DEUTSCHE BUNDESBANK
EUROSYSTEM

Discussion Paper

Deutsche Bundesbank
No 01/2026

Diagnostic tools for selecting the temporal resolution for seasonal adjustment

Daniel Ollech
Martin Stefan

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 email: www.bundesbank.de/contact

Internet http://www.bundesbank.de

DOI https://doi.org/10.71734/DP-2026-1
ISBN 978-3-98848-058-3
ISSN 2941-7503

Bundesbank Discussion Paper No 01/2026

Diagnostic tools for selecting the temporal resolution for seasonal adjustment*

Daniel Ollech
Deutsche Bundesbank

Martin Stefan
Deutsche Bundesbank

Abstract

Official statistics increasingly make use of higher-frequency time series. But when users ultimately are interested in a seasonally adjusted temporal aggregate of these data, we have to decide whether to perform seasonal adjustment or aggregation first. Consequently, we must weigh up the benefits of richer informational content against the increased computational requirements and the challenges presented by using more volatile and outlier-prone data. We examine this trade-off on simulated and real-world time series using a battery of diagnostics including revision size, tests on residual seasonal and calendar effects and linkage with target variables using leading adjustment procedures: DSA, WSA, X-13-ARIMA, and TRAMO-SEATS. We synthesise our findings into practical guidelines that help users choose the aggregation level that balances statistical quality and real-time usefulness.

Keywords: higher frequency time series; temporal aggregation; official time series

JEL classification: C13, C14, C22, C52

*Contact address: Daniel Ollech, Deutsche Bundesbank, Central Office, Directorate General Statistics, Wilhelm-Epstein-Strasse 14, 60431 Frankfurt am Main, Germany. Tel.: +49 69 9566 36250. Email: daniel.ollech@bundesbank.de. The authors thank Jakob Oberhammer for help with preparing initial computations as well as Anna Smyk and participants at the Bundesbank researcher's seminar for their helpful comments.

1 Introduction

Unlike forecasting, seasonal adjustment of a time series is a multi-purpose exercise. The purposes include accurately estimating the seasonal and calendar components, and obtaining a seasonally adjusted series that is informative for business cycle analysis and serves as a suitable input for subsequent economic forecasting. Because some of these objectives are latent and cannot be observed and at the weekly level, the reverse holds; their evaluation inevitably involves a degree of subjectivity. Analysts generally need to employ a battery of diagnostics to select the most appropriate set-up for seasonal adjustment.

One important choice is the temporal resolution for seasonal adjustment, i.e. the temporal aggregation level at which seasonal adjustment is conducted. This is especially relevant, when a higher-frequency time series is available. If for example a daily series is observed but the economic analysis focuses on the quarterly aggregate, we can choose from at least 4 temporal resolutions: the series can be adjusted at the daily, weekly, monthly or quarterly aggregation level.

Even when only lower frequency time series are available, the question of the optimal temporal resolution may arise: in a report for the U.S. Bureau of Economic Analysis, Moulton and Cowan (2016) point out that several monthly series that are part of the compilation of GDP and GDI data, showed no detectable seasonality at the monthly frequency and were therefore not adjusted, yet displayed seasonal patterns once they were aggregated to the quarterly level. Other series that had already been seasonally adjusted still exhibited seasonality after temporal aggregation. Both situations contribute to the residual seasonality observed in the published aggregates.

More generally, Ollech (2022) demonstrates that the properties we can detect in a seasonal time series depend on the level of temporal aggregation. For instance, the interaction between the annual seasonal pattern and the day-of-the-week effect is only discernible when the data are observed at sub-weekly periodicities. Indeed, ESS guidelines on seasonal adjustment take this into account and allow to include different effects at different temporal resolutions (Eurostat, 2024).

Here, we study how the choice of the temporal resolution impacts the seasonal adjustment results and which diagnostics to rely on to help guide this choice and discuss their limitations. Section 2 presents the seasonal adjustment methods considered and discusses the limitations of the methods at the different temporal resolutions. Section 3 gives an overview of appropriate diagnostical tools, while Section 4 illustrates the differences of seasonal adjustment results at different temporal resolutions using simulated data and applies the diagnostics to real-world series. Finally, Section 5 concludes.

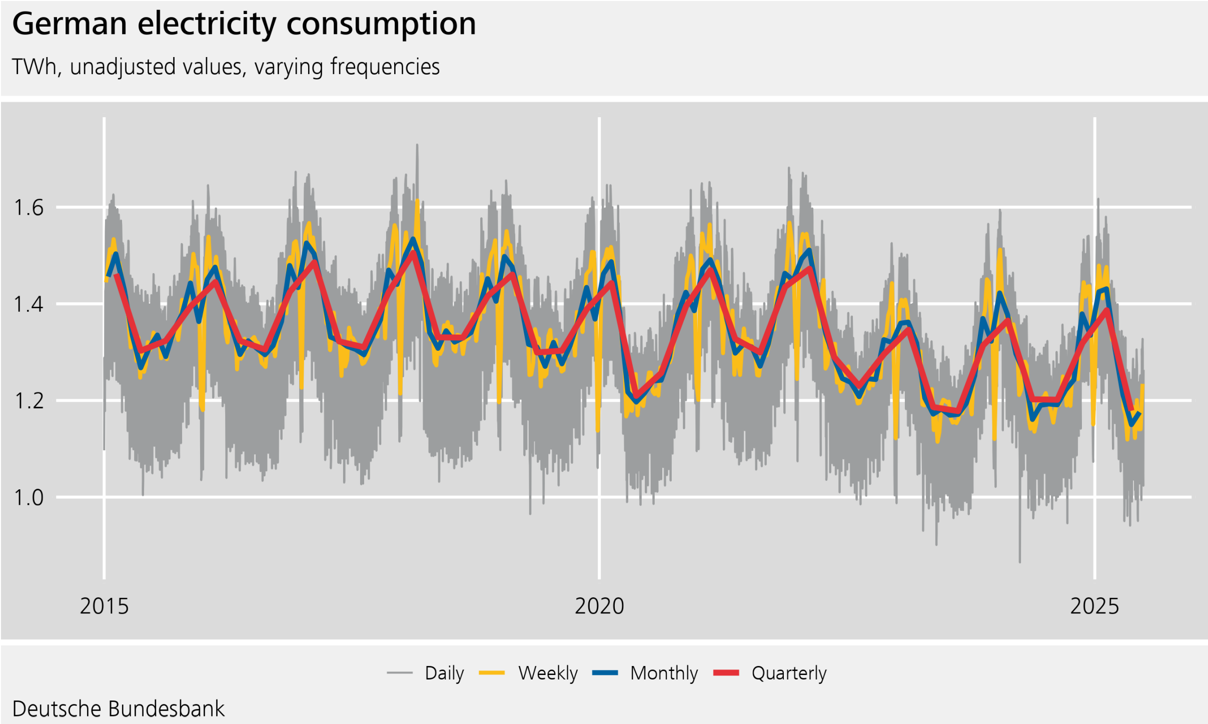

Figure 1: Electricity consumption (unadjusted)

2 Methodological considerations

Let $Y_t \; \forall \; t \in 1, \ldots, T$ be a time series observed at equally-spaced intervals and let B denote the backshift operator, so that $B^k Y_t = y_{t-k}$. A seasonal ARIMA$(p,d,q)(P,D,Q)_\tau$ model with seasonal period τ is written

$$\phi_p(B)\Phi_P(B^\tau)(1-B)^d(1-B^\tau)^D Y_t = \theta_q(B)\Theta_Q(B^\tau)\varepsilon_t, \qquad \varepsilon_t \overset{\text{i.i.d.}}{\sim} (0,\sigma^2), \tag{1}$$

where $\phi_p(B) = 1 - \phi_1 B - \ldots - \phi_p B^p$ and $\Phi_P(B^\tau) = 1 - \Phi_1 B^\tau - \ldots - \Phi_P B^{P\tau}$ are the non-seasonal and seasonal autoregressive (AR) polynomials, while $\theta_q(B) = 1 + \theta_1 B + \ldots + \theta_q B^q$ and $\Theta_Q(B^\tau) = 1 + \Theta_1 B^\tau + \ldots + \Theta_Q B^{Q\tau}$ are the corresponding moving-average (MA) polynomials. The operators $(1-B)^d$ and $(1-B^\tau)^D$ apply d orders of non-seasonal differencing and D orders of seasonal differencing, respectively.

How does the temporal aggregation level affect the ARIMA model? Rossana and Seater (1995) summarise the results derived by Brewer (1973) and Tiao (1972) as two-fold: 1) The AR order of the time series remains the same after temporal aggregation, except that the coefficients will decrease to zero, as k increases; all MA coefficients, except the first d, will decrease to zero as k increases. 2) Even random walks will lead to ARIMA(0,d,d) models through temporal aggregation, though the decrease in the number of observations increases standard errors, which may render the MA terms statistically insignificant and thus potentially undetectable

Based on the analysis of seasonal ARIMA models, Wei (1978) shows that estimation and forecasting become less efficient if temporal aggregates of the series are used, provided the non-seasonal part of the time series is non-stationary. In general, temporal aggregation reduces the complexity of the underlying ARIMA model; for instance, converting a monthly or quarterly series to an annual (or lower) frequency removes the seasonal component altogether (Silvestrini and Veredas, 2008).

The RegARIMA (Regression with ARIMA errors) model extends the ARIMA framework to allow for the inclusion of exogenous regressors, enabling the modelling of both the time series dynamics and the effects of external variables. In a RegARIMA model, the observed series Y_t is modelled as a linear combination of k regressors $X_{1,t}, \ldots, X_{k,t}$, plus an error term that follows a (seasonal) ARIMA process:

$$Y_t = \beta_0 + \beta_1 X_{1,t} + \cdots + \beta_k X_{k,t} + z_t, \tag{2}$$

where z_t is assumed to follow a seasonal ARIMA$(p, d, q)(P, D, Q)_\tau$ process as described above. Equivalently, the RegARIMA model can be written as

$$\phi_p(B)\Phi_P(B^\tau)(1-B)^d(1-B^\tau)^D z_t = \theta_q(B)\Theta_Q(B^\tau)\varepsilon_t, \tag{3}$$

with $z_t = Y_t - \beta_0 - \beta_1 X_{1,t} - \cdots - \beta_k X_{k,t}$. This approach allows the inclusion of external factors, such as calendar effects or other covariates and an automatic outlier detection based on the algorithm described by Chen and Liu (1993). The full RegARIMA model is also used to forecast the raw and linearised series, a practice that supports the revision policy frameworks promoted by Eurostat (2024) and lessens the reliance on asymmetric filters in the methods used to estimate the seasonal component.

2.1 Methods for seasonal and calendar adjustment

Let the time series Y_t be defined by an unobservable components model so that

$$Y_t = T_t + S_t + C_t + I_t \tag{4}$$

where T_t denotes the trend-cyclical component, S_t the seasonal component, C_t the calendar component, and I_t the irregular component. The objective of calendar and seasonal adjustment is to estimate and remove the calendar-related effects (S_t and C_t). In some cases, the model specified in Equation (4) might be augmented by further components, such as interaction effects.

To achieve the decomposition, all methods discussed below combine RegARIMA models – which capture selected calendar effects – with methods used to estimate the seasonal component that correct for periodically recurring patterns.

Seasonal adjustment of daily time series The daily seasonal adjustment (DSA) procedure of Ollech (2021) removes calendar and other periodic effects from daily time series in an iterative, multistep procedure. Its successor, DSA2, includes the recently developed extended airline model, which permits high-order RegARIMA specifications as well as the simultaneous estimation of several periodic patterns.[1] The method is built on the ARIMA (0,1,1)(0,1,1) model – better known as the Airline model introduced by Box and Jenkins (1970) – which has proven highly effective in capturing the dynamics

[1]The DSA2 method is implemented in the `dsa2` R package that can be downloaded from https://github.com/bbkrd/dsa2.

of the majority of seasonal time series (Fischer and Planas, 2000; Ollech and Webel, 2020) and has attractive features in the context of seasonal adjustment and time series decompositions (Aston and Koopman, 2006).

Let the backshift operator B for a non-integer $\tau \in \mathbb{R}_+$ be defined via a first-order Taylor approximation as follows:

$$B^\tau = (1 - \alpha_{\tau - \lfloor \tau \rfloor}) B^{\lfloor \tau \rfloor} + \alpha_{\tau - \lfloor \tau \rfloor} B^{\lceil \tau \rceil}. \tag{5}$$

With this definition, B^τ can be substituted into Equation 3 to obtain the extended airline model, which is capable of accommodating non-integer seasonal periodicities. Furthermore, by allowing for multiple seasonalities, the extended airline model enables the estimation of calendar effects while directly accounting for both weekly and annually recurring patterns.

The sequence of adjustment steps in DSA2 is as follows:

1. Calendar adjustment using the extended airline model
2. Adjustment for the day-of-the-week
3. Adjustment for the day-of-the-month
4. Adjustment for the day-of-the-year

A further advantage of DSA2 is its modular design: different seasonal-adjustment engines can be employed in step 2, 3, and 4. In contrast, DSA is limited to the STL method of Cleveland, Cleveland, McRae, and Terpenning (1990), which decomposes a time series into seasonal, trend-cyclical and irregular components by iteratively applying (seasonal) loess regressions and weighted moving averages. DSA2 additionally supports X-11 and a SEATS-type procedure based on the extended airline model (cf. Ollech, 2022). Steps 2–4 can be iterated multiple times to ensure convergence of the estimated periodic components.

Seasonal adjustment of weekly time series The weekly seasonal adjustment (WSA) procedure[2], described in Ollech (2022), uses the extended airline model to estimate calendar effects. Then either an X-11- or SEATS-type decomposition is used to extract the seasonal component. As discussed, the extended airline model allows non-integer periodicities. This is somewhat of a necessity for weekly time series, given that the number of observations per year is 52.18 on average.

Seasonal adjustment of monthly and quarterly time series The adjustment routines implemented in Eurostat's seasonal adjustment software JDemetra+ for monthly and quarterly time series are X-13-ARIMA and TRAMO-SEATS.

Both routines combine a RegARIMA-based pre-processor with a dedicated seasonal adjustment method. In addition to the calendar correction, the pre-processing includes algorithms for an automatic selection of the ARIMA order as designed by Gómez and

[2]WSA is available in the wsa() function in the aforementioned `dsa2` package.

Maravall (2001), and an automated choice between additive and multiplicative decompositions.[3]

In TRAMO-SEATS, the pre-processing is followed by the SEATS (Signal Extraction in ARIMA Time Series) for the estimation of the seasonal component. SEATS is a fully model-based procedure: using the estimated ARIMA specification, it constructs Wiener-Kolmogorov optimal filters that decompose the linearised series into four latent components: seasonal, trend-cycle, transitory (short-term) movements and an irregular remainder.

In X-13-ARIMA, the method X-11 is employed. This semi-parametric method iteratively applies weighted non-seasonal and seasonal moving averages to extract the seasonal, trend-cyclical and irregular components.

2.2 Frequency-dependent limitations

Table 1: Estimation possibilities and limitations across temporal resolutions

	Daily	Weekly	Monthly	Quarterly
Calendar effects	✓	✓	✓	(✓)
Time-varying weekly recurring pattern	✓	✗	(✗)	(✗)
Monthly recurring pattern	✓	(✓)	✗	✗
Interaction between weekday and calendar effects	✓	(✓)	(✗)	(✗)
Bridging days	✓	(✓)	(✓)	(✓)

Remark: ✓ = possible; (✓) = possible with caveats; (✗) = mostly not possible; ✗ = not possible.

As discussed above, the observable characteristics of seasonal time series depend on the temporal resolution. Table 1 highlights the key differences (for an exhaustive treatment, see Ollech, 2022).

Although calendar effects – influences related to the calendar constellation that are neither seasonal nor strictly periodic[4] such as moving holidays and the number of working-days in a period – can in principle be modelled with regression terms whenever suitable regressors are available, their estimation is weak at quarterly or lower periodicities. The variation in the length of a quarter, and hence in the number of working-days it contains, is too small to yield reliable identification. The same lack of variation hampers the detection

[3] The pre-processing routines in X-13-ARIMA and TRAMO-SEATS differ only in minor details. TRAMO-SEATS re-estimates the joint impact of all detected outliers more frequently, whereas X-13-ARIMA uses a higher default critical value for outlier inclusion. In the automatic ARIMA order selection they employ distinct unit-root cut-offs, impose different maximum ARMA orders, and vary in the use of exact versus conditional likelihood. Their automatic calendar-effect modules also rely on different statistical tests.

[4] Strictly speaking, the Western calendar repeats every 400 years, and with few exceptions even every 28 years. However, this periodic behaviour is, to the best of our knowledge, never exploited in practice, so we follow this terminology.

of any residual calendar effects, so the corresponding diagnostic tests are inevitably less powerful than we would like.

In daily time series, the effect of working-days manifests itself as a day-of-the-week pattern. This can be tackled by methods that allow this periodic effect to change over time. The RegARIMA models used in official statistics to estimate the working-day effect in monthly or quarterly time series generally assume a time-constant effect.[5] In weekly time series it will usually not be feasible to identify intra-weekly patterns.

Similarly, monthly recurring patterns can usually only be estimated from sub-monthly time series. Because there is a non-integer number of weeks per months, the estimation of such effects in weekly time series will usually be unreliable, especially if time-varying effects have to be estimated.

Some interaction effects can be captured by including suitable regressors in the model. However, for time series with a frequency lower than daily, it is often nearly impossible to identify such effects. This difficulty arises from the large number of additional regressors required and the limited variation in relevant calendar constellations present in lower-frequency series.

According to the official guidelines for seasonal adjustment (cf. Eurostat, 2024), bridge days should not be modelled or adjusted for in monthly or quarterly time-series. In lower-frequency data their impact is hard to isolate from other calendar effects, so any estimate would be unreliable.

For higher-frequency time series the situation is different. In daily time series, the influence of bridge days is often very apparent and suitable regressors can be constructed. In weekly data the bridge days typically fall in the same week as the associated holiday, making it unavoidable to include their effect in the estimation of the holiday effect.

3 Diagnostics

The purpose of seasonal and calendar adjustment is to remove all seasonal and calendar fluctuations from a time series, without extracting any of the other components into the seasonal and calendar effect, respectively. An adequately adjusted series should therefore display neither statistically significant seasonality (so-called residual seasonality) nor signs of over-adjustment. As McElroy and Roy (2022) discuss, however, the seasonal adjustment methods routinely used by official statistical agencies often yield adjusted series whose spectra exhibit pronounced troughs at the seasonal frequencies and corresponding negative spikes in the autocorrelation function. This phenomenon is typically dismissed as “a curious artefact of optimal signal extraction” (McElroy and Roy, 2022, p. 269) and accordingly, relatively few diagnostic tools exist for detecting seasonal over-adjustment.

Residual seasonality, in contrast, is always considered a problem to be avoided. Comparing residual-seasonality diagnostics can therefore help determine the most suitable level of temporal aggregation.

[5] Time-varying calendar effects can be handled, for instance, with rolling-window estimation, extended procedures such as STAHL (Haller, Daniel, and Bellone, 2025), structural time series models (Harvey, 1989) or enhancements of existing methods (For ARIMA and ETS such enhancements are implemented in the `smooth` package in R, see Svetunkov, 2017). Usually the identification of time-varying calendar effects is less stable than for time-constant effects.

Within the regression framework employed to estimate calendar effects, tests for residual calendar influences are straightforward. Appropriate regressors have to be constructed and can then be evaluated in a RegARIMA model fitted to the adjusted, aggregated series. The chosen ARIMA specification and its accompanying diagnostic measures can then be compared across alternative aggregation levels to identify the temporal resolution that delivers the best statistical fit.

In official statistics, seasonal adjustment results need to be computed, every time new data become available. Because each new observation adds information, previously published values typically change, especially the most recent ones. Analysing the size, direction, and pattern of these revisions and, to the extent that it does not reduce the quality of the adjustment itself, minimising them, is therefore of considerable practical importance.

Finally, many higher-frequency series are used as proxies for particular sectors of the economy or for lower-frequency aggregates. The strength and stability of their relationship to the target series thus provide an additional diagnostic criterion when selecting an appropriate aggregation level.

The diagnostics presented here should generally be applied to the data's final temporal aggregation level to make it possible to compare results obtained at different temporal resolutions.

3.1 Residual seasonality

McElroy and Roy (2022) examine several diagnostics for detecting residual seasonality, concentrating on the autocorrelation based QS test, spectral-based tests, and the ROOT diagnostic. For the latter they recommend using at least 20 years of data, noting that shorter samples render the test noticeably mis-sized. This requirement considerably restricts the practical applicability of the ROOT diagnostic, because many economic time series fall short of the two-decade threshold, especially if one limits the sample to periods free of methodological breaks or other structural changes in the data-generating process. In fact the ESS guidelines on seasonal adjustment recommend to split time series that are 20 years or longer (Eurostat, 2024)

Ollech and Webel (2022) demonstrate that the QS test and the ANOVA-type Friedman test complement each other in detecting different manifestations of seasonality. Using them in tandem increases the accuracy of classifying a series correctly as seasonal or non-seasonal. As discussed above, standard seasonal-adjustment procedures produce series with negative seasonal autocorrelation. The QS test statistic contains an indicator function that sets the statistic to zero whenever the first seasonal autocorrelation is negative. Accordingly, the p-value of the test applied to seasonally adjusted series is often exactly 1. In that case the QS test may give identical results for competing adjustment results.

So far, seasonality tests have not yet been evaluated systematically for higher-frequency time series. Ollech (2021) uses visual inspection of the spectrum to assess the presence of seasonal and periodic patterns in daily time series.

A thorough assessment of residual seasonality should therefore combine formal seasonality tests – such as the QS and Friedman tests – with spectral-based diagnostics.

When selecting the temporal resolution, tests for residual seasonality should be conducted at least at the aggregation level at which the data will be disseminated or analysed.

For example, if we adjust a daily time series and then aggregate it to a monthly frequency, the tests should be applied at the monthly level of temporal aggregation.

3.2 Residual calendar effects

As discussed above, residual calendar effects can be assessed by including suitable calendar regressors in a RegARIMA model for the already calendar- and seasonally-adjusted, aggregated series (For a detailled discussion about an appropriate construction of calendar regressors, we refer to Deutsche Bundesbank, 2012). The significance of these variables can then be tested with the usual t- or F-statistics. Possible regressors include:

- the generic working-/trading-day variables supplied by JDemetra+, and
- user-defined holiday or working-day variables that reflect a country's or a region's calendar.

For higher-frequency time series, additional calendar phenomena may be relevant. Ollech (2022) documents several such effects and shows how they can be handled within the adjustment procedure. Cross-seasonal effects, i.e. interdependencies between some or all periodic and (fixed) holiday effects, are modelled in DSA2 using dedicated regressors (see Section 2.1), these are often mere dummy variables. After they are compiled for the higher-frequency series, the regressors can be aggregated to the lower frequency and inserted into the RegARIMA model, allowing a direct test for any remaining cross-seasonal structure. The accuracy and power of this kind of statistical tests will decrease with the number of effects included. This can only partially be counter-acted by sensibly merging regressors.

Bridge days warrant special attention because current guidelines give contrasting recommendations for their treatment at different levels of aggregation (Eurostat, 2024). Especially in daily time series, failing to correct for bridge days typically leaves conspicuous spikes, so their adjustment is advisable. In monthly or quarterly series, however, their influence is diluted and very difficult to estimate with any precision. Consequently, searching for residual bridge-day effects in aggregated (e.g., monthly or quarterly) data may often be unnecessary and sometimes even be misleading.

To detect calendar effects, graphical tools can also be employed. Spectral analysis may be particularly useful in this context. For monthly time series, relevant spectral peaks associated with trading-day effects occur at 0.348 cycles per month, which is an alias of (30.4375 days per month / 7 days per week) = 4.348, as noted by Soukup and Findley (1999). Similarly, for quarterly time series, the calculation (91.3125 days per quarter / 7 days per week) = 13.04464 leads to spectral peaks at 0.04464 cycles per quarter. It is easy to see then, why this peak for quarterly time series may be over-shadowed by a dominant trend. Note also that the second quarter always spans over 91 days so – barring any additional holiday corrections – each day of the week occurs exactly 13 times. As a result, the variability in the number of trading-days per quarter is quite limited, making the identification of trading-day effects more challenging. Therefore, this graphical tool is generally more effective for monthly time series.

An alternative is the starting-day-of-the-period plot proposed by Cleveland and Devlin (1980). Using the seasonally (and trend) adjusted series, for all periods of equal length (e.g. all quarters with 92 days), a boxplot is created for each possible starting day of the week: one for all periods beginning on a Monday, another for those starting on a Tuesday, and so on. Examples are shown in Figure 7 and Figure 8. These plots can be used to identify patterns related to the day of the week.

3.3 Revisions

Revisions are an unavoidable side-effect of seasonal adjustment[6]: each time new raw observations become available, the previously published adjusted values may change. While the additional data generally improve the precision of the estimates, the attendant revisions can also alter earlier analyses and conclusions. To manage this trade-off, the Eurostat Guidelines on Seasonal Adjustment (Eurostat, 2024) outline several revision policies, taking into account specific properties of the series and the requirements of the data users. These policies share a common goal: minimising revisions that are economically irrelevant and thus unlikely to influence the interpretation of the current situation.

When conducting revision analysis, we must determine which data releases – known as vintages – we want to compare. The most common approaches are period-to-period revisions and first-to-final revisions.

Period-to-period revisions compare consecutive data vintages. For a series Y_t, which may be expressed in levels or as a growth rate, the period-to-period revisions are defined as

$$r_{t|T}^{\mathrm{PP}} = \hat{y}_{t|T} - \hat{y}_{t|T-1} \tag{6}$$

where $\hat{y}_{t|T}$ denotes the estimate for period t contained in the vintage released at time T, and $\hat{y}_{t|T-1}$ is the corresponding estimate published one release earlier.

First-to-final revisions measure the total change between the first release of an observation and the vintage that is regarded as final:

$$r_t^{\mathrm{FF}} = \hat{y}_{t|t_0} - \hat{y}_{t|T^*} \tag{7}$$

with $\hat{y}_{t|t_0}$ being the estimate first published at release date t_0, with $t_0 \geq t$ and $\hat{y}_{t|T^*}$ the estimate appearing in the final vintage issued at time T^*, with $T^* > t_0$. The horizon T^* is chosen far enough in the future that further revisions are assumed to be negligible.

To summarise the properties of any of the revision series $r_t^{\psi} \in \{r_t^{\mathrm{PP}}, r_t^{\mathrm{FF}}\}$ we consider standard forecast-error statistics.[7] For the T observations for which corresponding vintages are available,

[6]Seasonal adjustment methods that do not cause any revisions even if the results are re-estimated based on additional observations will typically not take advantage of all available information and thus not be efficient.

[7]Indeed, Nikolopoulos, Syntetos, Boylan, Petropoulos, and Assimakopoulos (2011) use an error statistic to select the optimal aggregation level for forecasts of demand arrivals. They use the mean absolute scaled error (MAsE) proposed by Hyndman and Koehler (2006) to select the optimal aggregation level for forecasts, where the scaling is done by dividing the absolute forecast error by the mean absolute error of the naive (seasonal) forecast. This kind of scaling will usually not be necessary for revision measures, and only applicable if an obvious method for comparison is available.

$$MAR = \frac{1}{T}\sum_{t=1}^{T}|r_t^{\psi}|, \tag{8}$$

$$RMSR = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(r_t^{\psi}\right)^2}. \tag{9}$$

The mean absolute revision (MAR) and root mean squared revision (RMSR) gauge the typical magnitude of a revision, RMSR giving greater weight to large revisions. These measures can be expressed in relative terms – e.g., as a percentage of the final estimate – to facilitate comparisons across series that differ in scale.

3.4 Relationship with target variable

Since the onset of the COVID-19 pandemic, economists have increasingly turned to higher-frequency data – either as higher-frequency proxies for certain branches of the economy (cf. Arshad and Beyer, 2023; Bricongne, Meunier, and Pouget, 2023; Lehmann and Möhrle, 2024) or as part of composite indicators intended to track the economy as a whole (e.g. Delle Monache, Emiliozzi, and Nobili, 2020; Eraslan and Götz, 2020; Fenz and Stix, 2021; Lewis, Mertens, Stock, and Trivedi, 2022; Lourenço and Rua, 2021; Wegmüller, Glocker, and Guggia, 2021; Woloszko, 2024). In most of these applications one or more lower-frequency target variables are available. The strength of the relationship between the proxy series Y_t and its target z_t therefore provides a natural measure for evaluating the quality of the seasonal adjustment.

The association between Y_t and z_t is most conveniently inspected with the sample **cross-correlation** function computed on stationary transformations (e.g. log-differences or percentage growth rates) of the variables (Kendall and Ord, 1983). In the context of identifying an appropriate temporal resolution for the seasonal adjustment the contemporaneous correlation of the growth rates already offers a parsimonious summary of their co-movement.

In a forecasting setting, a related metric is a test on **Granger causality**, which identifies the suitability of regressors as predictors in a regression, whose result might vary from the cross-correlational measures as they simultaneously include lags of the variables, if necessary (Granger, 1969).

Let m and n denote the chosen lag length of z_t and Y_t, respectively. The proxy is said to Granger-cause the target if, in the augmented autoregression

$$z_t = \alpha_0 + \alpha_1 z_{t-1} + \ldots + \alpha_m z_{t-m} + \beta_0 Y_t + \ldots + \beta_n Y_{t-n} + \varepsilon_t \sim \text{i.i.d. } (0, \sigma^2), \tag{10}$$

the null hypothesis $H_0 : \beta_0 = \ldots = \beta_n = 0$ is rejected, usually with an F- or Wald test. Because the regression controls for the target's own lags, the outcome of the test can differ markedly from simple contemporaneous correlations.

In some forecasting applications, it may even be feasible to include the different seasonal adjustment results into the evaluation stage. The prediction-accuracy metric can then inform the choice of the most appropriate temporal resolution.

An indicator of directional co-movement is the sign **congruence**, which quantifies the directional agreement between the series (see e.g. Eurostat, 2018). The sign congruence is defined as the number of times that the growth rate of the series share the same sign, i.e.

$$\text{Congruence} = \frac{1}{T}\sum_{t=1}^{T} \mathbb{1}_{\{\text{sign}(Y_t)=\text{sign}(z_t)\}}, \tag{11}$$

indicating that their short-term trajectories are synchronised. This measure is particularly useful for understanding whether the proxy and target consistently signal movement in the same direction.[8] The congruence can be normalised by comparing it to the naïve sign forecast, which always predicts the most frequently occurring sign – positive if positive values are more common, and negative otherwise. The scaled congruence is then given by

$$\text{Scaled congruence} = \frac{\text{Congruence}}{\max\left(\frac{1}{T}\sum_{t=1}^{T}\mathbb{1}_{\{\text{sign}(z_t)=+\}}, \frac{1}{T}\sum_{t=1}^{T}\mathbb{1}_{\{\text{sign}(z_t)=-\}}\right)} \tag{12}$$

When the series are expressed in comparable units, distance metrics such as the mean absolute error or the root-mean-square error (cf Section 3.3) of their growth rates offer an easily interpretable measure of the average discrepancy between proxy and target.

Taken together, these statistics characterise the strength, direction and predictive content of the relationship, all of which can guide the choice of the temporal resolution.

3.5 Comparison of ARIMA models

Earlier studies (see Section 2) suggest a certain alignment of the ARIMA models at different temporal resolutions. If, at one particular scale, the selected model is noticeably more complex than those at other scales, that time-aggregation level may be unsuitable. Before drawing a final conclusion, it may be worthwhile to refit the series with alternative ARIMA specifications. Often one can find a simpler model consistent with the models used at the other time scales that performs similarly, even if formal statistics initially favour the more complicated choice.

4 Illustration

4.1 Simulation

Our simulation study follows the procedure proposed by Ollech (2021) and implemented in the R package `tssim` (Version 0.2.7) to generate daily time series.[9] Each simulated series $\{Y_t\}$ is constructed multiplicatively as

$$Y_t = Y_t^{(\text{SA})} \cdot S_t^{(7)} \cdot S_t^{(365)}, \qquad Y_t^{(\text{SA})} \sim \text{ARIMA}(3,1,1), \tag{13}$$

[8] Pesaran and Timmermann (1992) derive a non-parametric tests of the accuracy of a regressor as a predictor of the sign of the growth rate of a target variable.

[9] To assess the robustness of our findings, we additionally applied the simulation algorithm developed by Cuevas and Quilis (2023), which is based on the time series models proposed by Harvey and Shephard (1993). The results were qualitatively similar to those obtained with our primary approach and are available upon request.

where $Y_t^{(\mathrm{SA})}$ is the seasonally adjusted component, and $S_t^{(7)}$ and $S_t^{(365)}$ are the weekly and annual seasonal factors, respectively. Calendar effects are deliberately excluded, as this is beyond the scope of this study.

For $\tau \in \{7, 365\}$ the initial seasonal factor $\tilde{S}_t^{(\tau)}$ is given by:

$$\begin{aligned} \tilde{S}_t^{(\tau)} &= S_{t-\tau}^{(\tau)} + \varepsilon_t, \\ \varepsilon_t &= \beta\, \varepsilon_{t-1} + u_t, \qquad u_t \sim \mathcal{N}(0, \sigma_u^2). \end{aligned}$$

Its magnitude is adjusted to a user-chosen level,

$$S_t^{(\tau)} \;=\; \frac{\tilde{S}_t^{(\tau)}}{\sigma_{\tilde{S}^{(\tau)}}}\, \sigma_{S^{(\tau)}},$$

where $\sigma_{\tilde{S}^{(\tau)}}$ denotes the empirical standard deviation of $\tilde{S}_t^{(\tau)}$.

We generate 1 000 independent series. For each replication the parameters are drawn with replacement from the following sets:

$$\begin{aligned} \sigma_{S^{(\tau)}} &\in \{1, 2, \ldots, 20\}, \\ \sigma^2_{u^{(7)}} &= \sigma_{S^{(7)}}/\omega^{(7)},\ \omega^{(7)} \in \{2,\ 20,\ 200\}, \\ \sigma^2_{u^{(365)}} &= \sigma_{S^{(365)}}/\omega^{(365)},\ \omega^{(365)} \in \{0.5,\ 5,\ 50\}, \\ \beta^{(7)} &\in \{0.1,\ 0.01,\ 0.001\}, \\ \beta^{(365)} &\in \{0.2,\ 0.02,\ 0.002\}. \end{aligned}$$

Because the data are simulated, the "true" seasonally adjusted series is known; we will consider it the target series. Even though we do include the relationship to the simulated seasonally adjusted series, the goal of this simulation study is not to determine which pairing of seasonal-adjustment method and temporal aggregation yields the overall best adjustment. Yet, it would be worth investigating, how combinations of the single adjustment results may optimise the accuracy.[10] The main objective of the present study is to analyse how the adjustment results relate to one another across different levels of temporal aggregation.

The results included in Table 2 and Table 3 show the correlations of the growth rates at different levels of temporal aggregation, restricted to the X-11 and SEATS based procedures, respectively. The presentation of all relationships, including across methods, can be found in Table 9 in Appendix A.

For both X-11 and SEATS, the highest correlations are observed between neighbouring aggregation levels. In the case of X-11, these correlations remain consistently high across all levels, likely due to the limited number of seasonal filters available compared to

[10] In the context of forecasting, different studies have investigated the benefits of combining the results from different levels of temporal aggregation to obtain improved forecast (e.g. Athanasopoulos, Hyndman, Kourentzes, and Petropoulos, 2017; Kourentzes, Rostami-Tabar, and Barrow, 2017). We do not find that combining the seasonal adjustment results based on different temporal resolutions generally improves the accuracy compared to the best single seasonal adjustment results.

Table 2: Correlations between growth rates for seasonal adjustment results at different levels of temporal aggregation, only X-11 results

	Daily	Weekly	Monthly	Quarterly	Target
Daily	1.000	0.963	0.974	0.965	0.943
Weekly	0.963	1.000	0.974	0.967	0.913
Monthly	0.974	0.974	1.000	0.982	0.931
Quarterly	0.965	0.967	0.982	1.000	0.924
Target	0.943	0.913	0.931	0.924	1.000

Note:
The Pearson-Bravais-type correlations are computed on seasonally adjusted series aggregated to quarterly frequency.

Table 3: Correlations between growth rates for seasonal adjustment results at different levels of temporal aggregation, only Seats results

	Daily	Weekly	Monthly	Quarterly	Target
Daily	1.000	0.922	0.889	0.883	0.846
Weekly	0.922	1.000	0.891	0.883	0.841
Monthly	0.889	0.891	1.000	0.959	0.947
Quarterly	0.883	0.883	0.959	1.000	0.929
Target	0.846	0.841	0.947	0.929	1.000

Note:
The Pearson-Bravais-type correlations are computed on seasonally adjusted series aggregated to quarterly frequency.

SEATS. For SEATS, however, there appears to be a distinct separation between higher and lower frequencies: daily and weekly series are highly correlated with each other (cor(Daily, Weekly) = 0.922), as are monthly and quarterly series (cor(Monthly, Quarterly) = 0.959). This is probably because, for daily and weekly time series, only the extended airline model can be applied, whereas for monthly and quarterly series, a broader set of ARIMA models is available. In contrast, X-11 exhibits a more gradual decline in correlation as the difference between aggregation levels increases. For practitioners working only with lower frequency time series, it is important to note that discrepancies between aggregation results are not limited to higher frequencies; monthly and quarterly series are also not perfectly correlated. This aligns with the findings in Table 10, which show that the monthly and quarterly aggregation levels yield the same sign in less than 95% of observations. Consequently, analysts working with both monthly and quarterly time series would encounter differing signals regarding the direction of the most recent development in approximately 5% of cases as a result of using different temporal resolutions.

4.2 Real world examples

Original values We examine the case for temporal disaggregation with 2 real-world examples. You can see these time series in Figures 1 and 2. Both time series are observed at a daily frequency and are also aggregated to weekly, monthly and quarterly series. Note that in the figures the aggregated variants of the time series have been shifted by 3, 15 and 45 days, respectively, to align them with the mid point of each period.

The first real-world time series in our analysis shows the development of the German electricity consumption since January 2015. These data are provided by the Bundesnetzagentur, the main regulative authority for infrastructure, energy, telecommunications, post and railways in Germany. Quite clearly these data contain multiple seasonal patterns. Except for a sharp decline between Christmas and New Year's, power consumption is the highest during the winter months. It then decreases during spring and reaches its lowest levels in the summer months of July, August and September. At this time electricity demand is usually around 30% lower than in January. After this low it increases again until the end of the year.

There is no discernible monthly recurring pattern, but a pronounced weekday effect. From Monday to Friday, power consumption is roughly constant at a fairly high rate. On Saturday it drops noticeably until it reaches its low point on Sunday, where demand is approximately 20% lower than during the week. In addition to these seasonal patterns there are also longer-term trends in the data. Most notably, there is a reduction in power consumption during 2020, when economic activity was hampered because of the COVID-19 pandemic. After a short-lived recovery in 2021 and 2022 we can since observe a downward trend, roughly coinciding with the ongoing recession of the German economy.

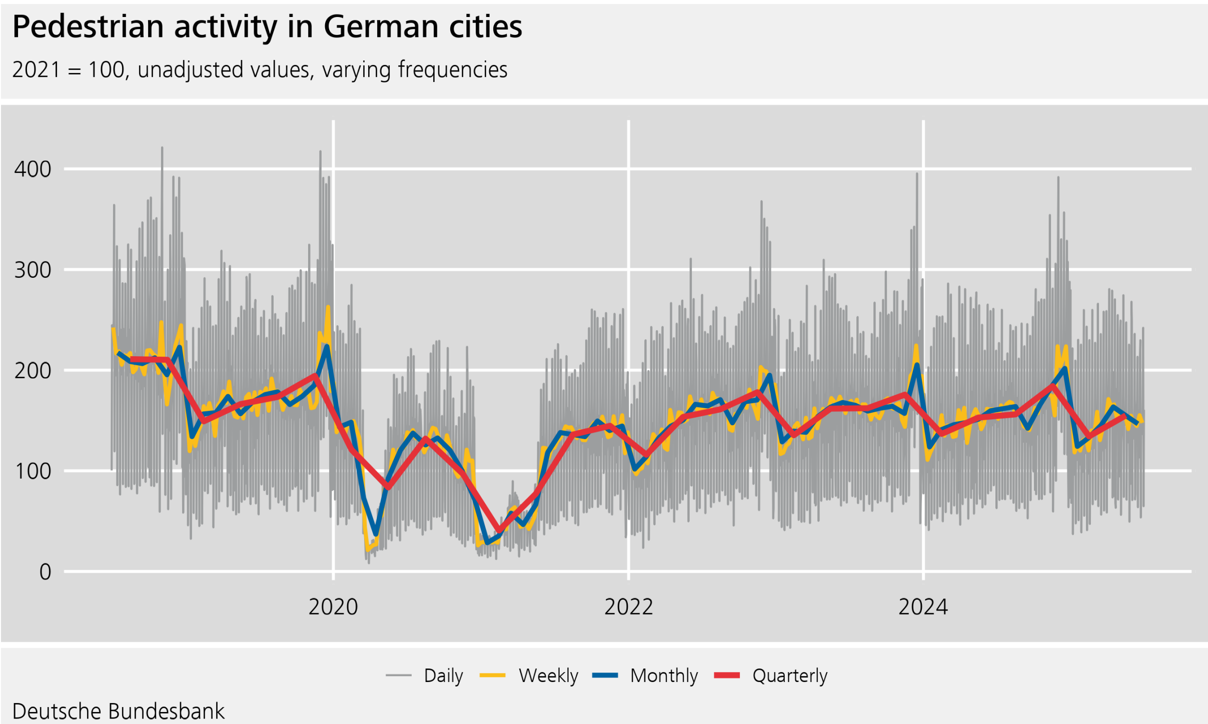

Figure 2: Pedestrian activity (unadjusted)

The second real-world time series estimates the pedestrian frequency in German cities since July 2018 using lasers. These data originate from hystreet.com GmbH, a private data provider that uses laser scanners to track foot traffic at prominent metering points in inner cities in Germany. The aggregated times series used in our analysis is a population-weighted average of the raw data from the metering points in 21 large German cities and is indexed to its 2021 average.

Albeit not as pronounced as in the electricity data, the pedestrian data too contain multiple seasonal patterns. First, we can see an increase of foot traffic throughout the year that culminates in the weeks before Christmas. Second, we find that on Fridays pedestrian activity increases by about 25% in comparison to the near constant level of foot traffic from Monday to Thursday. On Saturdays, the number of pedestrians increases by an additional 50%. On Sundays, when retail stores are closed in Germany, is about 40% of an average weekday or 75% lower than on Saturdays.

As with the electricity data, there is a significant downward trend of activity during the Coronavirus pandemic. The effects are however much stronger and unfold more rapidly. After the start of the first COVID-19 lockdown in mid-March 2020, within just one week the number of pedestrians shrinks by more than 2/3 in comparison to pre-pandemic levels. In mid-December 2020, when Germany entered its second lockdown period, the same pattern repeats once more.

In addition to these seasonal patterns, we find that both time series contain distinct dips and spikes pertaining to calendar-related effects such as holidays and bridge days. For example, all else equal, on Easter Monday or Pentecost Monday, pedestrian activity is 80 to 90% lower than on other days of the year.

Seasonal adjustment Section 2 describes the methods used to obtain seasonally adjusted data. The daily time series is adjusted using the X-11, SEATS and STL methods as implemented in the `dsa2` (Version 0.2.25). The weekly aggregates are adjusted using WSA with the X-11 and SEATS methods of the same package. For the monthly and quarterly series we use the X-13-ARIMA and TRAMO-SEATS routines of the `RJDemetra` (Version 0.2.8). Thus, in total we conduct nine sets of seasonal and calendar adjustments.

Naturally, there are more observations available when working with the daily data as opposed to the weekly, monthly or quarterly versions of the data. For the electricity time series we have a total of 3,834 daily observations, i.e. 10.5 years. The shorter pedestrian activity time series contains 2,557 daily observations, i.e 7.0 years. The number of observations in the weekly, monthly and quarterly aggregates, is lower by factors of 7, 30.4 and 91.3, respectively. Throughout all frequencies we work with a multiplicative transformation and a seasonal ARIMA model of order (0,1,1)(0,1,1).[11]

We vary, however, the choice of calendar regressors, depending on the level of temporal aggregation. For the daily frequency, we use 2 large sets of custom-made calendar regressors to account for all major public holidays in Germany and bridge days associated with them. To take into account interactions between the holiday and weekday components, we use additional regressors to account for instances where holidays fall on particular days of the week. In addition, we include further regressors for days that are not associated

[11] Notice that for the SEATS-adjustment of the daily time series we adjusted the logs of the original values, as the SEATS routine of the R package `rjd3highfreq` (Version 2.3.0) invoked by `dsa2` currently does not support multiplicative decompositions natively.

with public holidays in Germany but have nonetheless a significant impact on our time series of interest. For example, when adjusting the pedestrian data, we observe that foot traffic is significantly impacted by retail events such as Black Friday promotions or Sunday shopping exceptions during the Advent season.

When adjusting the weekly versions of our time series, we remove all regressors that pertain to days that occur within the same week, e.g. over the Easter weekend, or bridge days. Similarly, we remove those regressors that were included to account for holidays that fell on a weekend. At the monthly and quarterly frequency we use traditional working-day and trading-day regressors. As described in Section 4.2, the electricity series is characterized by significantly lower activity on Saturday and Sunday. As the difference between Saturday and Sunday as well as the difference among the working days is relatively small in comparison to the difference between working-days and the weekend, we use a working-day regressor when adjusting this time series. Conversely, pedestrian activity features great differences between weekdays, especially Saturdays and Sundays. In an ordinary week pedestrian activity spikes on Saturday and is the lowest on Sunday. Consequently, we use trading-day regressors to treat each day of the week separately.

In all estimations we work with automatic outlier detection. When adjusting the daily and weekly time series, we set the critical value for the detection of outliers equal to 8, to limit the number of automatically detected outliers to a moderate level. Most of these outliers are found at the start dates and end dates of the COVID-19 lockdowns in Germany, i.e. in the spring of 2020 and the 2020/2021 winter season. For the monthly and quarterly series we lower the critical value to 6, as the seasonal adjustment routines of the `RJDemetra` package, unlike `dsa2`, allow us to specify outliers manually. As with the adjustment of the higher-frequency series, we specify most of these outliers during the time of the COVID-19 pandemic.

When adjusting the daily electricity time series with the X-11 method we select a 3×15 and a 3×9 filter for the intra-weekly and the intra-yearly patterns. We do not estimate an intra-monthly effect, as such a pattern is not visible in the unadjusted values. Similarly, for the STL method we use a window of 25 and 15 observations in the local regressions for the seasonal components of the intra-weekly and the intra-yearly patterns but do not estimate an intra-monthly pattern. For the SEATS method we do not need to specify any filters, as the decomposition is fully based on an ARIMA model. For the X-11 and X-13-ARIMA methods used to adjust the weekly, monthly and quarterly time series we work with seasonal filters of length 3×5, 3×9 and 3×9, respectively.

When adjusting the daily pedestrian traffic time series with X-11, we work with a 3×15 filter for the day-of-the week effect and a 3×3 filter for the day-of-the year effect. Unfortunately, the `dsa2` package (in Version 0.2.25) does not allow for an even higher filter for the day-of-the week effect, which we would like to select to account for the very persistent intra-weekly pattern.[12] Analogously, for the STL approach, we include a relatively large number of 31 observations in the local regression for the day-of-the week effect and only 7 observations for the intra-yearly pattern. Because of the relatively low number of observations in this time series, `dsa2` auto-selects a 3×3 seasonal filter when adjusting the weekly pedestrian time series with X-11. Lastly, at the monthly and quarterly time frequencies, for which we work with the `RJDemetra` package, we use 3×9

[12] More precisely, the R package `rjd3x11plus` (Version 2.3.0) does not provide full asymmetric filters and thus the filter choice is limited by the number of seasonal cycles available.

seasonal filters in both cases.

Concerning the trend estimation, we select daily trend filters that encompass 13 observations for both the daily electricity time series and the daily pedestrian traffic time series and for both the X-11 method and the STL approach. For the yearly trend we work with filters that span 367 observations in the case of X-11 and 609 and 697 observations in the case of STL. For the weekly X-11 estimations, we smooth both time series over 55 observations and use a 13-term Henderson filter for the monthly and quarterly X-13 estimations. As was the case for the seasonal filters, there are no trend filters to be selected for the SEATS estimations.[13]

Adjustment results Based on these specifications we adjust each of the real-world time series 9 times: 3 times at the original daily frequency, and twice each at the weekly, monthly and quarterly frequencies. We then aggregate all non-quarterly adjusted time series to the quarterly level in order to facilitate comparison of the results.[14] Figures 3 and 4 show the adjusted series for the electricity time series, Figures 5 and 6 display the adjusted series for the pedestrian activity time series. For the electricity consumption, we can quite clearly see that the results are generally quite similar across the different adjustment methods. The differences appear to be a larger for the pedestrian activity. However, for this time series, in comparison to the electricity data, the intra-yearly fluctuations are much less dominant.

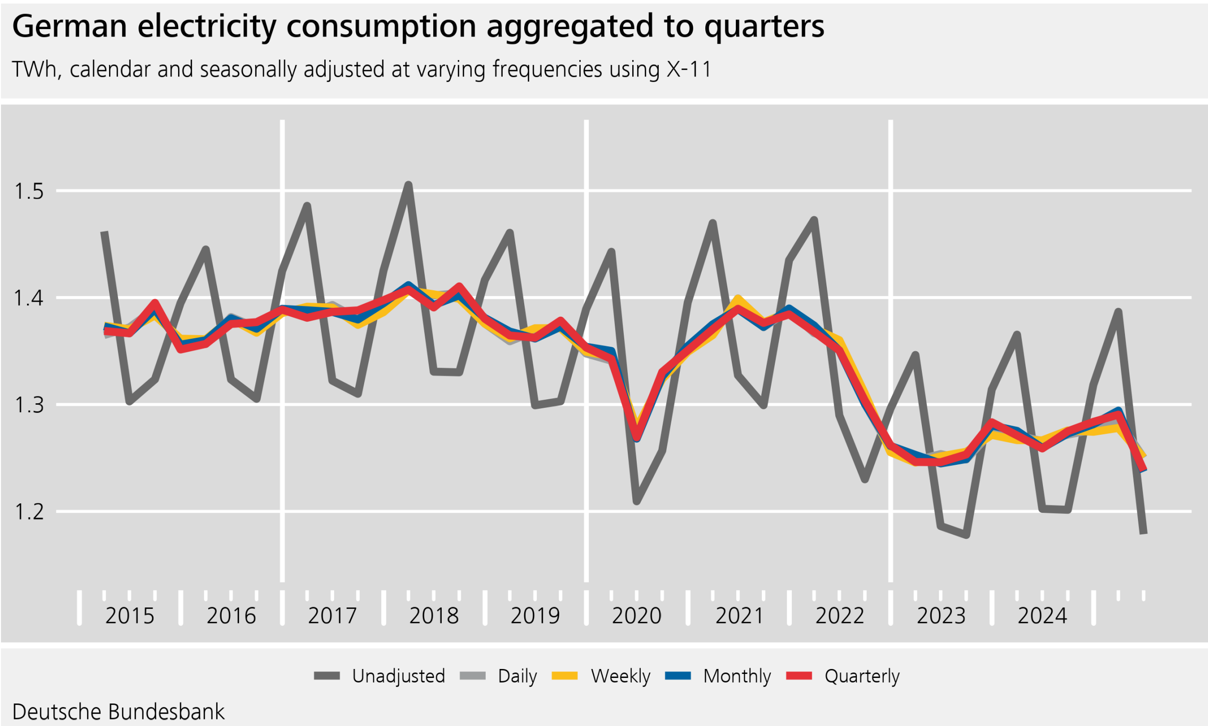

Figure 3: Electricity consumption (X-11)

[13]See the appendix for a detailed overview of all specifications and outliers used in the real-world examples.

[14]Note that the recommended temporal resolution at which seasonal adjustment is conducted may be differ if the target was a monthly instead of a quarterly time series.

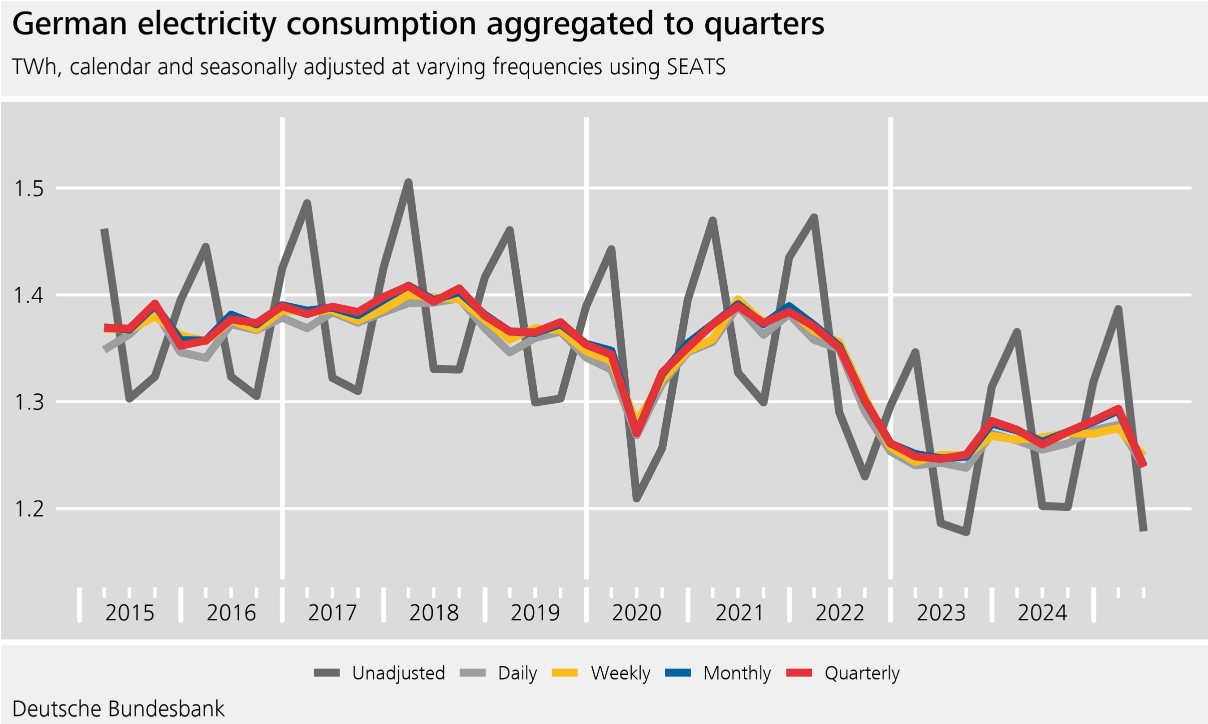

Figure 4: Electricity consumption (SEATS)

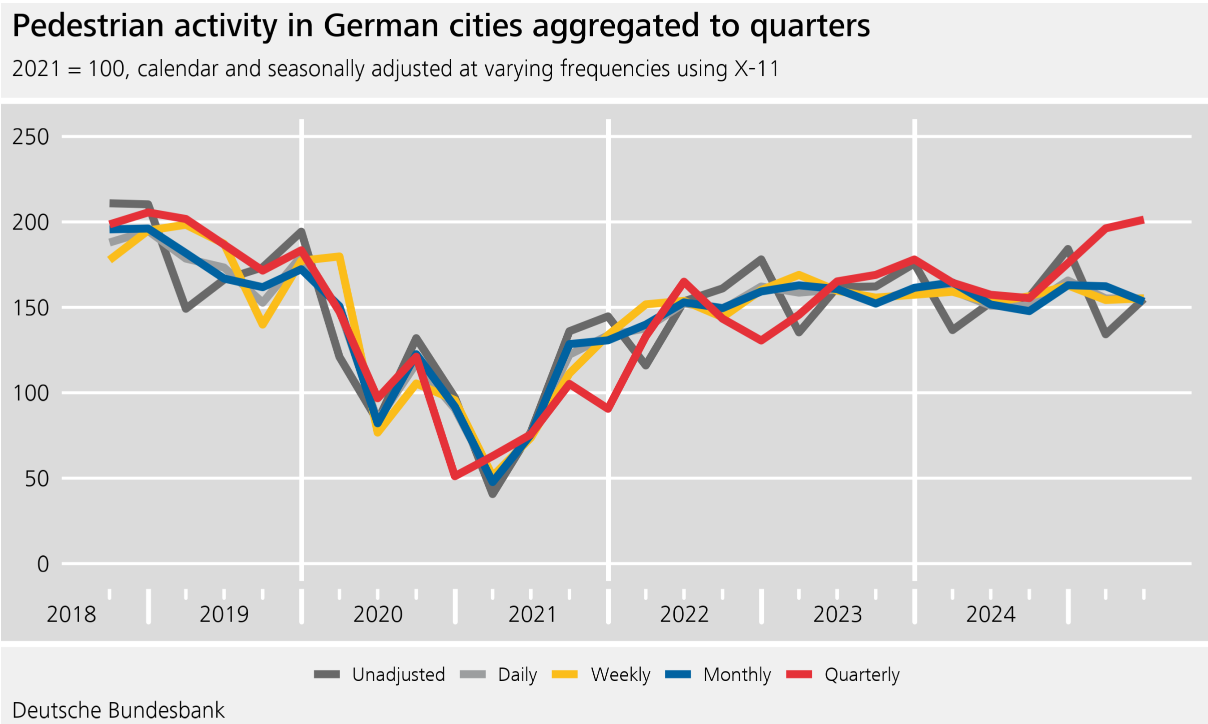

Figure 5: Pedestrian activity (X-11)

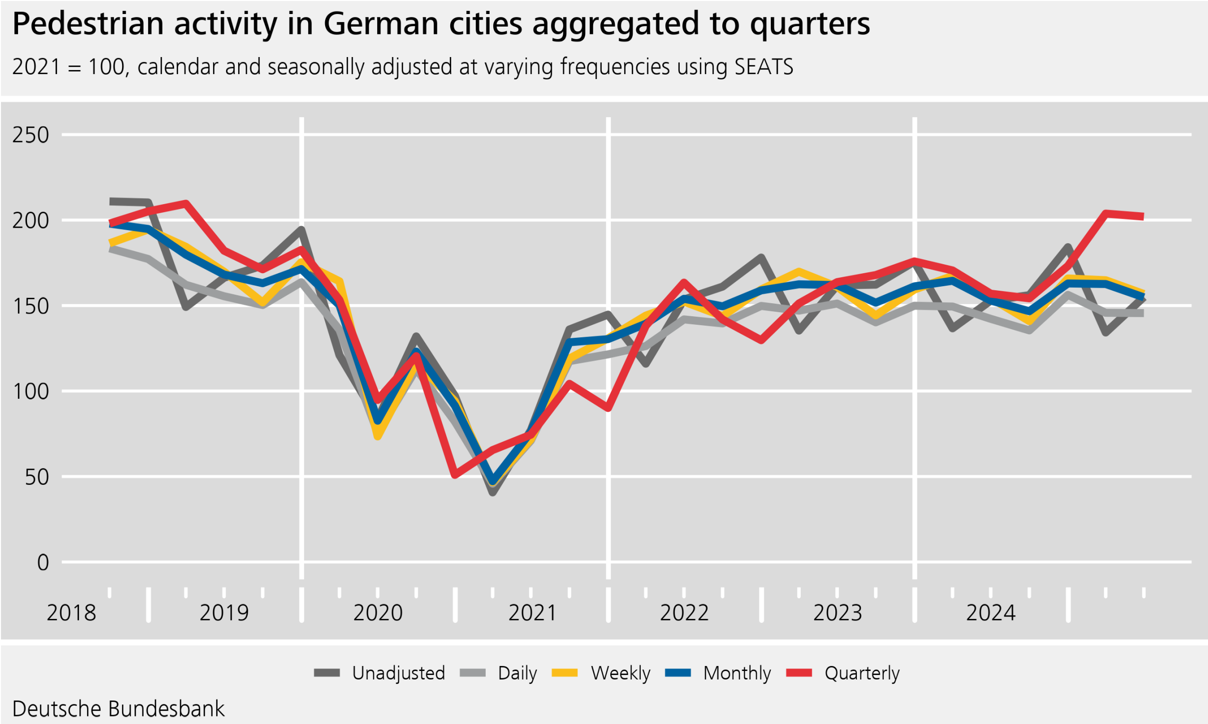

Figure 6: Pedestrian activity (SEATS)

Next, we apply the most relevant diagnostics discussed in Section 3 to the series, to analyse how the adjustments differ qualitatively across the nine cases for the 2 real-world examples. The results of these tests are displayed in Tables 4 and 5.

First, we test for **residual calendar effects**. To do so, we use the arima() function contained in the R package `stats` (Version 4.5.0) to estimate non-seasonal RegARIMA models of order $(0, 1, 1)$ for the adjusted time series that have been aggregated to quarterly frequency. For the electricity consumption time series we include a standard working-day regressor. For the pedestrian activity time series we use a trading-day approach, i.e. we specify individual regressors for the number of each day of the week contained in each quarter of our data sample. In doing so, we not only account for the difference between working-days and weekends but also the strong difference between Saturdays and Sundays (see Section 4.2).

In Tables 4 and 5 we report the p-values from a t-test of the statistical significance of the working-day regressor and F-tests of the joint significance of the trading-day regressors. For all specifications of the electricity series, we do not reject the null hypothesis of no working-day effect. Conversely, for the pedestrian series we reject the null of no trading-day effect in 8 of 9 cases. Only under the X-11 estimation at the weekly frequency do the tests find no evidence of untreated calendar effects.

Regarding trading-day effects, the spectra shown in Figures 10 and 11 mirror the results from the statistical tests. Pedestrian activity exhibits pronounced spikes at the frequencies associated with trading-day effects across all cases, even including the series adjusted at the weekly aggregation level using X-11. The spectrum for electricity consumption underscores the limitations of using spectral diagnostics for trading-day effects: the downward trend in the time series yields elevated spectral power at the lowest frequencies, which can obscure evidence of residual calendar effects. In this context, the starting-day-of-the-period plots – adjusted for calendar, season, and trend, and thus representing the irregular component – may offer a clearer perspective. As shown in Figures 7 and 8, differences between days of the week are apparent for both time series. Yet they also highlight the general difficulty of reliably identifying calendar effects in quarterly series: all subplots indicate higher irregular values for quarters beginning on a Wednesday. For the pedestrian activity, this calendar constellation occurs only once due to the short length of the series: in 2020-Q3, when restrictions following the first COVID-19 wave were lifted and pedestrian activity rebounded from its second-quarter levels. This single observation also drives the pattern observed in electricity consumption. Therefore, any apparent evidence of residual seasonality is likely spurious. This conclusion is supported by the statistical tests: if the 2020-Q3 observation is excluded, the F-tests for residual calendar effects in pedestrian activity are no longer significant.

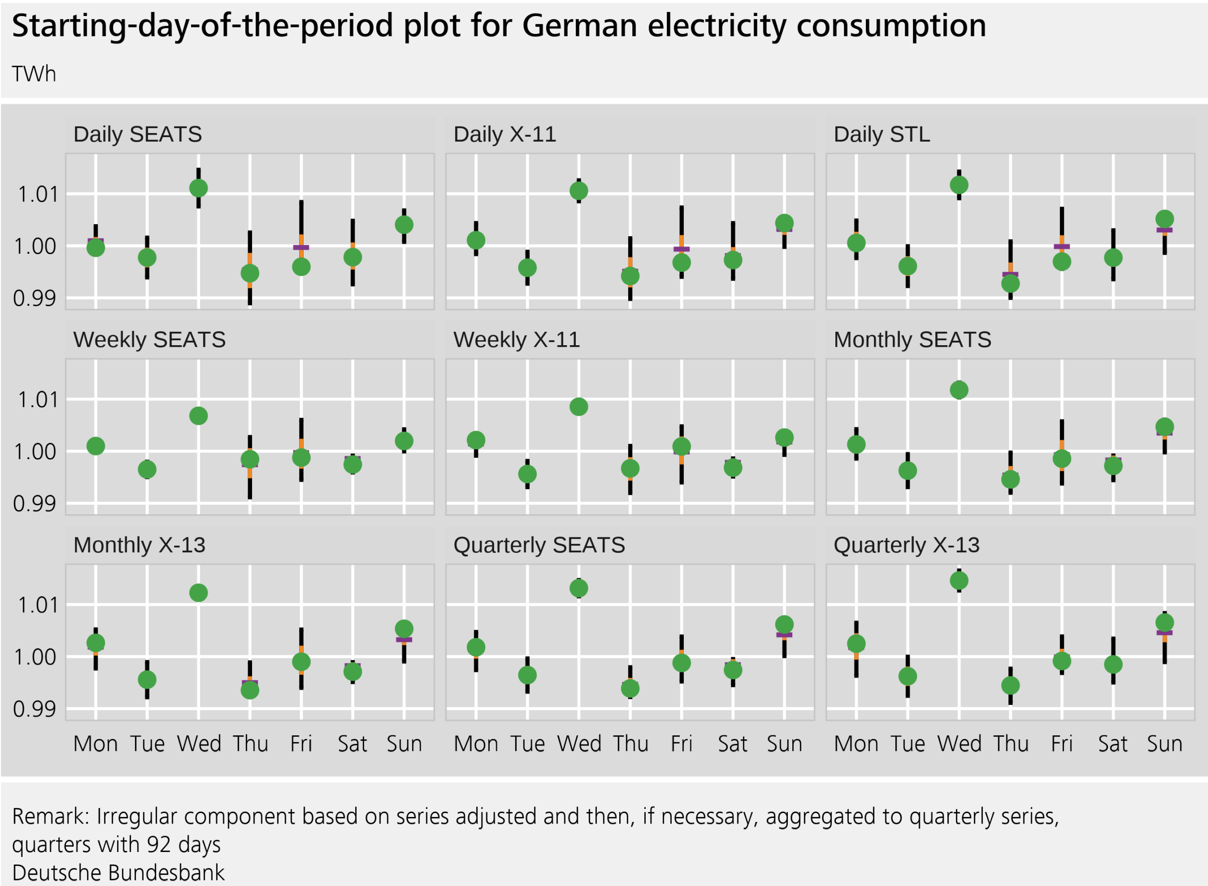

Figure 7: Starting-day-of-the-period plot for German electricity consumption

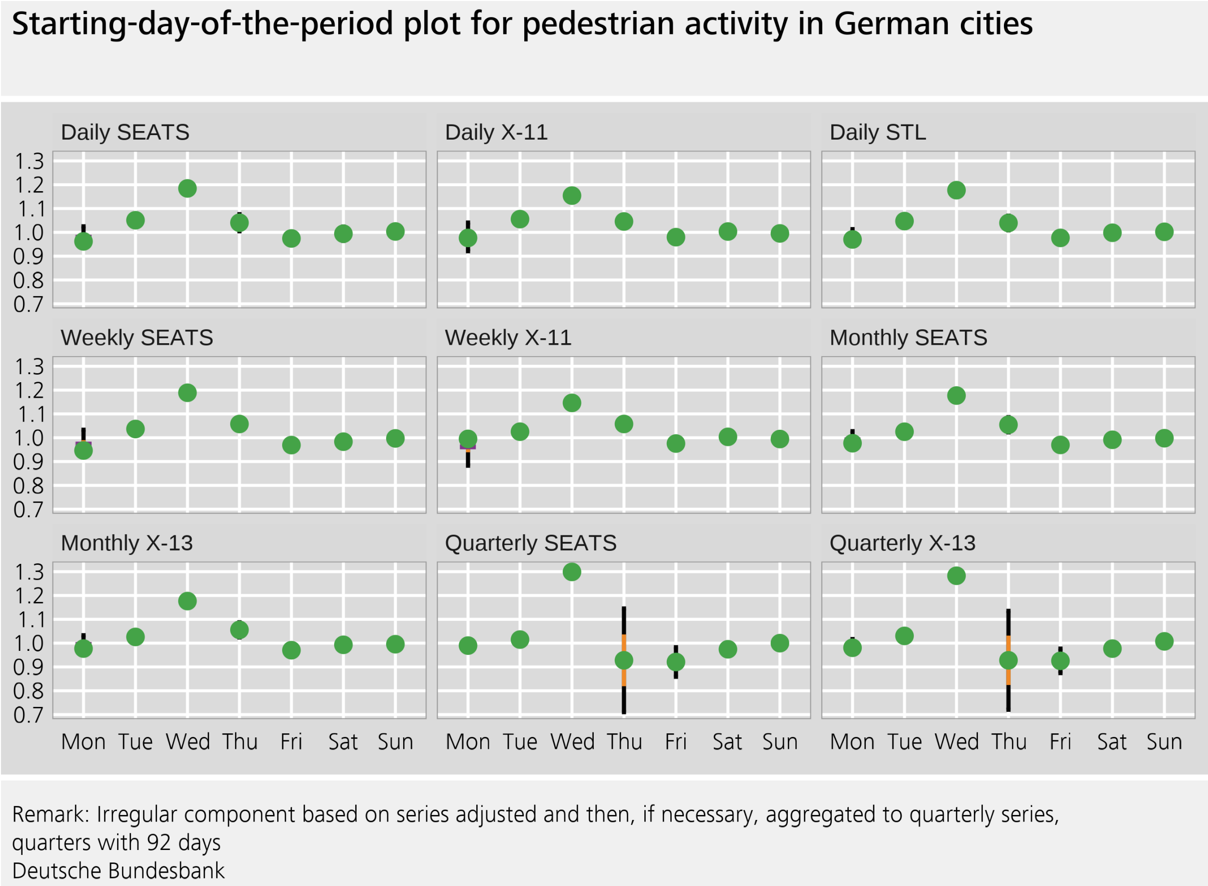

Figure 8: Starting-day-of-the-period plot for pedestrian activity in German cities

To test for **residual seasonality** in the adjusted and quarterised time series, we use 2 tests from the R package `seastests` (Version 0.15.4): the QS test and the Friedman test. Under the null, both tests assume that the series is non-seasonal.[15] As indicated by the p-values in Tables 4 and 5, we generally do not reject the null hypotheses at the 5% significant level for both the electricity consumption and pedestrian activity series. The p-values of the Friedman test are generally lower for pedestrian series, the X-11 adjustment at the daily temporal resolution is even marginally significant at the 10% level. These findings match the visual inspection of the adjusted time series. We do not discern seasonal patterns in any of our time series before aggregating them to quarters.

For electricity consumption, the spectral plots (Figures 10 and 11) show notable spikes at the seasonal frequency for the monthly and quarterly aggregations; even the series adjusted at the daily level using SEATS exhibits a minuscule uptick. In the case of pedestrian activity there are slightly elevated levels at the seasonal frequency in most of the spectra. As none of these spikes translate into visible residual seasonality in the adjusted series, we would not recommend ruling out any temporal resolution solely on the basis of these spectral diagnostics.

To examine the the revision propensity at each temporal resolution, we conduct a **revision analysis**. To this end, we first restrict our data sample to the first 5 years of each time series and then iteratively increase it by one quarter until the end of our time

[15]For a detailed discussion of these tests and their test statistics see e.g. Ollech and Webel (2020).

series. For each of the resulting vintages, we then repeat the process of aggregation and seasonal adjustment using the same sets of specifications that we used for the complete time series. Finally, we compare the adjusted time series aggregated to quarters and compute the revision metrics: the mean absolute revision (MAR) and the root mean squared revision (RMSR), both in %. We compute both of these measures on a period-to-period basis as well as based on the first and the last vintage.

Our results reveal a large difference between the 2 real-world examples. For the electricity demand series, the largest mean absolute revision between the first and the last vintage amounts to a mere 0.53% (SEATS, daily). Whereas, for the pedestrian activity time series, the smallest mean absolute revision is equal to 1.80% and thus more than three times larger. For the electricity series, revisions are lowest when adjusting the monthly aggregates. Revisions of the pedestrian traffic series are particularly high when adjusting at a quarterly frequency. These findings illustrate that in terms of the propensity for revisions, there may not be a single frequency at which seasonal adjustment is generally best.

We synthesise the preceding diagnostics by computing **mean ranks** that combine all metrics into a single summary measure. For electricity consumption, the preferred aggregation level depends on our choice of adjustment method used. Across methods, using SEATS at the monthly aggregation and STL at the daily resolution yield the lowest mean ranks. When restricting attention to X-11-based procedures, the daily aggregation performs best. Looking at Figures 3 and 4, we see that the choice of a temporal aggregation level does only have a modest impact on the adjustment result; practical considerations (e.g., timeliness or computational cost) may therefore guide the final choice.

By contrast, the results for the pedestrian activity depend strongly on the chosen aggregation level. Quarterly aggregation produces elevated end-of-sample values, driven by rapidly increasing amplitudes of the seasonal and trend components in the RegARIMA-based forecasts (not shown). The mean-rank summary accordingly disfavors the quarterly aggregation. Across methods, the monthly aggregation performs best, particularly with X-13-ARIMA. If a higher frequency is required, STL at the daily aggregation level is a suitable candidate.

Lastly, we study the **relationship to target series**, i.e how well our 2 higher-frequency time series are suited to nowcast traditional economic indicators. For the electricity time series this is the output in the German production sector. For the pedestrian activity time series we consider retail turnover in sales rooms in Germany. Both of these are originally monthly time series, of which we compute quarterly averages. Then, we use 2 measures, to see how well our adjusted time series are able to match their target time series. First, we compute the scaled congruence between the adjusted time series and the target time series, i.e. the share of observations for which the growth rate of the adjusted time series has the same sign as the growth rate of the target time series, scaled by a naïve sign forecast (see Section 3). If this scaled congruence ratio exceeds a value of 1, the adjusted time series does provide some additional informational value on the future trajectory of the target time series. Bear in mind, however, that the sign congruence measure only considers the direction in which the adjusted time series and its target move but not by how much. To also capture this quantitative aspect, we compute a second measure, which is the simple correlation coefficient between the growth rates of the adjusted time series and the growth rates of its target.

We find that, for both proxies, the scaled sign congruence takes values both above and below 1. Nevertheless, some adjusted series may still be useful for nowcasting their respective target series. In line with this, we observe contemporaneous correlation coefficients between 0.55 and 0.63 for the electricity consumption. The relationship between proxy and target in terms of correlation of growth rates is highest at the monthly aggregation level, despite the scaled sign congruence being below 1 in those cases. Taken together, we recommend using STL at the daily aggregation level for the electricity consumption.

As with other diagnostics discussed, the results for the quarterly pedestrian series using X-13-ARIMA and TRAMO-SEATS are discouraging, with correlations of 0.16 and 0.13, respectively.[16] For other candidates, correlations are substantially higher, consistently exceeding 0.7. At the daily and monthly aggregation levels, X-11-based adjustments outperform their SEATS-based counterparts in terms of the scaled sign congruence and at the weekly level, the reverse holds. We would therefore advise to use daily or monthly X-11 or weekly SEATS, if the relationship to a target is to be maximised.

[16] To gauge typical proxy-target relationships, we computed several contemporaneous correlations between individual WAI series and corresponding German target series. Using quarterly growth rates, the correlations range from 0.43 (credit-card payments vs. retail turnover in sales rooms) to 0.97 (global flight numbers vs. exports). Excluding the pandemic period (2020-Q2 to 2022-Q1) markedly weakens these relationships: the highest correlation falls to 0.65 (credit-card payments vs. turnover in accommodation and food-service activities) and several correlations become statistically insignificant. These results echo the evidence in the forecasting literature: the informational edge of higher-frequency indicators becomes most valuable in periods of abrupt economic stress, such as the COVID-19 shock (see, for example, Jardet and Meunier, 2022)

Table 4: German electricity consumption

	Daily			Weekly		Monthly		Quarterly	
	X-11	SEATS	STL	X-11	SEATS	X-13	SEATS	X-13	SEATS
Relationship to target series									
Scaled congruence	1.10	1.10	1.10	0.96	0.96	0.96	0.96	1.19	1.05
Correlation	0.57	0.55	0.59	0.58	0.55	0.63	0.62	0.59	0.60
Remaining calendar effects									
F-test / t-test	0.833	0.524	0.945	0.836	0.612	0.724	0.904	0.814	0.848
Starting-day-of-the-period plot	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)
Remaining seasonality									
QS test	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Friedman test	0.948	0.696	0.868	0.971	0.840	0.241	0.696	0.218	0.266
Visual inspection									
Adjusted series	✓	✓	✓	✓	✓	✓	✓	✗	✗
Spectrum	✓	(✓)	✓	✓	✓	(✗)	(✗)	(✗)	(✗)
Revision metrics (in %)									
MAR (period to period)	0.05	0.07	0.05	0.09	0.06	0.05	0.05	0.08	0.07
RMSR (period to period)	0.07	0.09	0.06	0.13	0.11	0.06	0.06	0.10	0.09
MAR (first to last)	0.30	0.53	0.35	0.32	0.21	0.19	0.19	0.38	0.26
RMSR (first to last)	0.34	0.56	0.38	0.42	0.29	0.24	0.24	0.49	0.33
Ranking									
Mean rank	2.62	4.38	2.31	3.19	3.31	3.12	2.12	5.06	3.75

Notes:

[1] Scaled congruence: fraction of obs. with same sign in growth rates, scaled by naive sign forecast.

[2] For the statistical tests the table reports p-values.

[3] Visual inspections: Check marks (in parentheses) indicate (mostly) unproblematic results.

[4] The ranks are calculated, so that the best entry is assigned rank 1.

Table 5: German pedestrian frequency

	Daily			Weekly		Monthly		Quarterly	
	X-11	SEATS	STL	X-11	SEATS	X-13	SEATS	X-13	SEATS
Relationship to target series									
Scaled congruence	1.07	0.94	0.94	1.00	1.07	1.07	0.93	0.87	1.00
Correlation	0.72	0.70	0.72	0.72	0.73	0.73	0.74	0.16	0.13
Remaining calendar effects									
F-test / t-test	0.013	0.003	0.000	0.601	0.000	0.000	0.000	0.000	0.000
Starting-day-of-the-period plot	(✗)	(✗)	(✗)	(✗)	(✗)	(✗)	(✗)	(✗)	(✓)
Remaining seasonality									
QS test	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Friedman test	0.086	0.241	0.334	0.158	0.158	0.334	0.334	0.978	0.362
Visual inspection									
Adjusted series	✓	✓	✓	✓	✓	✓	✓	✓	✓
Spectrum	(✗)	(✗)	(✗)	(✗)	(✗)	(✗)	(✗)	✓	(✗)
Revision metrics (in %)									
MAR (period to period)	0.66	0.68	0.74	0.86	0.56	0.44	0.43	1.85	1.64
RMSR (period to period)	0.92	0.80	0.99	1.50	0.67	0.51	0.52	2.21	1.99
MAR (first to last)	4.15	4.15	1.97	3.12	2.15	1.80	2.10	8.91	9.20
RMSR (first to last)	5.19	5.07	2.44	5.75	2.60	2.29	2.65	10.04	9.77
Ranking									
Mean rank	3.75	3.38	2.88	3.62	3.31	2.56	2.75	4.06	4.31

Notes:

[1] Scaled congruence: fraction of obs. with same sign in growth rates, scaled by naive sign forecast.

[2] For the statistical tests the table reports p-values.

[3] Visual inspections: Check marks (in parentheses) indicate (mostly) unproblematic results.

[4] The ranks are calculated, so that the best entry is assigned rank 1.

5 Summary

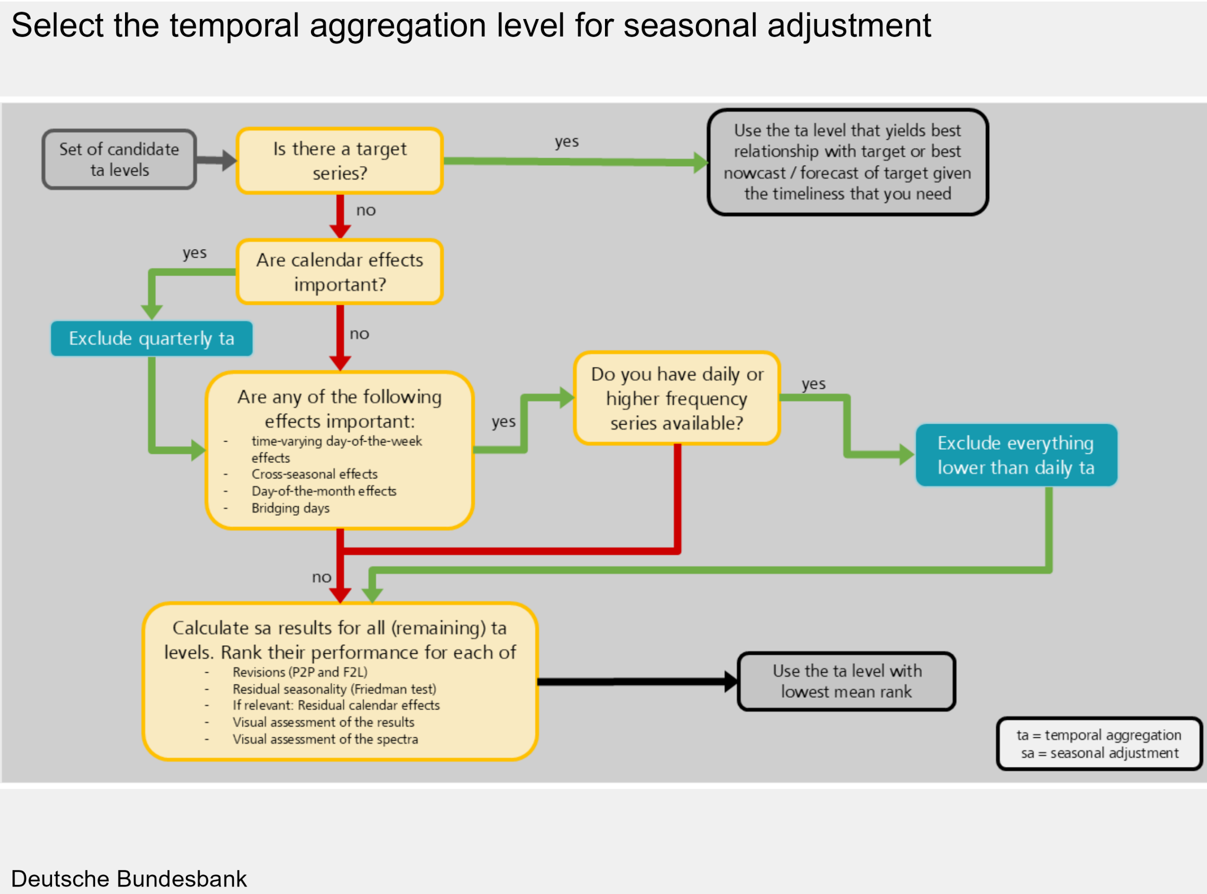

Figure 9: Decision tree for selection of temporal aggregation level

We present a stylised decision tree in Figure 9 to guide the choice of temporal aggregation for seasonal adjustment in practice. First, define the use case and filter out aggregation levels that cannot serve it. If, say, we want to use the series as a higher-frequency proxy for a quarterly target, quarterly aggregation offers little advantage in terms of timeliness. So, we would only be interested in results based on the quarterly aggregation level if the target series is available with a sizeable delay.

Then, if the sole objective is to proxy a given target, optimise the proxy-target relationship: select the aggregation level that maximises alignment with the target. When the proxy enters a forecasting model, out-of-sample prediction accuracy metrics can be used directly to select the optimal temporal resolution.

If there is no target series, or the relationship to a target is only one of the objectives, we can consider the properties of the series at hand. When calendar effects are very important, avoid a quarterly aggregation. As we have seen, the identification of (residual) calendar effects at that frequency is difficult and often unreliable. If additional effects, such as cross-seasonal effects are important, consider only a daily or higher frequency. Such effects can typically be estimated reliably only at higher frequencies.

For the remaining candidate temporal resolutions, evaluate the diagnostics and metrics discussed in this paper. Summarise them via mean ranks and choose the option with the lowest mean rank.

Future work should explore principled weighting schemes for the diagnostics-tailored to the intended use, sample length, and timeliness requirements to place greater emphasis on the most relevant criteria.

References

Arshad, S. and R. C. Beyer (2023). Tracking economic fluctuations with electricity consumption in bangladesh. *Energy Economics 123*, 1–14.

Aston, J. A. and S. J. Koopman (2006). A non-gaussian generalization of the airline model for robust seasonal adjustment. *Journal of Forecasting 25*(5), 325–349.

Athanasopoulos, G., R. J. Hyndman, N. Kourentzes, and F. Petropoulos (2017). Forecasting with temporal hierarchies. *European Journal of Operational Research 262*(1), 60–74.

Box, G. E. P. and G. Jenkins (1970). *Time Series Analysis: Forecasting and Control.* San Francisco: Holden-Day.

Brewer, K. R. W. (1973). Some consequences of temporal aggregation and systematic sampling for arma and armax models. *Journal of Econometrics 1*(2), 133–154.

Bricongne, J.-C., B. Meunier, and S. Pouget (2023). Web-scraping housing prices in real-time: The covid-19 crisis in the uk. *Journal of Housing Economics 59*, 101906.

Chen, C. and L.-M. Liu (1993). Joint Estimation of Model Parameters and Outlier Effects in Time Series. *Journal of the American Statistical Association 88*(421), 284–297.

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

Cleveland, W. S. and S. J. Devlin (1980). Calendar effects in monthly time series: detection by spectrum analysis and graphical methods. *Journal of the American Statistical Association 75*(371), 487–496.

Cuevas, Á. and E. M. Quilis (2023). Seasonal adjustment methods for daily time series. a comparison by a monte carlo experiment.

Delle Monache, D., S. Emiliozzi, and A. Nobili (2020). Tracking Economic Growth during the Covid-19: a Weekly Indicator for Italy. *Bank of Italy, Mimeo*.

Deutsche Bundesbank (2012, December). Calendar effects on economic activity. *Monthly Report 64*(12), 51–60.

Eraslan, S. and T. Götz (2020). Weekly Activity Index for the German Economy. Available under URL https://www.bundesbank.de/wai.

Eurostat (2018). ESS Guidelines on Temporal Disaggregation, Benchmarking and Reconciliation. Publications Office of the European Union, Luxembourg, ISBN 978-92-79-98006-0.

Eurostat (2024). ESS Guidelines on Seasonal Adjustment. 2024 edition. Publications Office of the European Union, Luxembourg. ISBN 978-92-79-45176-8.

Fenz, G. and H. Stix (2021). Monitoring the Economy in Real Time with the Weekly OeNB GDP Indicator: Background, Experience and Outlook. *Monetary Policy & the Economy*.

Fischer, B. and C. Planas (2000). Large scale fitting of regression models with arima errors. *Journal of Official Statistics 16*(2), 173.

Gómez, V. and A. Maravall (2001). Automatic Modeling Methods for Univariate Series. In D. Peña, G. C. Tiao, and R. S. Tsay (Eds.), *A Course in Time Series Analysis*, pp. 171–201. New York: Wiley.

Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: journal of the Econometric Society*, 424–438.

Haller, V., S. Daniel, and B. Bellone (2025). STAHL: Seasonal, trend, and holiday decomposition with loess. *Journal of Official Statistics*, 0282423X251344219.

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

Harvey, A. C. and N. Shephard (1993). *Structural time series models*, Volume 11, pp. 261–302.

Hyndman, R. J. and A. B. Koehler (2006). Another look at measures of forecast accuracy. *International journal of forecasting 22*(4), 679–688.

Jardet, C. and B. Meunier (2022). Nowcasting world gdp growth with high-frequency data. *Journal of Forecasting 41*(6), 1181–1200.

Kendall, Maurice; Stuart, A. and J. K. Ord (1983). *Kendall's advanced theory of statistics.* (4 ed.), Volume 3. Griffin: London.

Kourentzes, N., B. Rostami-Tabar, and D. K. Barrow (2017). Demand forecasting by temporal aggregation: Using optimal or multiple aggregation levels? *Journal of Business Research 78*, 1–9.

Lehmann, R. and S. Möhrle (2024). Forecasting regional industrial production with novel high-frequency electricity consumption data. *Journal of Forecasting 43*(6), 1918–1935.

Lewis, D. J., K. Mertens, J. H. Stock, and M. Trivedi (2022). Measuring real activity using a weekly economic index. *Journal of Applied Econometrics 37*(4), 667–687.

Lourenço, N. and A. Rua (2021). The Daily Economic Indicator: Tracking Economic Activity Daily During the Lockdown. *Economic Modelling 100*, 1–10.

McElroy, T. and A. Roy (2022). A review of seasonal adjustment diagnostics. *International Statistical Review 90*(2), 259–284.

Moulton, B. R. and B. D. Cowan (2016). Residual seasonality in gdp and gdi.

Nikolopoulos, K., A. A. Syntetos, J. E. Boylan, F. Petropoulos, and V. Assimakopoulos (2011). An aggregate–disaggregate intermittent demand approach (adida) to forecasting: an empirical proposition and analysis. *Journal of the Operational Research Society 62*(3), 544–554.

Ollech, D. (2021). Seasonal adjustment of daily time series. *Journal of Time Series Econometrics 13*(2).

Ollech, D. (2022). Economic analysis using higher-frequency time series: Challenges for seasonal adjustment. *Empirical Economics*, 1–24.

Ollech, D. and K. Webel (2020). A random forest-based approach to identifying the most informative seasonality tests. Discussion Paper No 55/2020, Deutsche Bundesbank.

Ollech, D. and K. Webel (2022). A random forest-based approach to combining and ranking seasonality tests. *Journal of Econometric Methods*.

Pesaran, M. H. and A. Timmermann (1992). A simple nonparametric test of predictive performance. *Journal of Business & Economic Statistics 10*(4), 461–465.

Rossana, R. J. and J. J. Seater (1995). Temporal aggregation and economic time series. *Journal of Business & Economic Statistics 13*(4), 441–451.

Silvestrini, A. and D. Veredas (2008). Temporal aggregation of univariate and multivariate time series models: A survey. *Journal of Economic Surveys 22*(3), 458–497.

Soukup, R. J. and D. F. Findley (1999). On the Spectrum Diagnostics Used by X-12-ARIMA to Indicate the Presence of Trading Day Effects after Modeling or Adjustment. In *Proceedings of the American Statistical Association, Business and Economic Statistics Section*, pp. 144–149.

Svetunkov, I. (2017). Statistical Models Underlying Functions of 'smooth' Package for R. Working paper, Lancaster University Management School.

Tiao, G. C. (1972). Asymptotic behaviour of temporal aggregates of time series. *Biometrika 59*(3), 525–531.

Wegmüller, P., C. Glocker, and V. Guggia (2021). Weekly Economic Activity: Measurement and Informational Content. Grundlagen für die Wirtschaftspolitik 17, Staatssekretariat für Wirtschaft SECO.

Wei, W. W. (1978). Some consequences of temporal aggregation in seasonal time series models. In *Seasonal analysis of economic time series*, pp. 433–448. NBER.

Woloszko, N. (2024). Nowcasting with panels and alternative data: The oecd weekly tracker. *International Journal of Forecasting 40*(4), 1302–1335.

A Appendix

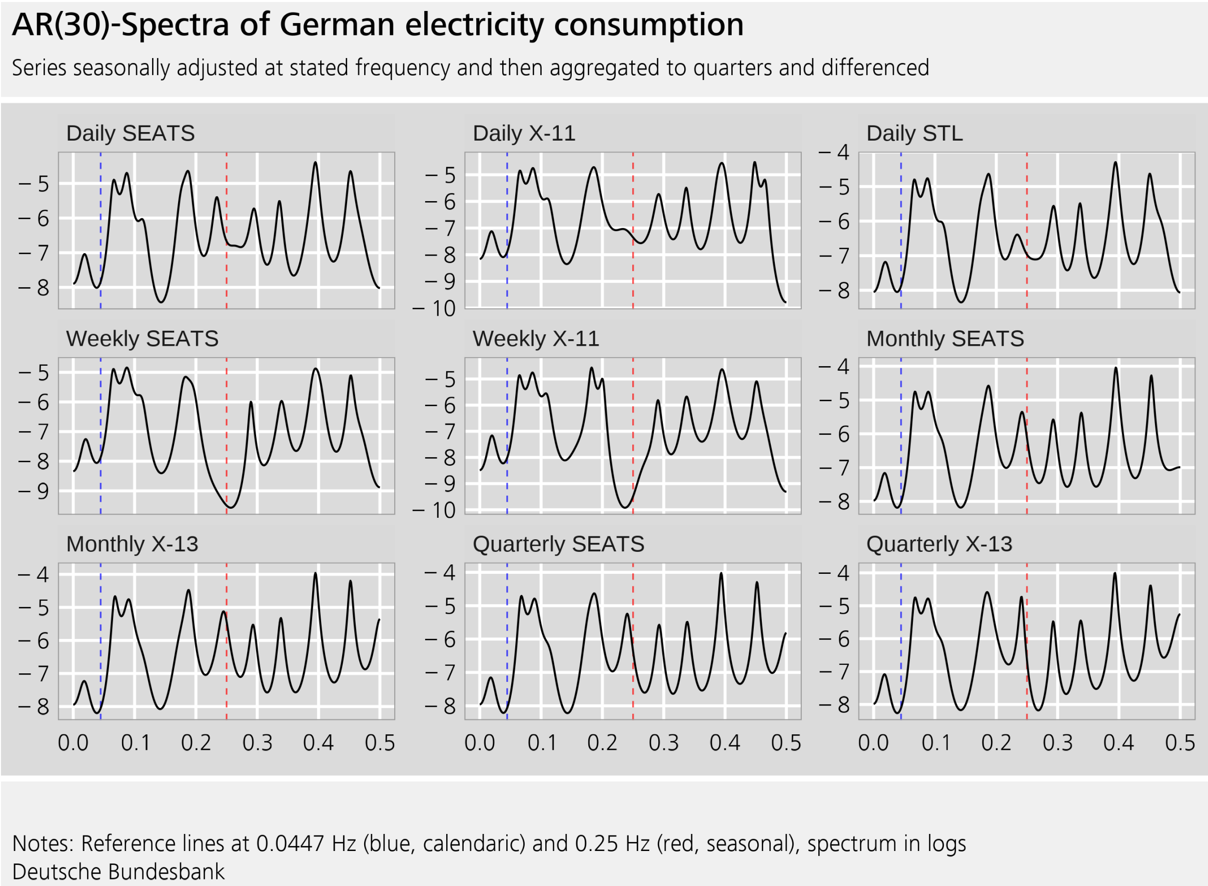

Figure 10: Spectra for German electricity consumption

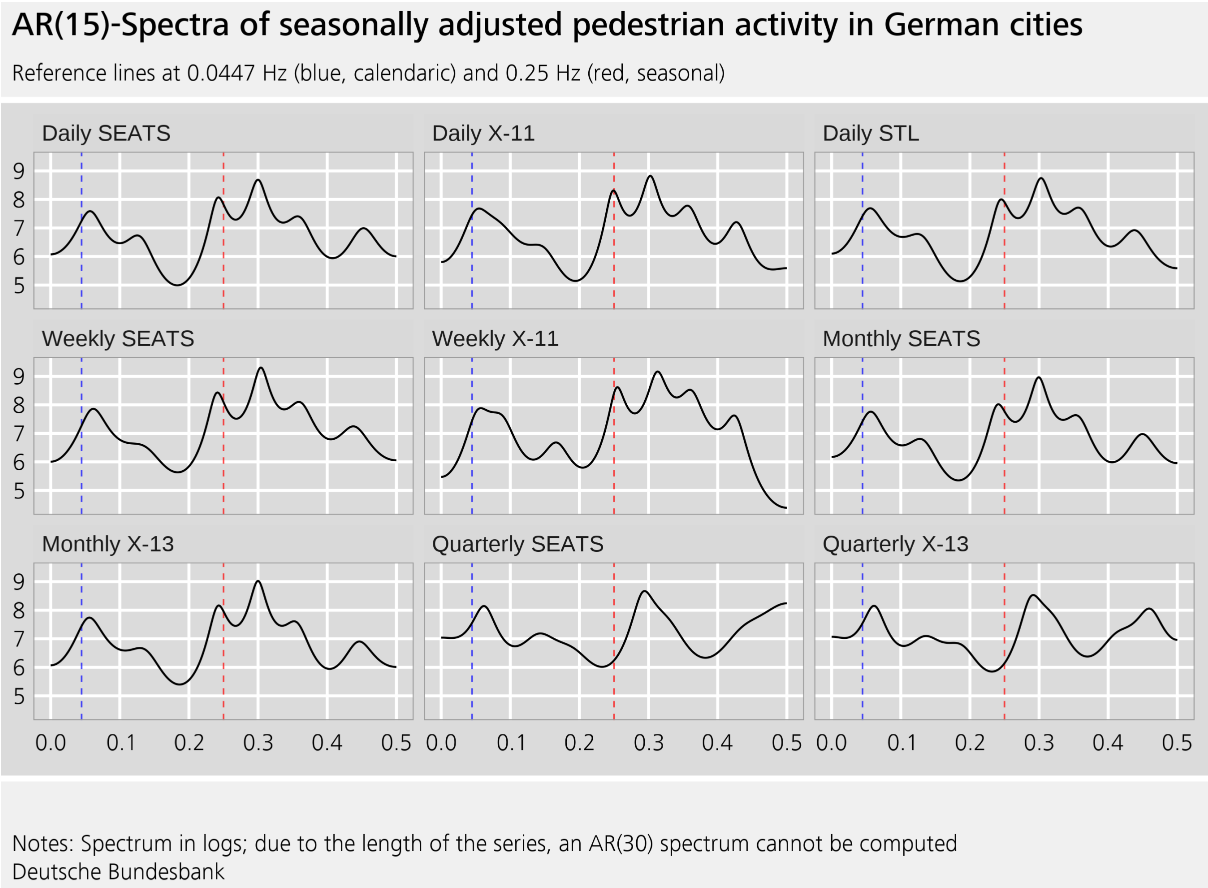

Figure 11: Spectra for pedestrian activity in German cities

Table 6: Pre-processing of electricity

	Daily	Weekly	Monthly	Quarterly
Observations	3834	547	126	42
Decomposition	Mult.	Mult.	Mult.	Mult.
ARIMA	(0,1,1)(0,1,1)	(0,1,1)(0,1,1)	(0,1,1)(0,1,1)	(0,1,1)(0,1,1)
Crit.val.	8	8	6	6
Outliers	WO (2018-04-30) AO (2024-05-01) AO (2025-05-01)	-	LS (4-2020) LS (8-2020) LS (4-2025)	AO (II-2020) LS (II-2025)
Calendar	Jan1Sun HolyThursday GoodFriday HolySaturday EasterSunday EasterMonday EasterMondayAft1Day AscensionBef1Day Ascension AscensionAft1Day CorpusChristi CorpusChristiAft1Day May1Bridge PentecostBef1Day PentecostMonday PentecostAft1Day Oct3Bridge Nov1Bridge Dec25Sat Dec24Sun Dec25Sun Dec26Sun PostXmasSun10d DstSpring DstAutumn ReformationDayBridge NationalHolWeekday NationalHolSat XmasPeriodWeekday CarnivalMonday	CarnivalMonday GoodFriday EasterMonday Pentecost CorpusChristi	WorkingDay	WorkingDay

Notes:
[1] Due to technical reasons, we conducted a log-transformation for the daily SEATS model outside the adjustment routine itself.

Table 7: Pre-processing of pedestrians

	Daily	Weekly	Monthly	Quarterly
Observations	2557	365	84	28
Decomposition	Mult.	Mult.	Mult.	Mult.
ARIMA	(0,1,1)(0,1,1)	(0,1,1)(0,1,1)	(0,1,1)(0,1,1)	(0,1,1)(0,1,1)
Crit.val.	8	8	6	6
Outliers	LS (2020-03-18) LS (2020-04-19) LS (2020-05-02) LS (2020-12-16) LS (2021-05-20)	LS (2020-03-22) LS (2020-03-29) LS (2020-04-26) LS (2020-12-27)	LS (3-2020) LS (4-2020) LS (5-2020) LS (12-2020) LS (1-2021) LS (6-2021) LS (12-2021) LS (1-2022)	AO (II-2020) AO (IV-2020) AO (IV-2021)
Calendar	Jan1Sun HolyThursday GoodFriday EasterMonday EasterMondayAft1Day May1Bridge PentecostMonday PentecostAft1Day Ascension AscensionAft1Day CorpusChristi CorpusChristiAft1Day Oct3Bridge Nov1Bridge Dec25Sat PreXmasSun3d Dec25Sun Dec26Sun Dec31Sun ThanksgivingFriday XmasPeriodWeekday LabourDayWeekdays LabourDaySaturday GermanUnityWeekdays GermanUnitySaturday AdventSun	GoodFriday PentecostMonday	Monday Tuesday Wednesday Thursday Friday Saturday	Monday Tuesday Wednesday Thursday Friday Saturday

Notes:
[1] Due to technical reasons, we conducted a log-transformation for the daily SEATS model outside the adjustment routine itself.

Table 8: Filter selection

	Electricity consumption		Pedestrian activity	
	Seasonal	Trend	Seasonal	Trend
Daily				
X-11	3x15 ($\tau = 7$)	13 ($\tau = 7$)	3x15 ($\tau = 7$)	13 ($\tau = 7$)
	3x9 ($\tau = 365$)	367 ($\tau = 365$)	3x3 ($\tau = 365$)	367 ($\tau = 365$)
SEATS	-	-	-	-
STL	25 ($\tau = 7$)	13 ($\tau = 7$)	31 ($\tau = 7$)	13 ($\tau = 7$)
	15 ($\tau = 365$)	609 ($\tau = 365$)	7 ($\tau = 365$)	697 ($\tau = 365$)
Weekly				
X-11	3x5	55	3x3	55
SEATS	-	-	-	-
Monthly				
X-11	3x9	13-Henderson	3x9	13-Henderson
SEATS	-	-	-	-
Quarterly				
X-11	3x9	13-Henderson	3x9	13-Henderson
SEATS	-	-	-	-

Notes:
[1] For the X-11-type models ”Seasonal” refers to the seasonal movering average filter. For the STL-model it refers to number of observations included in the local regressions calculated to obtain the seasonal component.

Table 9: Correlations between growth rates for seasonal adjustment results at different levels of temporal aggregation

		Daily			Weekly		Monthly		Quarterly		
	Original	X-11	SEATS	STL	X-11	SEATS	X-11	SEATS	X-11	SEATS	Target
Original	1.000	0.602	0.530	0.597	0.582	0.540	0.584	0.604	0.582	0.604	0.621
Daily											
X-11	0.602	1.000	0.889	0.980	0.963	0.884	0.974	0.967	0.965	0.946	0.943
SEATS	0.530	0.889	1.000	0.881	0.907	0.922	0.907	0.889	0.906	0.883	0.846
STL	0.597	0.980	0.881	1.000	0.950	0.875	0.969	0.966	0.964	0.946	0.948
Weekly											
X-11	0.582	0.963	0.907	0.950	1.000	0.920	0.974	0.954	0.967	0.937	0.913
SEATS	0.540	0.884	0.922	0.875	0.920	1.000	0.905	0.891	0.905	0.883	0.841
Monthly											
X-11	0.584	0.974	0.907	0.969	0.974	0.905	1.000	0.967	0.982	0.947	0.931
SEATS	0.604	0.967	0.889	0.966	0.954	0.891	0.967	1.000	0.959	0.959	0.947
Quarterly											
X-11	0.582	0.965	0.906	0.964	0.967	0.905	0.982	0.959	1.000	0.953	0.924
SEATS	0.604	0.946	0.883	0.946	0.937	0.883	0.947	0.959	0.953	1.000	0.929
Target	0.621	0.943	0.846	0.948	0.913	0.841	0.931	0.947	0.924	0.929	1.000

Note:
The Pearson-Bravais-type correlations are computed on growth rates of seasonally adjusted series. When necessary, the data are first aggregated to quarterly frequency.

Table 10: Congruence analysis for seasonal adjustment results at different levels of temporal aggregation

		Daily			Weekly		Monthly		Quarterly		
	Original	X-11	SEATS	STL	X-11	SEATS	X-11	SEATS	X-11	SEATS	Target
Original	100.0	70.4	68.5	70.1	69.4	68.6	70.0	71.4	70.2	73.5	71.1
Daily											
X-11	70.4	100.0	87.3	94.4	93.0	86.5	94.2	93.5	92.5	90.8	90.1
SEATS	68.5	87.3	100.0	86.1	88.7	91.2	88.8	86.9	88.8	86.3	83.7
STL	70.1	94.4	86.1	100.0	91.2	85.5	93.2	93.1	92.4	90.5	89.8
Weekly											
X-11	69.4	93.0	88.7	91.2	100.0	88.6	93.7	91.5	92.6	89.3	87.3
SEATS	68.6	86.5	91.2	85.5	88.6	100.0	87.8	86.7	88.2	86.0	82.9
Monthly											
X-11	70.0	94.2	88.8	93.2	93.7	87.8	100.0	93.0	94.7	90.6	88.9
SEATS	71.4	93.5	86.9	93.1	91.5	86.7	93.0	100.0	91.8	93.4	90.3
Quarterly											
X-11	70.2	92.5	88.8	92.4	92.6	88.2	94.7	91.8	100.0	90.6	88.2
SEATS	73.5	90.8	86.3	90.5	89.3	86.0	90.6	93.4	90.6	100.0	88.5
Target	71.1	90.1	83.7	89.8	87.3	82.9	88.9	90.3	88.2	88.5	100.0

Note:
Congruence is the proportion of observations – aggregated to quarterly frequency when necessary – in which the growth rates of the two series have the same sign.

Table 11: Mean absolute differences for seasonal adjustment results at different levels of temporal aggregation

		Daily			Weekly		Monthly		Quarterly		
	Original	X-11	SEATS	STL	X-11	SEATS	X-11	SEATS	X-11	SEATS	Target
Original	0.0	9.9	10.3	10.0	10.1	10.4	10.0	9.8	10.0	9.5	9.7
Daily											
X-11	9.9	0.0	2.4	0.9	1.8	2.9	1.1	1.2	1.2	1.6	1.8
SEATS	10.3	2.4	0.0	2.4	2.4	2.0	2.1	2.4	2.0	2.4	3.0
STL	10.0	0.9	2.4	0.0	2.0	3.0	1.2	1.3	1.3	1.6	1.8
Weekly											
X-11	10.1	1.8	2.4	2.0	0.0	2.2	1.6	2.0	1.7	2.2	2.5
SEATS	10.4	2.9	2.0	3.0	2.2	0.0	2.6	2.9	2.6	2.9	3.4
Monthly											
X-11	10.0	1.1	2.1	1.2	1.6	2.6	0.0	1.2	0.8	1.5	2.0
SEATS	9.8	1.2	2.4	1.3	2.0	2.9	1.2	0.0	1.4	1.2	1.8
Quarterly											
X-11	10.0	1.2	2.0	1.3	1.7	2.6	0.8	1.4	0.0	1.5	2.1
SEATS	9.5	1.6	2.4	1.6	2.2	2.9	1.5	1.2	1.5	0.0	2.0
Target	9.7	1.8	3.0	1.8	2.5	3.4	2.0	1.8	2.1	2.0	0.0

Note:
The mean absolute difference is computed for the series aggregated to a quarterly level after seasonal adjustment.

Table 12: Mean absolute differences of growth rates for seasonal adjustment results at different levels of temporal aggregation

		Daily			Weekly		Monthly		Quarterly		
	Original	X-11	SEATS	STL	X-11	SEATS	X-11	SEATS	X-11	SEATS	Target
Original	0.0	8.9	9.3	8.9	9.0	9.3	9.0	8.8	9.0	8.5	8.7
Daily											
X-11	8.9	0.0	2.1	0.9	1.2	2.3	1.0	1.1	1.2	1.4	1.7
SEATS	9.3	2.1	0.0	2.1	1.7	1.3	1.8	2.0	1.8	2.1	2.6
STL	8.9	0.9	2.1	0.0	1.4	2.3	1.1	1.1	1.2	1.4	1.6
Weekly											
X-11	9.0	1.2	1.7	1.4	0.0	1.8	1.0	1.4	1.1	1.6	2.0
SEATS	9.3	2.3	1.3	2.3	1.8	0.0	2.0	2.2	1.9	2.3	2.8
Monthly											
X-11	9.0	1.0	1.8	1.1	1.0	2.0	0.0	1.1	0.8	1.4	1.8
SEATS	8.8	1.1	2.0	1.1	1.4	2.2	1.1	0.0	1.3	1.1	1.6
Quarterly											
X-11	9.0	1.2	1.8	1.2	1.1	1.9	0.8	1.3	0.0	1.4	1.9
SEATS	8.5	1.4	2.1	1.4	1.6	2.3	1.4	1.1	1.4	0.0	1.8
Target	8.7	1.7	2.6	1.6	2.0	2.8	1.8	1.6	1.9	1.8	0.0

Note:
The mean absolute difference is computed for the growth rates of the series aggregated to a quarterly level after seasonal adjustment.