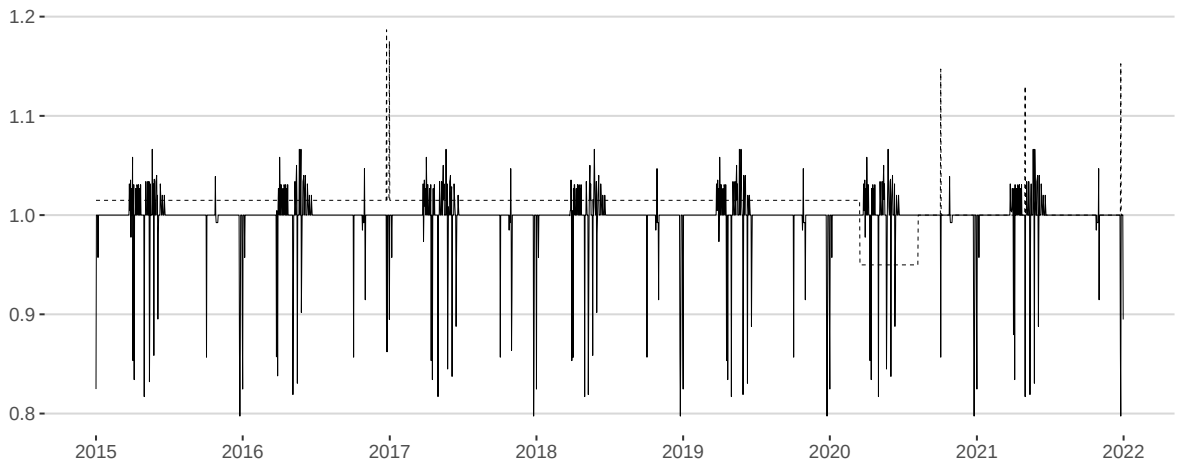


Figure 4: Estimated calendar (*solid*) and outlier (*dashed*) components of daily electricity consumption.

4.1.2 Pretreatment

We fit model (8) 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, associated bridging days and daylight saving time (DST). We also run automatic detection of additive outliers, level shifts and lag-1 switch interventions and eventually keep all calendar and outlier regression variables with absolute t-values larger than 5.[4] The only exception is the March DST dummy, which has an absolute t-value of 4.6 but is kept in the model for the sake of coherence.

Table 2 reports the estimated fractional Airline model. The estimated seasonal MA parameters exceed 0.85, reflecting again the strong DOW and DOY patterns. Five additive outliers have been automatically detected, the dates of which coincide with the Saturdays of Christmas and New Year's Eve 2016, German Unification Day 2020, Labour Day 2021 and Christmas Day 2021. Each estimated effect is positive and hence partially cancels the respective negative holiday effect, which underlines the general difficulty of estimating reliably time-constant regression effects of rare events. It may also indicate the presence of cross-dependencies between the DOW pattern and fixed holidays, which could be solved in theory by adding appropriate regression variables to model (8) but is infeasible here as the sample size is just seven years. Two level shifts have also been automatically detected with dates related to the onset and recovery phase of the first COVID-19 wave. The estimated calendar and outliers components are shown in Figure 4.

4.1.3 Seasonal adjustment

The DOW and DOY patterns are extracted sequentially from linearised electricity consumption with the extended X-11 and AMB approaches. The X-11 approach utilises a 9-term third-order Henderson kernel and a 3×9 seasonal filter for extracting the DOW

[4]This threshold accounts nicely for the effective number of observations during automatic outlier detection as it is close to 4.9, which is the X-13 default critical value at the 1% level of significance according to the implemented modified formula of Ljung (1993).

Table 2: Estimated pretreatment model (8) for logged daily electricity consumption.

Event	Date	Weight	Estimate	SE	t-value
FIXED HOLIDAYS					
New Year's Day	1 Jan	1.0	−0.193	0.010	−19.973
Epiphany	6 Jan	0.4	−0.109	0.019	−5.738
Labour Day	1 May	1.0	−0.202	0.008	−24.457
German Unification Day	3 Oct	1.0	−0.155	0.008	−18.225
500th Reformation Day	31 Oct 2017	1.0	−0.139	0.018	−7.609
All Saints' Day	1 Nov	0.7	−0.127	0.013	−10.094
Christmas Eve	24 Dec	1.0	−0.148	0.011	−13.991
Christmas Day	25 Dec	1.0	−0.226	0.012	−18.222
Boxing Day	26 Dec	1.0	−0.148	0.010	−14.999
New Year's Eve	31 Dec	1.0	−0.111	0.010	−10.680
MOVING HOLIDAYS					
Good Friday		1.0	−0.185	0.007	−25.936
Easter Monday		1.0	−0.212	0.007	−29.629
Ascension		1.0	−0.233	0.008	−27.395
Pentecost Monday		1.0	−0.217	0.007	−30.981
Corpus Christi		0.7	−0.198	0.012	−16.315
BRIDGING DAYS					
Ascension Friday		1.0	−0.110	0.009	−12.890
Corpus Christi Friday		0.7	−0.083	0.012	−6.829
DAYLIGHT SAVING TIME					
March	Last Sunday	1.0	−0.032	0.007	−4.567
October	Last Sunday	1.0	0.053	0.007	7.580
ADDITIVE OUTLIERS					
Christmas Eve 2016	24 Dec 2016		0.156	0.020	7.702
New Year's Eve 2016	31 Dec 2016		0.147	0.020	7.402
German Unification Day 2020	3 Oct 2020		0.137	0.019	7.052
Labour Day 2021	1 May 2021		0.122	0.019	6.406
Christmas Day 2021	25 Dec 2021		0.142	0.020	6.968
LEVEL SHIFTS					
	16 Mar 2020		−0.066	0.007	−9.622
	8 Aug 2020		0.051	0.007	7.478
MA PARAMETERS					
θ_1			−0.569	0.014	−40.708
θ_7			0.854	0.013	66.243
$\theta_{365.2425}$			0.874	0.009	99.268

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. Fixed holidays falling onto a Sunday are treated as Sundays. Regression variables for moving holidays and associated bridging days have been centred by removing daily means calculated over the entire sample.

pattern, and a 371-term third-order Henderson kernel and a 3×3 seasonal filter for extracting the DOY pattern. The choice of the short trend filter in the first step ensures that the DOY dynamics are passed on to the DOW-adjusted data as part of the trend-cycle, and the choice of the short seasonal filter in the second step is essentially dictated by sample size. In either step, asymmetric trend filters are obtained through the "cut-and-normalise" method, and the default σ-limits of $(1.5, 2.5)$ are applied. The AMB approach is run in default mode.

Figure 5 (a) shows that the two estimated DOW patterns are very stable and similar in size and shape with the unsurprising exception of slightly less moving seasonality in the canonically smooth AMB estimates. Panel (b) reveals that the two estimated DOY patterns capture nicely the U-shaped infra-yearly dynamics and spike year-end troughs seen in unadjusted electricity consumption (Figure 2 (a)). Again, the AMB estimates are visibly smoother than the X-11 estimates, which also fluctuate much more rapidly as a result of the short DOY seasonal filter. Overall, either approach manages to remove successfully the most distinct repetitive dynamics of the raw series (Figure 6), bearing in mind that such a mere graphical diagnosis is qualitative and might be subject to misjudgement.

4.1.4 Constrained DOW pattern

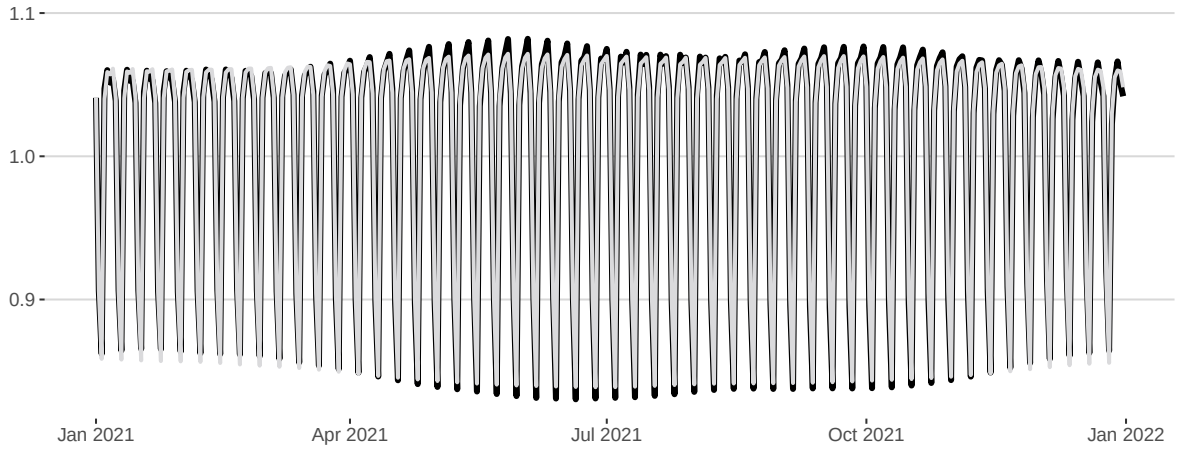
The DOW and DOY patterns extracted with the extended X-11 and AMB approaches cover unconstrained stable and moving seasonality. However, Figure 2 (a) and (c) already provided visual evidence that the DOW pattern between Christmas and New Year (C2NY) looks noticeably different from the rest of the year. To compare the unconstrained and constrained estimates, stable DOW and DOY patterns are extracted with Prophet from linearised electricity consumption, setting $J(7) = 3$ and $J(365.2425) = 10$ in (20) and $\mathcal{C} = \{\text{W52}, \text{W53}\}$ in (5) to impose the C2NY constraint. A piecewise linear trend is also specified alongside automatic change-point selection.

Figure 5 (c) shows the two estimated stable DOW patterns. Compared to the unconstrained estimates, the constrained C2NY estimates are lower from Monday through 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.

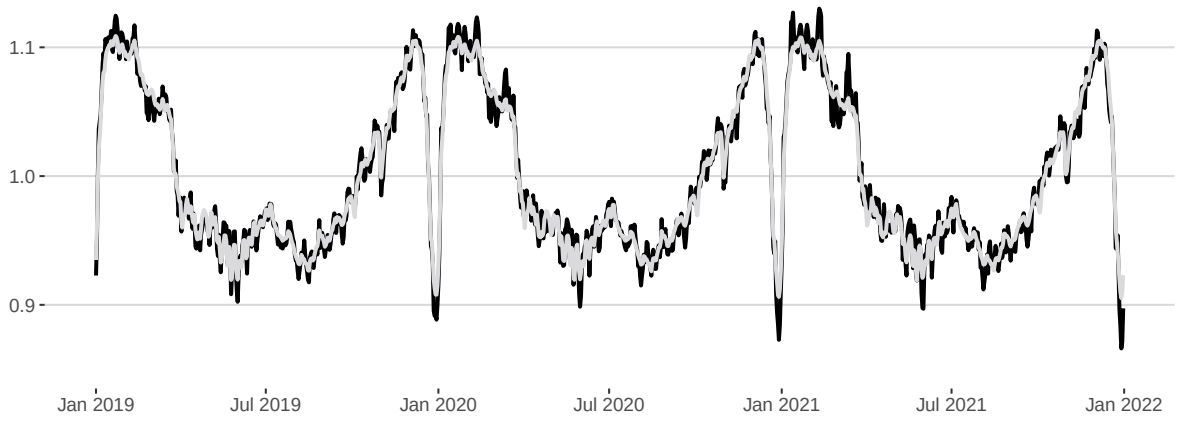
4.2 Hourly counts of TARGET2 customer payments

Hourly counts of customer payments (HCP) in the TARGET2 system in Germany are considered from 2 January to 31 December 2019 (Figure 1 (a)), where counts are measured by the introduction time in the settlement queue during the daytime settlement cycle from 07:00 to 17:00. Hence, the series consists of 10 observations per TARGET2 business day and a total of 2,550 irregularly spaced observations in 2019. Some stylised facts, including the seasonal profile, have already been described in Section 2.1 and visualised in Figure 1 (e)–(f). The latter revealed a dominant hour-of-the-day (HOD) pattern, which is confirmed by spectral graphs (Figure 7) and mainly driven by counts peaking at the early business hours. Being somewhat pronounced around the turn of the week, this "early-bird" effect also seems to introduce an hour-of-the-week (HOW) pattern. We now show how the HCP series can be seasonally adjusted with the STS approach.

(a) DOW pattern: extended X-11 (*black*) and AMB (*grey*) approaches



(b) DOY pattern: extended X-11 (*black*) and AMB (*grey*) approaches



(c) DOW pattern: Prophet without (*solid*) and with constraints to W01–W51 (*dashed*) and W52–W53 (*dotted*)

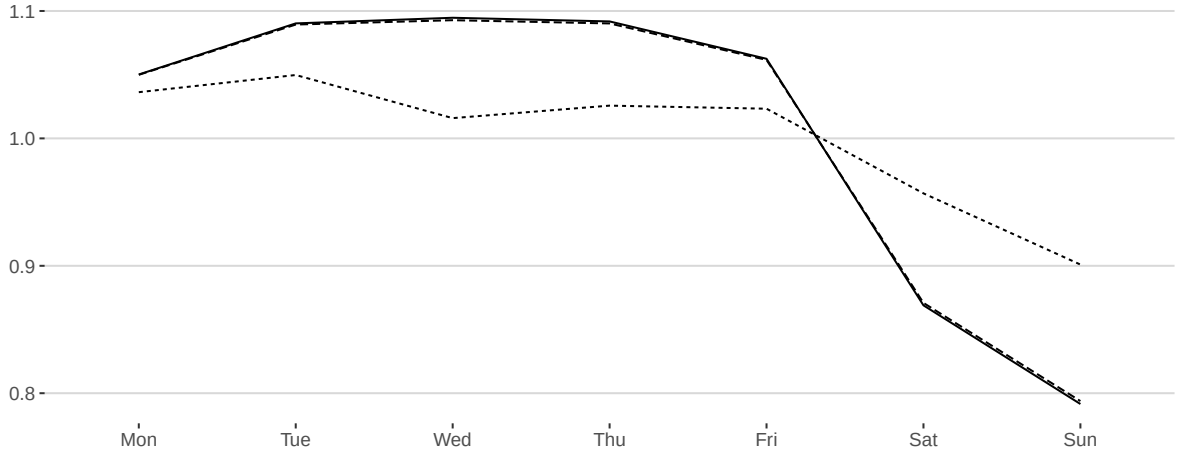
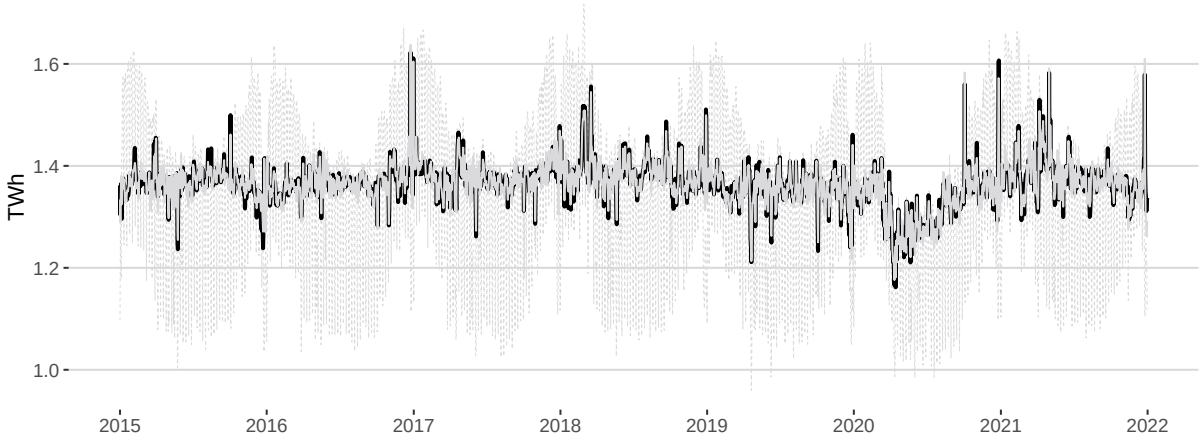


Figure 5: Estimated day-of-the-week (DOW) and day-of-the-year (DOY) patterns for daily electricity consumption.

(a) Time series plot



(b) Estimated periodogram (logged series)

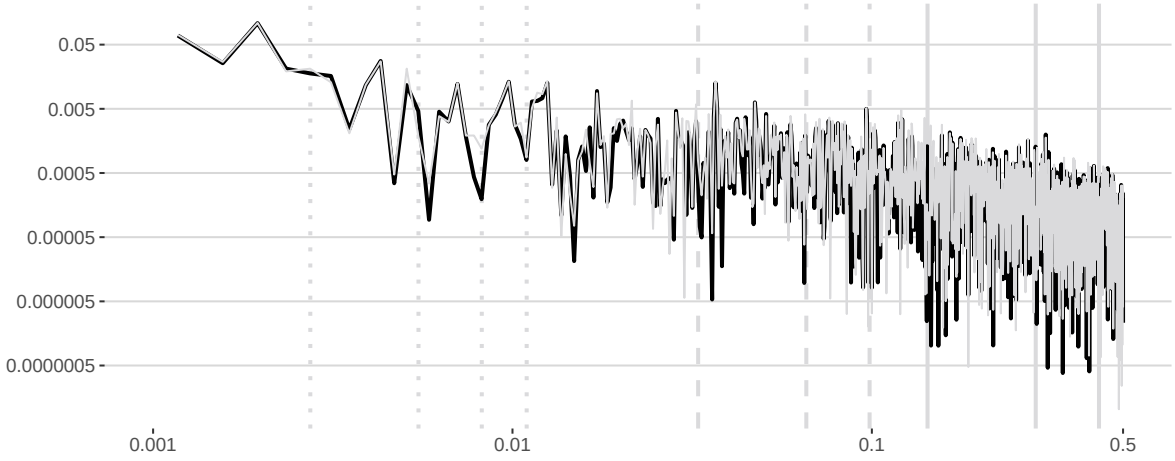


Figure 6: Seasonally adjusted electricity consumption obtained from the extended X-11 (*black*) and AMB (*grey*) approaches. Dashed grey line in Panel (a) corresponds to unadjusted series. Grey verticals in Panel (b) correspond to selected seasonal harmonics of the DOY (*dotted*), DOM (*dashed*) and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

In a first step, a tentative basic structural model (BSM) is fitted to the logged HCP series. The trend-cyclical component in (1) is assumed to follow a local linear trend, which is given by

$$\begin{aligned} t_t &= t_{t-1} + \nu_{t-1} + \xi_t, \quad \xi_t \overset{iid}{\sim} \mathcal{N}\left(0, \sigma_\xi^2\right), \\ \nu_t &= \nu_{t-1} + \zeta_t, \qquad\quad \zeta_t \overset{iid}{\sim} \mathcal{N}\left(0, \sigma_\zeta^2\right), \end{aligned}$$

where the level and slope disturbances, $\{\xi_t\}$ and $\{\zeta_t\}$, are mutually uncorrelated. The seasonal component's HOD pattern in (2) is modelled according to the West-Harrison representation (10) with $\tau = 10$ and trigonometric seasonality (Example 2), the HOW pattern is modelled with the aid of 20 dummies—one for each business hour on Mondays and Fridays—and the irregular component in (1) is assumed to be Gaussian white noise. In addition, a set of 42 dummy variables is considered to account for the following calen-

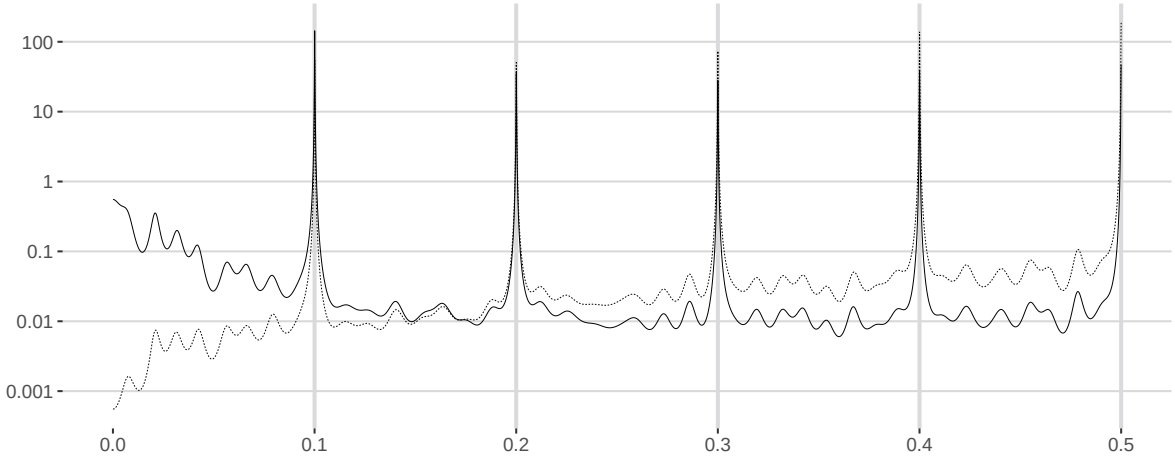


Figure 7: Estimated autoregressive spectrum of logged (*solid*) and differenced logged (*dashed*) HCP series, scaled by 2π. Grey verticals correspond to HOD harmonics.

Table 3: Estimated calendar and outlier effects for the logged HCP series.

Event	Est.	SE	t-value	Event	Est.	SE	t-value
Post-TARGET2 holidays				HOW dummies			
Easter Tuesday	0.284	0.061	4.641	Monday 07:00	0.088	0.018	4.995
Fixed holidays				Monday 09:00	0.139	0.017	8.150
German Unification Day	−0.333	0.060	−5.510	Monday 11:00	0.097	0.017	5.698
All Saints' Day	−0.417	0.060	−6.894	Friday 07:00	0.079	0.018	4.486
Christmas Eve	−0.692	0.060	−11.465	Additive outliers			
Moving holidays				2019-04-23 09:00	0.544	0.112	4.867
Ascension	−0.541	0.060	−8.953	2019-11-18 16:00	0.566	0.110	5.130
Pentecost Monday	−0.338	0.060	−5.599	2019-12-30 10:00	0.545	0.113	4.844
End of quarter							
Q3	0.369	0.060	6.109				
Q4	−0.489	0.086	−5.686				

dar events: (a) banking days that follow immediately after the TARGET2 holiday (New Year's Day, Good Friday, Easter Monday, Labour Day, Christmas Day, Boxing Day); (b) banking days that fall onto the other public holidays considered for electricity consumption; (c) the last banking day of the reserve maintenance periods (RMP) as indicated by the European Central Bank (12 March, 16 April, 11 June, 30 July, 17 September, 29 October and 17 December 2019); (d) the last banking day of each quarter. For the sake of parsimony, the hourly effect of each of the HOW and calendar dummies is assumed constant over the entire banking day. Finally, additive outliers and level shifts are automatically detected, using a critical value of 4.5 which corresponds to the 5% level of significance (see Footnote 4). The corresponding state space model is estimated with the Kalman filter and smoother, using a diffuse square root initialisation.

In a second step, the least significant HOW or calendar dummy is removed and the tentative BSM is re-estimated (including automatic outlier detection) until all calendar and outlier regression variables have absolute t-values larger than 4.5. The resulting final BSM contains four HOW dummies associated with Monday's early business hours

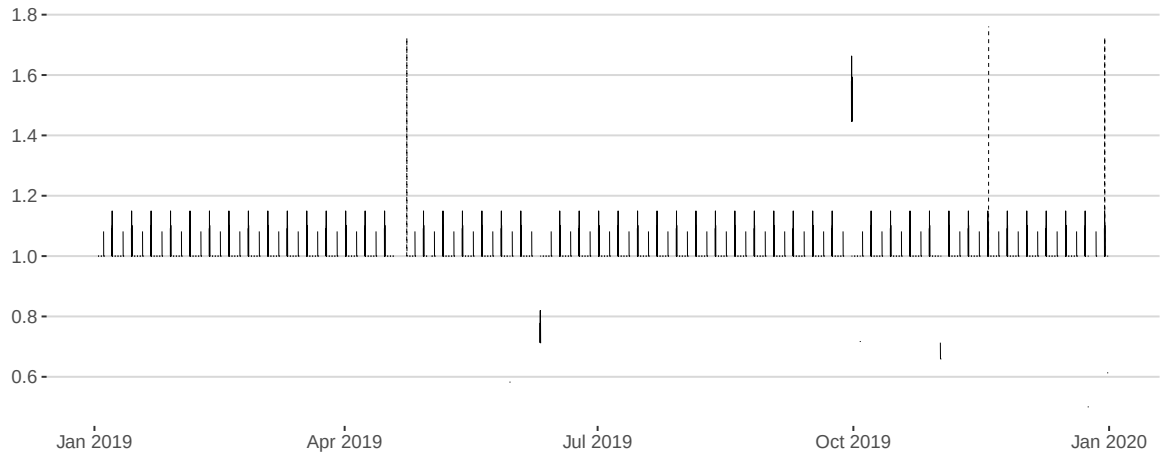


Figure 8: Estimated combined HOW-calendar (*solid*) and outlier (*dashed*) components of the HCP series.

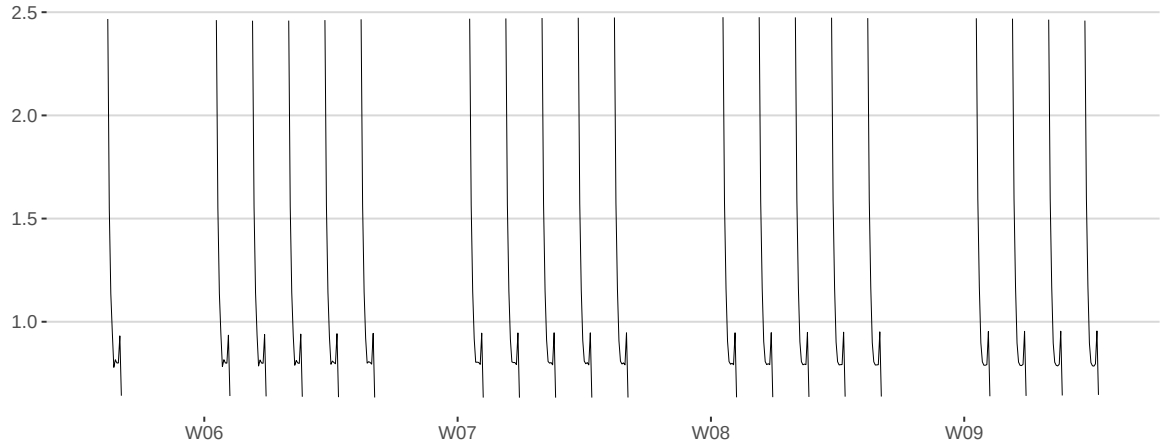
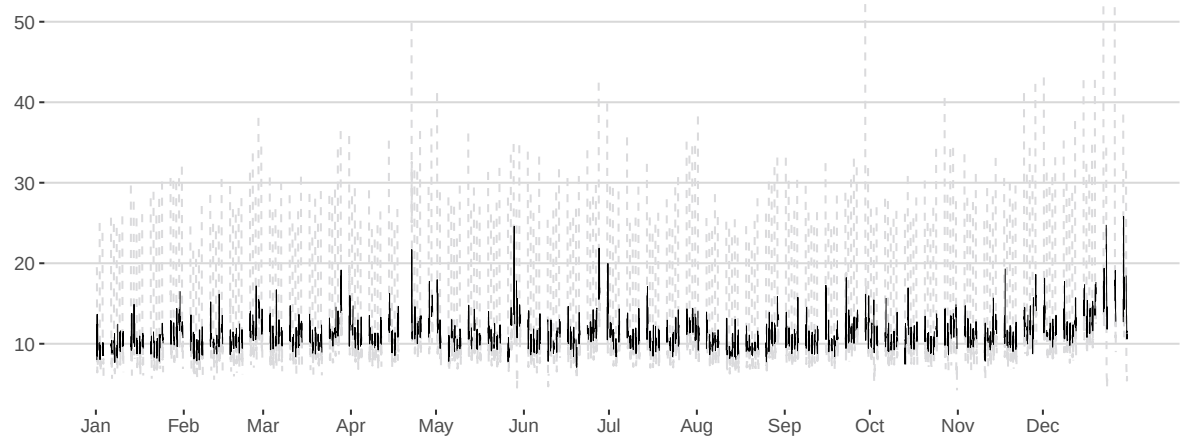


Figure 9: Smoothed estimates of the HOD pattern for the HCP series (1 to 28 February 2019).

and Friday's opening business hour, eight calendar dummies and three additive outliers (Table 3).[5] None of the RMP dummies is included in the final BSM, which may not come as a surprise given the current monetary policy environment and the fact that the data concerns customer payments. More surprisingly, the end-of-quarter dummies for Q1 and Q2 also fail to have sufficiently significant estimated effects, displaying t-values of approximately 3 and 4, respectively, during variable selection. The included calendar effects are in line with earlier descriptive analyses (Figure 1 (b)–(d)), although it should be kept in mind that the sample contains only one observation for each calendar event. The difficulty in estimating those effects from such a few cases is also mirrored by the fact that two out of the three automatically detected additive outliers are close in time to calendar events, such as the one found for 09:00 on the Tuesday after Easter. The estimated combined HOW-calendar and outlier components are shown in Figure 8.

[5]Running variable selection with the fractional Airline model (8) with $\mathcal{S} = \{10\}$ yields essentially the same results, except for an additional yet barely significant level shift for 07:00 on 28 May 2019.

(a) Time series plot



(b) Estimated periodogram (logged series)

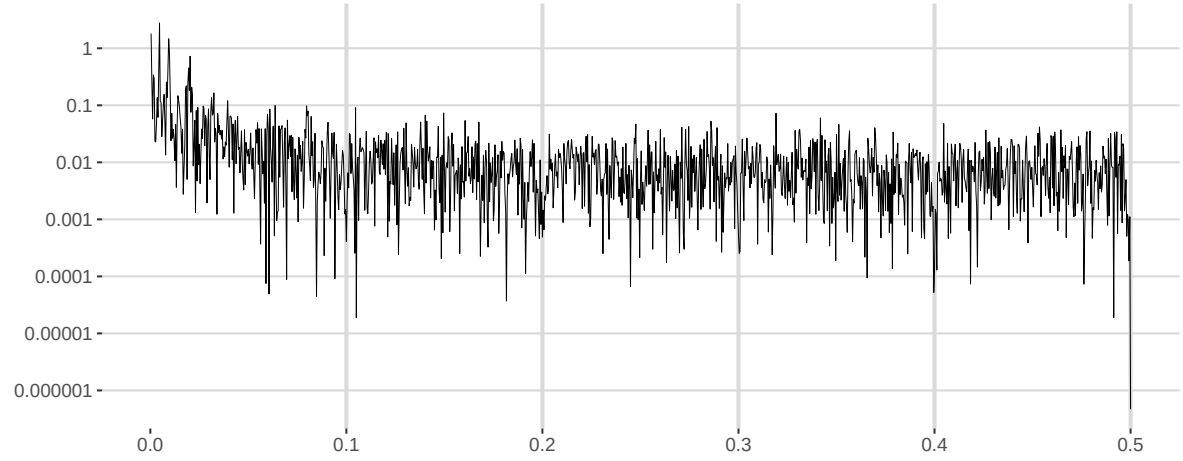


Figure 10: Seasonally adjusted HCP series obtained from the STS approach. Dashed line in Panel (a) corresponds to unadjusted series. Grey verticals in Panel (b) correspond to HOD harmonics.

The q-ratios, that is the estimated disturbance variances relative to that of the irregular component, for the final BSM are given by $\hat{\sigma}^2_\xi/\hat{\sigma}^2_i = 0.133$, $\hat{\sigma}^2_\zeta/\hat{\sigma}^2_i = 0$ and $\hat{\sigma}^2_\omega/\hat{\sigma}^2_i = 1.691 \times 10^{-4}$, so that the local linear trend has turned into a local linear trend with drift. Figure 9 shows the smoothed estimates of the HOD pattern, which capture nicely the "early-bird" effects in the data. Once these and the deterministic combined HOW-calendar effects have been removed, the seasonally adjusted HCP series tends to settle somewhere between 10,000 and 15,000 transactions per hour with a substantial variance reduction of more than 91% compared to the unadjusted data. Overall, the STS approach provides satisfactory results as it manages to remove successfully the most distinct repetitive swings in the HCP series (Figure 10).

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 economic data in official statistics. In seeking to meet those demands, more and more attention has been paid to infra-monthly time series which often display not only seasonal behaviour of known forms but also new stylised facts not observable in monthly and quarterly economic data. Examples include irregular spacing, coexistence of multiple seasonal patterns with integer versus non-integer seasonal periodicities and granular calendar variation. We discussed a broad range of potential issues in the modelling and seasonal adjustment of such peculiar data, highlighting that traditional approaches in official statistics are often inapplicable. Employing a unified latent component model, we reviewed recent developments, made mostly within the past 10 to 15 years, that can handle a fair amount of infra-monthly data peculiarities but seem to have been given little recognition in official statistics so far. We finally illustrated some of these developments, using daily realised electricity consumption and hourly transaction counts in the TARGET2 system in Germany.

Notwithstanding these methodological achievements, it seems that more research is needed to elevate seasonal adjustment of infra-monthly economic data above its current experimental stage in official statistics. Availability of tailored exploratory tools, methods for identifying relevant seasonal patterns, or even cycles, and statistical tests and quality diagnostics for seasonal adjustment adequacy is limited for such data and, therefore, developing such tools could be a first step into this direction. Dealing with increasingly complex models will certainly be another challenge. Dimension reduction techniques, functional components, smooth transition models and continuous-time models or discrete Fourier transforms of periodic time series models (e.g. Anderson, Sabzikar, and Meerschaert, 2021; Chambers, 1999; Chambers and McGarry, 2002; He, Kang, Teräsvirta, and Zhang, 2019; Martín-Rodríguez and Cáceres-Hernández, 2010; Zamani, Haghbin, Hashemi, and Hyndman, 2022) could be interesting approaches to developing flexible yet parsimonious and potentially multivariate models. What is more, some basis functions, such as (cyclical) cubic splines, and continuous-time models in general are directly applicable to irregularly spaced time series without the need for prior data regularisation. Hybrids of non-parametric and model-based approaches (e.g. McElroy and Monsell, 2017) may also pave the way for the addition of correlated latent components (e.g. Hindrayanto, Jacobs, Osborn, and Tian, 2019; McElroy and Maravall, 2014) or GARCH-type heteroskedasticity (e.g. Koopman et al., 2007). Companion advances in the development of fast algorithms and parallel computing could round off any of these approaches with respect to (potentially robust) model estimation. Assuming that the availability of and demand for infra-monthly economic time series will continue to increase, mass production of seasonal adjustments based mainly upon automatic procedures is likely to become another challenge in the future, especially for empirical-based methods, such as the extended STL and X-11 approaches which, for example, currently lack automatic selection rules for trend and seasonal filters given the data.

References

Abeln, B. and J. P. A. M. Jacobs (2015, July). Seasonal adjustment with and without revisions: A comparison of X-13ARIMA-SEATS and CAMPLET. CAMA Working Paper 25/2015.

Abeln, B., J. P. A. M. Jacobs, and P. Ouwehand (2019, April). CAMPLET: Seasonal Adjustment Without Revisions. *Journal of Business Cycle Research 15*(1), 73–95.

Aguilar, P., C. Ghirelli, M. Pacce, and A. Urtasun (2021). Can news help measure economic sentiment? An application in COVID-19 times. *Economics Letters 199*, Article 109730.

Alonso, A. M., C. García-Martos, J. Rodríguez, and M. J. Sánchez (2011, May). Seasonal Dynamic Factor Analysis and Bootstrap Inference: Application to Electricity Market Forecasting. *Technometrics 53*(2), 137–151.

Alvarez, S. E. and S. M. Lein (2020, November). Tracking inflation on a daily basis. *Swiss Journal of Economics and Statistics 156*, Article 18.

Anderson, P. L., F. Sabzikar, and M. M. Meerschaert (2021, July). Parsimonious time series modeling for high frequency climate data. *Journal of Time Series Analysis 42*(4), 442–470.

Ansley, C. F. and R. Kohn (1990, July). Filtering and Smoothing in State Space Models with Partially Diffuse Initial Conditions. *Journal of Time Series Analysis 11*(4), 275–293.

Aprigliano, V., G. Ardizzi, and L. Monteforte (2019, October). Using Payment System Data to Forecast Economic Activity. *International Journal of Central Banking 15*(4), 55–80.

Asai, M., M. McAleer, and S. Peiris (2020, October). Realized stochastic volatility models with generalized Gegenbauer long memory. *Econometrics & Statistics 16*, 42–54.

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

Babbage, C. (1856, March). Analysis of the Statistics of the Clearing House During the Year 1839. *Journal of the Statistical Society of London 19*(1), 28–48.

Bell, W. R. (1984). Seasonal Decomposition of Deterministic Effects. Research Report 84/01, Statistical Research Division, U.S. Bureau of the Census.

Biancotti, C., A. Rosolia, G. Veronese, R. Kirchner, and F. Mouriaux (2021, February). Covid-19 and official statistics: a wakeup call? Occasional Paper No 605, Banca d'Italia, Rome.

Bisaglia, L., S. Bordignon, and F. Lisi (2003). k-Factor GARMA models for intraday volatility forecasting. *Applied Economics Letters 10*(4), 251–254.

Bordignon, S., M. Carporin, and F. Lisi (2007, August). Generalised long-memory GARCH models for intra-daily volatility. *Computational Statistics & Data Analysis 51*(12), 5900–5912.

Burman, J. P. (1980). Seasonal Adjustment by Signal Extraction. *Journal of the Royal Statistical Society: Series A (General) 143*(3), 321–337.

Buss, G. (2016, April). Real-Time Signal Extraction with Regularized Multivariate Direct Filter Approach. *Journal of Forecasting 35*(3), 206–216.

Byrd, R. H., P. Lu, J. Nocedal, and C. Zhu (1995). A Limited Memory Algorithm for Bound Constrained Optimization. *SIAM Journal on Scientific Computing 16*(5), 1190–1208.

Cabrero, A., G. Camba-Mendez, A. Hirsch, and F. Nieto (2009, April). Modelling the Daily Banknotes in Circulation in the Context of the Liquidity Management of the European Central Bank. *Journal of Forecasting 28*(3), 194–217.

Cancelo, J. R. and A. Espasa (1996, September). Modelling and Forecasting Daily Series of Electricity Demand. *Investigaciones Economicas 20*(3), 359–376.

Cancelo, J. R., A. Espasa, and R. Grafe (2008, October–December). Forecasting the electricity load from one day to one week ahead for the Spanish system operator. *International Journal of Forecasting 24*(4), 588–602.

Chambers, M. J. (1999, March). A Note on Modelling Seasonal Processes in Continuous Time. *Journal of Time Series Analysis 20*(2), 139–143.

Chambers, M. J. and J. S. McGarry (2002, April). Modeling Cyclical Behavior with Differential-Difference Equations in an Unobserved Components Framework. *Econometric Theory 18*(2), 387–419.

Cleveland, R. B., W. S. Cleveland, J. E. McRae, and I. Terpenning (1990, March). STL: A Seasonal-Trend Decomposition Procedures Based on Loess (with comments and rejoinder). *Journal of Official Statistics 6*(1), 3–73.

Cleveland, W. P., T. Evans, and S. Scott (2018). Weekly Seasonal Adjustment: A Locally-Weighted Regression Approach. In G. L. Mazzi, D. Ladiray, and D. A. Riester (Eds.), *Handbook on Seasonal Adjustment*, Chapter 28, pp. 735–755. Luxembourg: Publications Office of the European Union.

Cleveland, W. P. and M. R. Grupe (1983). Modeling Time Series When Calendar Effects Are Present (with comments). In A. Zellner (Ed.), *Applied Time Series Analysis of Economic Data*, pp. 57–73. Washington, D.C.: U.S. Department of Commerce, Bureau of the Census.

Cleveland, W. P. and S. Scott (2007, June). Seasonal Adjustment of Weekly Time Series with Application to Unemployment Insurance Claims and Steel Production. *Journal of Official Statistics 23*(2), 209–221.

Coffinet, J., J.-B. Delbos, V. Kaiser, J.-N. Kien, E. Kintzler, A. Lestrade, M. Mouliom, T. Nicolas, J.-C. Bricongne, and B. Meunier (2020, September–October). Tracking the economy during the Covid-19 pandemic: the contribution of high-frequency indicators. *Bulletin de la Banque de France 231*, Article 5.

Cordeiro, C. and M. M. Neves (2009, June). Forecasting Time Series with BOOT.EXPOS Procedure. *REVSTAT Statistical Journal 7*(2), 135–149.

Cottet, R. and M. Smith (2003, December). Bayesian Modeling and Forecasting of Intraday Electricity Load. *Journal of the American Statistical Association 98*(464), 839–849.

Cox, M., J. Triebel, S. Linz, C. Fries, L. F. Flores, A. Lorenz, D. Ollech, A. Dietrich, J. LeCrone, and K. Webel (2020, August). Daily truck toll mileage index based on digital process data from toll collection system. *Wirtschaft und Statistik 72*(4), 63–76. In German, English version available under URL https://www.destatis.de/EN/Methods/WISTAScientificJournal/Downloads/truck-toll-mileage-index-042020.html.

Czaplicki, N. and R. J. Hutchinson (2020, January). Using Daily Payment Processor Data to Determine Existence and Length of Retail Shopping Event Effects. Working Paper ADEP-WP-012020-03, U.S. Census Bureau.

de Jong, P. (1991, June). The Diffuse Kalman Filter. *The Annals of Statistics 19*(2), 1073–1083.

de Jong, P. and S. Chu-Chun-Lin (2003, March). Smoothing With An Unknow Initial Condition. *Journal of Time Series Analysis 24*(2), 141–148.

de Jong, P. and J. Penzer (1998, June). Diagnosing Shocks in Time Series. *Journal of the American Statistical Association 93*(442), 796–806.

De Livera, A. M., R. J. Hyndman, and R. D. Snyder (2011, December). Forecasting Time Series With Complex Seasonal Patterns Using Exponential Smoothing. *Journal of the American Statistical Association 106*(496), 1513–1527.

Dickopf, X., C. Janz, and T. Mucha (2019, December). Moving from GDP flash estimates to GDP nowcasts: first results of a feasibility study to further accelerate early GDP estimation. *Wirtschaft und Statistik 2019*(6), 47–58. In German, English version available under URL https://www.destatis.de/EN/Methods/WISTAScientificJournal/Downloads/moving-gdp-flash-estimates-to-gdp-nowcasts-062019.html.

Dokumentov, A. and R. J. Hyndman (2015, June). STR: A Seasonal-Trend Decomposition Procedure Based on Regression. Working Paper 13/15, Monash University, Department of Econometrics and Business Statistics.

Doornik, J. A., J. L. Castle, and D. F. Hendry (2022, April–June). Short-term forecasting of the coronavirus pandemic. *International Journal of Forecasting 38*(2), 453–466.

Dordonnat, V., S. J. Koopman, M. Ooms, A. Dessertaine, and J. Collet (2008, October–December). An hourly periodic state space model for modelling French national electricity load. *International Journal of Forecasting 24*(4), 566–587.

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

Eckert, F. and H. Mikosch (2020, August). Mobility and sales activity during the Corona crisis: daily indicators for Switzerland. *Swiss Journal of Economics and Statistics 156*, Article 9.

Eraslan, S. and T. Götz (2021, July). An unconventional weekly activity index for Germany. *Economics Letters 204*, Article 109881.

Farley, D. E. and Y.-Y. C. O'Brien (1987, September). Seasonal Adjustment of the Monetary Stock in the United States. *Journal of Official Statistics 3*(3), 223–233.

Fenz, G. and H. Stix (2021, March). 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.

Findley, D. F. (2005, June). Some Recent Developments and Directions in Seasonal Adjustment. *Journal of Official Statistics 21*(2), 343–365.

Gasser, T. and H.-G. Müller (1979). Kernel Estimation of Regression Functions. In T. Gasser and M. Rosenblatt (Eds.), *Smoothing Techniques for Curve Estimation*, Volume 757 of *Lecture Notes in Mathematics*, pp. 23–68. Heidelberg: Springer.

Giraitis, L. and R. Leipus (1995, January). A Generalized Fractionally Differencing Approach in Long-Memory Modeling. *Lithuanian Mathematical Journal 35*(1), 53–65.

Gómez, V. and A. Maravall (1994, June). Estimation, Prediction, and Interpolation for Nonstationary Series With the Kalman Filter. *Journal of the American Statistical Association 89*(426), 611–624.

Gould, P. G., A. B. Koehler, J. K. Ord, R. D. Snyder, R. J. Hyndman, and F. Vahid-Araghi (2008, November). Forecasting Time Series With Multiple Seasonal Patterns. *European Journal of Operational Research 191*(1), 207–222.

Grassi, S., G. L. Mazzi, and T. Proietti (2018). Automatic Outlier Detection for the Basic Structural Time Series Model. In G. L. Mazzi, D. Ladiray, and D. A. Riester (Eds.), *Handbook on Seasonal Adjustment*, Chapter 8, pp. 169–194. Luxembourg: Publications Office of the European Union.

Gray, H. L., N.-F. Zhang, and W. A. Woodward (1989, May). On Generalized Fractional Processes. *Journal of Time Series Analysis 10*(3), 233–257.

Grun-Rehomme, M., F. Guggemos, and D. Ladiray (2018). Asymmetric Moving Averages Minimizing Phase Shift. In G. L. Mazzi, D. Ladiray, and D. A. Riester (Eds.), *Handbook on Seasonal Adjustment*, Chapter 15, pp. 391–413. Luxembourg: Publications Office of the European Union.

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

Harvey, A. C. and S. J. Koopman (1993, December). Forecasting Hourly Electricity Demand Using Time-Varying Splines. *Journal of the American Statistical Association 88*(424), 1228–1236.

Harvey, A. C., S. J. Koopman, and M. Riani (1997, July). The Modeling and Seasonal Adjustment of Weekly Observations. *Journal of Business & Economic Statistics 15*(3), 354–368.

He, C., J. Kang, T. Teräsvirta, and S. Zhang (2019, October). The shifting seasonal mean autoregressive model and seasonality in the Central England monthly temperature series, 1772–2016. *Econometrics & Statistics 12*, 1–24.

Hillmer, S. C. and G. C. Tiao (1982, March). An ARIMA Model-Based Approach to Seasonal Adjustment. *Journal of the American Statistical Association 77*(377), 63–70.

Hindrayanto, I., J. P. A. M. Jacobs, D. R. Osborn, and J. Tian (2019, December). Trend-Cycle-Seasonal Interactions: Identification and Estimation. *Macroeconomic Dynamics 23*(8), 3163–3188.

Holan, S. H. and T. S. McElroy (2012). Bayesian Seasonal Adjustment of Long Memory Time Series. In W. R. Bell, S. H. Holan, and T. S. McElroy (Eds.), *Economic Time Series – Modeling and Seasonality*, pp. 403–429. Boca Raton: CRC 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., H. Shen, and A. Buja (2008). Functional Principle Components Analysis via Penalized Rank One Approximation. *Electronic Journal of Statistics 2*, 678–695.

Huang, J. Z., H. Shen, and A. Buja (2009, December). The Analysis of Two-Way Functional Data Using Two-Way Regularized Singular Value Decompositions. *Journal of the American Statistical Association 104*(488), 1609–1620.

Hyndman, R. J. and S. Fan (2010, May). Density Forecasting for Long-Term Peak Electricity Demand. *IEEE Transactions on Power Systems 25*(2), 1142–1153.

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

Hyndman, R. J., A. B. Koehler, R. D. Snyder, and S. Grose (2002, July–September). A state space framework for automatic forecasting using exponential smoothing methods. *International Journal of Forecasting 18*(3), 439–454.

Jarmin, R. S. (2019, Winter). Evolving Measurement for an Evolving Economy: Thoughts on the 21st Century US Economic Statistics. *Journal of Economic Perspectives 33*(1), 165–184.

Kang, Y., R. J. Hyndman, and K. Smith-Miles (2017, April–June). Visualising forecasting algorithm performance using time series instance spaces. *International Journal of Forecasting 33*(2), 345–358.

Keane, M. and T. Neal (2021, January). Consumer panic in the COVID-19 pandemic. *Journal of Econometrics 220*(1), 86–105.

Koopman, S. J. (1993, March). Disturbance Smoother for State Space Models. *Biometrika 80*(1), 117–126.

Koopman, S. J. (1997, December). Exact Initial Kalman Filtering and Smoothing for Non-stationary Time Series Models. *Journal of the American Statistical Association 92*(440), 1630–1638.

Koopman, S. J. and J. Durbin (2003, January). Filtering and Smoothing of State Vector for Diffuse State-Space Models. *Journal of Time Series Analysis 24*(1), 85–98.

Koopman, S. J. and M. Ooms (2003, November). Time Series Modelling of Daily Tax Revenues. *Statistica Neerlandica 57*(4), 439–469.

Koopman, S. J. and M. Ooms (2006, November). Forecasting daily time series using periodic unobserved components time series models. *Computational Statistics & Data Analysis 51*(2), 885–903.

Koopman, S. J., M. Ooms, and M. A. Carnero (2007, March). Periodic Seasonal Reg-ARFIMA-GARCH Models for Daily Electricity Spot Prices. *Journal of the American Statistical Association 102*(477), 16–27.

Ladiray, D., J. Palate, G. L. Mazzi, and T. Proietti (2018). Seasonal Adjustment of Daily and Weekly Data. In G. L. Mazzi, D. Ladiray, and D. A. Riester (Eds.), *Handbook on Seasonal Adjustment*, Chapter 29, pp. 757–783. Luxembourg: Publications Office of the European Union.

Lee, S., Y. Liao, M. H. Seo, and Y. Shin (2021, January). Sparse HP filter: Finding kinks in the COVID-19 contact rate. *Journal of Econometrics 220*(1), 158–180.

Leschinski, C. and P. Sibbertsen (2019, January). Model order selection in periodic long memory models. *Econometrics & Statistics 9*, 78–94.

Lewis, D. J., K. Mertens, J. H. Stock, and M. Trivedi (2022, June–July). Measuring Real Activity Using a Weekly Economic Index. *Journal of Applied Econometrics 37*(4), 667–687.

Li, S. and O. Linton (2021, January). When will the Covid-19 pandemic peak? *Journal of Econometrics 220*(1), 130–157.

Lin, W., J. Z. Huang, and T. S. McElroy (2020, July). Time Series Seasonal Adjustment Using Regularized Singular Value Decomposition. *Journal of Business & Economic Statistics 38*(3), 487–501.

Liou, S.-M., J.-L. Lin, and C.-N. Peng (2012, October). Using High Frequency Data to Model Moving Holiday Effects: An Empirical Investigation of Taiwanese Monetary Aggregates. *Taiwan Economic Forecast and Policy 43*(1), 171–192.

Liu, J. M., R. Chen, L.-M. Liu, and J. L. Harris (2006, December). A Semi-Parametric Time Series Approach in Modeling Hourly Electricity Loads. *Journal of Forecasting 25*(8), 537–559.

Ljung, G. M. (1993). On Outlier Detection in Time Series. *Journal of the Royal Statistical Society: Series B (Methodological) 55*(2), 559–567.

Lourenço, N. and A. Rua (2021, July). The Daily Economic Indicator: tracking economic activity daily during the lockdown. *Economic Modelling 100*, Article 105500.

Martín-Rodríguez, G. and J. J. Cáceres-Hernández (2005, May). Modelling the hourly Spanish electricity demand. *Economic Modelling 22*(3), 551–569.

Martín-Rodríguez, G. and J. J. Cáceres-Hernández (2010, January). Splines and the proportion of the seasonal period as a season index. *Economic Modelling 27*(1), 83–88.

McElroy, T. S. (2008, August). Matrix Formulas for Nonstationary ARIMA Signal Extraction. *Econometric Theory 24*(4), 988–1009.

McElroy, T. S. (2017, October). Multivariate Seasonal Adjustment, Economic Identities, and Seasonal Taxonomy. *Journal of Business & Economic Statistics 35*(4), 611–625.

McElroy, T. S. and S. H. Holan (2012, October). On the Computation of Autocovariances for Generalized Gegenbauer Processes. *Statistica Sinica 22*(4), 1661–1687.

McElroy, T. S. and S. H. Holan (2016, September). Computation of the autocovariances for time series with multiple long-range persistencies. *Computational Statistics & Data Analysis 101*, 44–56.

McElroy, T. S. and A. Maravall (2014, July). Optimal Signal Extraction with Correlated Components. *Journal of Time Series Econometrics 6*(2), 237–273.

McElroy, T. S. and B. C. Monsell (2017, October). Issues Related to the Modeling and Adjustment of High Frequency Time Series. Research Report No 2017-08, Center for Statistical Research & Methodology, U. S. Census Bureau, Washington, D. C.

McElroy, T. S., B. C. Monsell, and R. J. Hutchinson (2018, January). Modeling of Holiday Effects and Seasonality in Daily Time Series. Research Report No 2018-01, Center for Statistical Research & Methodology, U. S. Census Bureau, Washington, D. C.

McElroy, T. S. and T. Trimbur (2015, March). Signal Extraction for Non-Stationary Multivariate Time Series with Illustrations for Trend Inflation. *Journal of Time Series Analysis 36*(2), 209–227.

McElroy, T. S. and M. Wildi (2020, April). The Multivariate Linear Prediction Problem: Model-Based and Direct Filtering Solutions. *Econometrics & Statistics 14*, 112–130.

Mestekemper, T., M. Windmann, and G. Kauermann (2010, October–December). Functional hourly forecasting of water temperature. *International Journal of Forecasting 26*(4), 684–699.

Monteiro, A., R. Menezes, and M. E. Silva (2017, November). Modelling spatio-temporal data with multiple seasonalities: The NO_2 Portuguese case. *Spatial Statistics 22*(2), 371–387.

Morf, M. and T. Kailath (1975, August). Square-Root Algorithms for Least-Squares Estimation. *IEEE Transactions on Automatic Control 20*(4), 487–497.

Morf, M., G. S. Sidhu, and T. Kailath (1974, August). Some New Algorithms for Recursive Estimation in Constant, Linear, Discrete-Time Systems. *IEEE Transactions on Automatic Control 19*(4), 315–323.

Nyamela, Y., V. Plakandaras, and R. Gupta (2020). Frequency-dependent real-time effects of uncertainty in the United States: evidence from daily data. *Applied Economics Letters 27*(19), 1562–1566.

Ollech, D. (2021, July). Seasonal Adjustment of Daily Time Series. *Journal of Time Series Econometrics 13*(2), 235–264.

Ord, J. K., A. B. Koehler, and R. D. Snyder (1997, December). Estimation and Prediction for a Class of Dynamic Nonlinear Statistical Models. *Journal of the American Statistical Association 92*(440), 1621–1629.

Pedregal, D. J. and P. C. Young (2006, January–March). Modulated cycles, an approach to modelling periodic components from rapidly sampled data. *International Journal of Forecasting 22*(1), 181–194.

Pierce, D. A., M. R. Grupe, and W. P. Cleveland (1984, July). Seasonal Adjustment of the Weekly Monetary Aggregates: A Model-Based Approach. *Journal of Business & Economic Statistics 2*(3), 260–270.

Proietti, T. (2000, April). Comparing Seasonal Components for Structural Time Series Models. *International Journal of Forecasting 16*(2), 247–260.

Proietti, T. and A. Luati (2008, December). Real Time Estimation in Local Polynomial Regression, with Application to Trend-Cycle Analysis. *Annals of Applied Statistics 2*(4), 1523–1553.

Proietti, T. and D. J. Pedregal (2022). Seasonality in High Frequency Time Series. *Econometrics & Statistics*. https://doi.org/10.1016/j.ecosta.2022.02.001.

Puindi, A. C. and M. E. Silva (2021). Dynamic structural models with covariates for short-term forecasting of time series with complex seasonal patterns. *Journal of Applied Statistics 48*(5), 804–826.

Radermacher, W. J. (2019). Governing-by-the-numbers/Statistical governance: Reflections on the future of official statistics in a digital and globalised society. *Statistical Journal of the IAOS 35*(4), 519–537.

Radermacher, W. J. (2020). *Official Statistics 4.0 – Verified Facts for People in the 21st Century*. Heidelberg: Springer.

Ramanathan, R., R. Engle, C. W. J. Granger, F. Vahid-Araghi, and C. Brace (1997, June). Short-run forecasts of electricity loads and peaks. *International Journal of Forecasting 13*(2), 161–174.

Rodriguez, A. and E. Ruiz (2009, March). Bootstrap Prediction Intervals in State-Space Models. *Journal of Time Series Analysis 30*(2), 167–178.

Seiler, P. (2020, September). Weighting bias and inflation in the time of COVID-19: evidence from Swiss transaction data. *Swiss Journal of Economics and Statistics 156*, Article 13.

Shiskin, J., A. H. Young, and J. C. Musgrave (1967). The X-11 Variant of the Census Method II Seasonal Adjustment Program. Technical Paper No 15, U.S. Department of Commerce, Bureau of the Census, Washington, D.C.

Soares, L. J. and M. C. Medeiros (2008, October–December). Modeling and forecasting short-term electricity load: a comparison of methods with an application to Brazilian data. *International Journal of Forecasting 24*(4), 630–644.

Soares, L. J. and L. R. Souza (2006, January–March). Forecasting electricity demand using generalized long memory. *International Journal of Forecasting 22*(1), 17–28.

Taylor, J. W. (2003). Short-Term Electricity Demand Forecasting Using Double Seasonal Exponential Smoothing. *Journal of the Operational Research Society 54*(8), 799–805.

Taylor, J. W. (2008, October–December). An evaluation of methods for very short-term load forecasting using minute-by-minute British data. *International Journal of Forecasting 24*(4), 645–658.

Taylor, J. W. (2010a, October–December). Exponentially weighted methods for forecasting intraday time series with multiple seasonal cycles (with comments and reply). *International Journal of Forecasting 26*(4), 627–660.

Taylor, J. W. (2010b, July). Triple seasonal methods for short-term electricity demand forecasting. *European Journal of Operational Research 204*(1), 139–152.

Taylor, J. W., L. M. de Menezes, and P. E. McSharry (2006, January–March). A comparison of univariate methods for forecasting electricity demand up to a day ahead. *International Journal of Forecasting 22*(1), 1–16.

Taylor, J. W. and P. E. McSharry (2007, November). Short-Term Load Forecasting Methods: An Evaluation Based on European Data. *IEEE Transactions on Power Systems 22*(4), 2213–2219.

Taylor, J. W. and R. D. Snyder (2012, December). Forecasting intraday time series with multiple seasonal cycles using parsimonious seasonal exponential smoothing. *Omega 40*(6), 748–757.

Taylor, S. J. and B. Letham (2018). Forecasting at Scale. *The American Statistician 72*(1), 37–45.

Tissot, B. and B. De Beer (2020, November). Implications of Covid-19 for official statistics: a central banking perspective. IFC Working Paper No 20, Bank for International Settlements, Basel.

Voges, M. and P. Sibbertsen (2021, July). Cyclical fractional cointegration. *Econometrics & Statistics 19*, 114–129.

Wang, X., K. Smith, and R. J. Hyndman (2006, November). Characteristic-Based Clustering for Time Series Data. *Data Mining and Knowledge Discovery 13*(3), 335–364.

Webel, K. (2020). Challenges and Recent Developments in the Seasonal Adjustment of Daily Time Series. JSM Proceedings, Business and Economic Statistics Section, Alexandria, VA: American Statistical Association, pp. 1971–1997.

Wegmüller, P., C. Glocker, and V. Guggia (2021, April). Weekly economic activity: Measurement and informational content. Grundlagen für die Wirtschaftspolitik Nr. 17, Staatssekretariat für Wirtschaft SECO, Bern, Schweiz.

Weinberg, J., L. D. Brown, and J. R. Stroud (2007, December). Bayesian Forecasting of an Inhomogeneous Poisson Process with Applications to Call Center Data. *Journal of the American Statistical Association 102*(480), 1185–1198.

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

Wildi, M. (2005). *Signal Extraction: Efficient Estimation, 'Unit-Root'-Tests and Early Detection of Turning Points*, Volume 547 of *Lecture Notes in Economics and Mathematical Systems*. Heidelberg: Springer.

Wildi, M. (2008). *Real-Time Signal Extraction: Beyond Maximum Likelihood Principles*. Berlin: Springer.

Wildi, M. and T. S. McElroy (2016, June). Optimal Real-Time Filters for Linear Prediction Problems. *Journal of Time Series Econometrics 8*(2), 155–192.

Wildi, M. and T. S. McElroy (2019, July–September). The trilemma between accuracy, timeliness and smoothness in real-time signal extraction. *International Journal of Forecasting 35*(3), 1072–1084.

Winters, P. R. (1960, April). Forecasting Sales by Exponentially Weighted Moving Averages. *Management Science 6*(3), 324–342.

Woloszko, N. (2020). Tracking activity in real time with Google Trends. OECD Economics Department Working Paper No. 1634.

Woodward, W. A., Q. C. Cheng, and H. L. Gray (1998, July). A k-Factor GARMA Long-Memory Model. *Journal of Time Series Analysis 19*(4), 485–504.

Wu, L. S.-Y., J. R. M. Hosking, and N. Ravishanker (1993). Reallocation Outliers in Time Series. *Journal of the Royal Statistical Society: Series C (Applied Statistics) 42*(2), 301–313.

Zamani, A., H. Haghbin, M. Hashemi, and R. J. Hyndman (2022, March). Seasonal functional autoregressive models. *Journal of Time Series Analysis 43*(2), 197–218.