Asymmetric Linear Filters for Seasonal Adjustment, Applications to the Covid-19 Impact in Monthly Time Series

Alain Quartier-la-Tente[1][*]
Dominique Ladiray[2]

Abstract
The COVID-19 crisis highlights that business cycle analysis, and in particular the early detection of turning points, is a major topic in the analysis of economic outlook. Moving averages, or linear filters, are ubiquitous in business cycle extraction and seasonal adjustment methods. For example, the X-12ARIMA seasonal adjustment method uses Henderson moving averages and composite moving averages to estimate the main components of a time series, while TRAMO-SEATS uses Wiener-Kolmogorov filters. Symmetric filters are applied to the center of the series, but when it comes to estimate the most recent points, all of these methods must rely on asymmetric filters which introduce phase-shifts and delays in the detection of turning points. Construction of good asymmetric filters, in terms of fidelity, revisions, smoothness and timeliness, is still an open topic. This presentation describes and compares the recent approaches on trend-cycle extraction and asymmetric filters, and applies them to the COVID-19 economic crisis. All the methods are implemented in the ℝ package rjdfilters (https://github.com/palatej/rjdfilters) and the results can be easily reproduced.

Seasonal adjustment, asymmetric filters, turning points, Covid-19.

[1]LEMNA

[2]Independant

Corresponding author: alain.quartier@yahoo.fr,

Contents

1. Introduction

Moving averages, or linear filters, are ubiquitous in business cycle extraction and seasonal adjustment methods. For example, the X-12ARIMA seasonal adjustment method uses Henderson moving averages and composite moving averages to estimate the main components of a time series, while TRAMO-SEATS uses Wiener-Kolmogorov filters. Symmetric filters are applied to the center of the series, but when it comes to estimate the most recent points (i.e.: for real time estimates), all of these methods must rely on asymmetric filters. For example, even if X-12ARIMA or TRAMO-SEATS apply symmetrical averages to the forecasts obtained from an ARIMA model of the series, in reality it consists in applying asymmetric filters at the end of the series, because the predicted values are linear combinations of past values.

If the classic asymmetric moving averages have good properties regarding the future revisions induced by the process[1], they create phase shifts that impact the real-time estimation of turning points, introducing time delay in the detection.

This presentation is part of an ongoing research on the description and comparison of recent approaches around trend-cycle extraction and asymmetric filters: local polynomial filters of Proietti and Luati (2008), filters based on Reproducing Kernel Hilbert Space detailed by Dagum and Bianconcini (2008) and methods based on an optimization of filters properties of Grun-Rehomme et al. (2018) or Wildi and McElroy (2019)

---

[1]See for example Pierce (1980).

2. From seasonal adjustment to trend-cycle estimation

To perform seasonal adjustment, most of the algorithm decompose the data in several unknown components: trend-cycle, seasonal and irregular component. In X-12ARIMA and TRAMO-SEATS (the most popular seasonal adjustment methods), prior to the decomposition, the initial series is pre-adjustment from deterministic effects (outliers, calendar effects). Thus, the estimation of trend-cycle component is technically linked to the seasonal component.

Moreover, since the trend-cycle extraction methods are applied to seasonally adjusted data, both problems cannot be separated. In this paper, we study trend-cycle extraction methods that can be implemented in the seasonal adjustment method X-12ARIMA. This is why almost all the methods described are implemented in the core libraries of JDemetra+ [2], the seasonal adjustment software recommended by Eurostat. An ⓡ interface is implemented in the package `rjdfilters`[3].

In the literature, different approaches were considered for trend-cycle extraction[4]. Among the most popular, we can cite the Model-Based Approach and the non-parametric extraction methods:

- The Model-Based Approach assumes the specification of a stochastic time series model for the trend (ARIMA model, state space model, etc.) and the estimates is obtained by minimizing a penalty function, generally the mean squared error. This is for example the case of the Kalman filter, the Wiener-Kolmogorov filter (used in TRAMO-SEATS) and the Direct Filter Approach of Wildi and McElroy (2019) (section 4.4).
- The non-parametric extraction methods do not assume that the structure of a model is fixed and can be easily applied to any time series. This is for example the case of the Henderson and Musgrave filters (used in X-12ARIMA). Classical methods can be seen as local polynomial regression, approach generalized by Proietti and Luati (2008) (section 4.1).
  Non-parametric estimators can also be reproduced by exploiting Reproducing Kernel Hilbert Space (RKHS) methodology, like it is done by Dagum and Bianconcini (2008) (section 4.2).

Grun-Rehomme et al. (2018) (section 4.3) drawn a general unifying approach that allows a theoretical link between Musgrave non-parametric filter and the Direct Filter Approach. They also proposed a global procedure to build asymmetric moving averages, that allows to minimize the phase shift effects and which is described in this report.

---

[2]See https://github.com/jdemetra/jdemetra-core.
[3]Available at https://github.com/palatej/rjdfilters.
[4]See for example Alexandrov et al. (2012) for a review of some modern approaches to trend extraction.

3. Basic properties of moving average

In this section we describe some definitions and properties of moving average and linear filters (see for example Ladiray (2018) for more details).

Let p et f two integers, a moving average M_θ or M is defined by a set of coefficients $\theta = (\theta_{-p}, \ldots, \theta_f)'$ such as for all time series X_t:

$$M_\theta(X_t) = \sum_{k=-p}^{+f} \theta_k X_{t+k}$$

- $p + f + 1$ is called the *moving average order*.
- When $p = f$ the moving average is said to be *centered*. If we also have $\forall k : \theta_{-k} = \theta_k$, the moving average M_θ is said to be *symmetric*. In this case, the quantity $h = p = f$ is called the *bandwidth*.

Let $X_t = e^{-i\omega t}$. The result of the moving average M_θ in X_t is:

$$Y_t = M_\theta X_t = \sum_{k=-p}^{+f} \theta_k e^{-i\omega(t+k)} = \left(\sum_{k=-p}^{+f} \theta_k e^{-i\omega k} \right) \cdot X_t.$$

The function $\Gamma_\theta(\omega) = \sum_{k=-p}^{+f} \theta_k e^{-i\omega k}$ is called the *transfer function* or *frequency response function*. It can be rewritten as:

$$\Gamma_\theta(\omega) = G_\theta(\omega) e^{-i\Phi_\theta(\omega)}$$

where $G_\theta(\omega) = |\Gamma_\theta(\omega)|$ is the *gain* or *amplitude* function and $\Phi_\theta(\omega)$ is the *phase shift* or *time shift* function. For all symmetric moving average we have $\Phi_\theta(\omega) \equiv 0$ (mod π).

To sum up, applying a moving average to a harmonic times series affects it in in two different ways:
- by multiplying it by an amplitude coefficient $G_\theta(\omega)$;
- by "shifting" it in time by $\Phi_\theta(\omega)/\omega$, which directly affects the detection of turning points.

4. Description of the different approaches

4.1 Local polynomial filters
In this subsection we detail the filters that arise from fitting a local polynomial to our time series, as described by Proietti and Luati (2008). Local polynomial filters encompass classical filters like Henderson and Musgrave filters.

4.1.1 Symmetric filter
We assume that our time series y_t can be decomposed as

$$y_t = \mu_t + \varepsilon_t$$

where μ_t is the signal (trend) and $\varepsilon_t \overset{i.i.d}{\sim} \mathcal{N}(0, \sigma^2)$ is the noise[5]. We assume that μ_t can be locally approximated by a polynomial of degree d of the time t between y_t and the neighboring observations $(y_{t+j})_{j \in \llbracket -h,h \rrbracket}$. Then $\mu_t \simeq m_t$ with:

$$\forall j \in \llbracket -h, h \rrbracket : y_{t+j} = m_{t+j} + \varepsilon_{t+j}, \quad m_{t+j} = \sum_{i=0}^{d} \beta_i j^i$$

---

[5]The time series is therefore seasonally adjusted.

This signal extraction problem is then equivalent to the estimation of $m_t = \beta_0$. In matrix notation we can write:

$$y = X\beta + \varepsilon, \quad \text{with:} \begin{cases} y &= \begin{pmatrix} y_{t-h} & \cdots & y_{t+h} \end{pmatrix}' \\ \beta &= \begin{pmatrix} \beta_1 & \cdots & \beta_d \end{pmatrix}' \\ \varepsilon &= \begin{pmatrix} \varepsilon_{t-h} & \cdots & \varepsilon_{t+h} \end{pmatrix}' \end{cases}$$

and:

$$X = \begin{pmatrix} 1 & -h & \cdots & (-h)^d \\ 1 & -(h-1) & \cdots & (-(h-1))^d \\ \vdots & \vdots & \cdots & \vdots \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 1 & h-1 & \cdots & (h-1)^d \\ 1 & h & \cdots & h^d \end{pmatrix}.$$

The estimation of β is done by the weighted least squares (WLS), which consists of minimizing the following objective function:

$$S(\hat{\beta}_0, \ldots, \hat{\beta}_d) = \sum_{j=-h}^{h} \kappa_j (y_{t+j} - \hat{\beta}_0 - \hat{\beta}_1 j - \cdots - \hat{\beta}_d j^d)^2$$

where κ_j is a set of weights called *kernel*. We have $\kappa_j \geq 0 : \kappa_{-j} = \kappa_j$, and with $K = diag(\kappa_{-h}, \ldots, \kappa_h)$, the estimate of β can be written as $\hat{\beta} = (X'KX)^1 X'Ky$. With $e_1 = \begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix}'$, the estimate of the trend is:

$$\hat{m}_t = e_1 \hat{\beta} = w'y = \sum_{j=-h}^{h} w_j y_{t-j} \text{ with } w = KX(X'KX)^{-1}e_1$$

The goal of the kernel function is to weight the observations according to their distance from time t. There exist many different kernels and its choice can lead to some well-known estimators: Nadaraya-Watson estimator ($d = 0$), Macaulay filter, Loess, etc. Here, we will only focus on the kernel that produces the Henderson filter, widely used for trend-cycle estimation (for example it's the filter used in the seasonal adjustment software X-12ARIMA):

$$\kappa_j = \left[1 - \frac{j^2}{(h+1)^2} \right] \left[1 - \frac{j^2}{(h+2)^2} \right] \left[1 - \frac{j^2}{(h+3)^2} \right]$$

To conclude, the estimate of the trend $\hat{m}_t$ can be obtained applying the symmetric filter w to y_t[6]. Moreover, $X'w = e_1$ so:

$$\sum_{j=-h}^{h} w_j = 1, \quad \forall r \in [\![1,d]\!] : \sum_{j=-h}^{h} j^r w_j = 0$$

Hence, the filter w preserve deterministic polynomial of order d.

---

[6] w is symmetric due to the symmetry of the kernel weights κ_j.

4.1.2 Asymmetric filter

To deal with the problem of the variance of the estimates of the real-time filters, Proietti and Luati (2008) suggest a general of asymmetric filters to make a trade-off between bias and variance. It is a generalization of Musgrave asymmetric filters (used in the seasonal adjustment algorithm X-12ARIMA, see Musgrave (1964)). They propose four classes of asymmetric filters, obtained by minimizing mean square revision error (with the symmetric filter) subject to some constraints:

1. *Direct asymmetric filters* (DAF): the asymmetric filter preserves the same polynomial constraints than the symmetric filter. This is equivalent to apply the symmetric methods to the available observations

2. *Linear-Constant* (LC): $y_t = a + \delta t + \varepsilon_t$, $\varepsilon_t \sim \mathcal{N}(0, \sigma^2)$ and the asymmetric filter only preserves constant trends. We obtain the so-called Musgrave filters when the Henderson kernel is used.

3. *Quadratic-Linear* (QL): $y_t = a + bt + \delta t^2 + \varepsilon_t$, $\varepsilon_t \sim \mathcal{N}(0, \sigma^2)$ and the asymmetric filter only preserves preserves linear trends.

4. *Cubic-Quadratic* (CQ): $y_t = a + bt + ct^2 + \delta t^3 + \varepsilon_t$, $\varepsilon_t \sim \mathcal{N}(0, \sigma^2)$ and the asymmetric filter only preserves preserves quadratic trends.

The last three filters depends on the ratio δ^2/σ^2 that can be related to the "I-C" ratio $R = \frac{\bar{I}}{\bar{C}} = \frac{\sum|I_t - I_{t-1}|}{\sum|C_t - C_{t-1}|}$ ($\delta_1/\sigma = 2/(R\sqrt{\pi})$), the ratio between the expected absolute difference of the irregular and of the trend-cycle. In the seasonal adjustment method X-12ARIMA, the I-C ratio is used to determine the bandwidth to use for the Henderson filter.

> **Summary — Local polynomial filters**
>
> **Advantages**:
> - Simple models with an easy interpretation.
> - The asymmetric linear filter is independent of the date of estimation, but it depends on the data if we calibrate the filter with the "I-C" ratio.
>
> **Drawbacks**:
> - Timeliness is not controlled.
>
> Ⓡ **function**: `rjdfilters::lp_filter()`.

4.2 Reproducing Kernel Hilbert Space

Classical non-parametric filters (Henderson, LOESS, Hodrick-Prescott) can be characterized using the Reproducing Kernel Hilbert Space (RKHS) methodology, as described by Dagum and Bianconcini (2008). They use this theory to derive linear filters and associated asymmetric filters.

A RKHS is a Hilbert space characterized by a kernel that reproduces, via an inner product defined by a density function $f_0(t)$, every function of the space. Therefore, a kernel estimator K_p of order p (i.e.: that reproduce without distortion a polynomial trend of degree $p - 1$) can be decomposed into the product of

a reproducing kernel R_{p-1}, belonging to the space of polynomials of degree $p-1$, and a probability density function f_0.

For any sequence $(P_i)_{0 \leq i \leq p-1}$ of orthonormal polynomials in $\mathbb{L}^2(f_0)$[7], the kernel estimator K_p is defined by:

$$K_p(t) = \sum_{i=0}^{p-1} P_i(t) P_i(0) f_0(t)$$

The weights of a symmetric filter are then derived with:

$$\forall j \in [\![-h, h]\!] : w_j = \frac{K_p(j/b)}{\sum_{i=-h}^{h} K_p(i/b)}$$

where b is a time-invariant global bandwidth parameter.

The goal of Dagum and Bianconcini (2008) is to derive asymmetric filters from the Henderson symmetric filter, therefore the ideal kernel function would be the Henderson one. However, the Henderson density is a function of the bandwidth and needs to be calculated any time m changes (as its corresponding orthonormal polynomial). That's why the authors use the biweight kernel to approximate the Henderson kernel (for $h \geq 24$ they suggest considering the triweight kernel).

The asymmetric weighted are obtained adapting the kernels to the length of the asymmetric filters:

$$\forall j \in [\![-h, q]\!] : w_{a,j} = \frac{K_p(j/b)}{\sum_{i=-h}^{q} K_p(i/b)}$$

In Bee Dagum and Bianconcini (2015), the authors suggest performing an optimal bandwidth selection (parameter b). It can for example be chosen minimizing the mean squared revision error ($b_{q,\gamma}$) or the timeliness ($b_{q,\varphi}$):

$$b_{q,\gamma} = \min_{b_q} \sqrt{2 \int_0^{\pi} |\Gamma_s(\omega) - \Gamma_\theta(\omega)|^2 d\omega}$$

$$b_{q,\varphi} = \min_{b_q} \int_0^{2\pi/12} \rho_s(\lambda) \rho_\theta(\lambda) \sin^2\left(\frac{\varphi_\theta(\omega)}{2}\right) d\omega$$

Where Γ_s and ρ_s are the transfer function and the gain function of the symmetric filter and Γ_θ, ρ_θ and φ_θ are the transfer function, the gain function and the phase shift function of the asymmetric filter that depend on the bandwidth b.

---

[7]$\mathbb{L}^2(f_0)$ is the Hilbert space defined by the inner product:

$$\langle U(t), V(t) \rangle = \mathbb{E}[U(t)V(t)] = \int_{\mathbb{R}} U(t)V(t)f_0(t)dt$$

Summary — RKHS filters

Advantages:
- The asymmetric linear filter is independent of the data and of the date of estimation.
- Method can be generalized to create filter that could be applied to series with irregular frequency series (for example with a lot of missing values).

Drawbacks:
- With some parameters there are several minima that lead to very different results (for example with $h = 6$ and $q = 5$).

ℝ **function**: `rjdfilters::rkhs_filter()`.

4.3 General optimization problem: FST filters

Grun-Rehomme et al. (2018) defined a general approach to derive linear filters, based on an optimization problem of three criteria. Contrary to the local polynomial approach of Proietti and Luati (2008), this method enables to control the size of the timeliness in order to minimize the phase-shift.

The following criteria are used:
- *Fidelity*, F_g: it's the variance reduction ratio. It is called "Fidelity" because we want the output signal to be as close as possible to the input signal where the noise component is removed

$$F_g(\theta) = \sum_{k=-p}^{+f} \theta_k^2$$

F_g can be rewritten in a positive quadratic form: $F_g(\theta) = \theta' F \theta$ with F the identity matrix of order $p + f + 1$.

- *Smoothness*, S_g: it's the Henderson smoothness criterion (sum of the squared of the third difference of the coefficients of the filter). It measures the flexibility of the coefficient curve of a filter and the smoothness of the trend.

$$S_g(\theta) = \sum_j (\nabla^3 \theta_j)^2$$

S_g could also be rewritten in a positive quadratic form: $S_g(\theta) = \theta' S \theta$ with S a symmetric matrix of order $p + f + 1$.

- *Timeliness*, T_g: it measures the phase shift between input and output signal for specific frequencies. When a linear filter is applied, the level input signal is also altered by the gain function. Therefore, it is natural to consider that the higher the gain is, the higher the phase shift impact is. That's why the timeliness criterion depends on the gain and phase shift functions (ρ_θ and φ_θ), the link between both functions being made by a penalty function f.

$$T_g(\theta) = \int_0^{2\pi/12} f(\rho_\theta(\omega), \varphi_\theta(\omega)) d\omega$$

For the penalty function, the authors suggest taking $f: (\rho, \varphi) \mapsto \rho^2 \sin(\varphi)^2$. Indeed, for this function, the timeliness criterion is analytically solvable ($T_g = \theta' T \theta$ with T a square symmetric matrix of order $p + f + 1$), which is better in a computational point of view.

The asymmetric filters are computed minimizing a weighted sum of the past three criteria, subject to some constraints. Those constraints are usually polynomial preservation.

$$\begin{cases} \min_{\theta} & J(\theta) = \alpha F_g(\theta) + \beta S_g(\theta) + \gamma T_g(\theta) \\ s.t. & C\theta = a \end{cases}$$

The conditions $\alpha, \beta, \gamma \geq 0$ and $\alpha\beta \neq 0$ guarantee that $J(\theta)$ is a strictly convex function: therefore the optimization problem has a unique solution.

The Henderson symmetric filters can for example be computed with

$$C = \begin{pmatrix} 1 & \cdots & 1 \\ -h & \cdots & h \\ (-h)^2 & \cdots & h^2 \end{pmatrix}, \quad a = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \alpha = \gamma = 0, \quad \beta = 1$$

Summary — FST filters

Advantages:
- The asymmetric linear filter is independent of the symmetric filter, the data and the date of estimation.
- Unique solution to the optimization problem.
- The approach can be customized adding new criteria.

Drawbacks:
- The different criteria are not normalized: the associated weights cannot be compared.

Ⓡ function: `rjdfilters::fst_filter()`.

4.4 Data-dependent filter

In Wildi and McElroy (2019), the authors proposed a data-dependent approach to derive linear filters. They decompose the mean square revision error in a trilemma between three quantities: *accuracy*, *timeliness* and *smoothness*. Contrary to Ladiray (2018), this decomposition implies that the values of the three criteria are comparable.

Let:
- $\{x_t\}$ be our input time series;
- $\{y_t\}$ the target signal, i.e. the result of a symmetric filter, and Γ_s, ρ_s and φ_s the associated frequency response, gain and phase shift functions.
- $\{\hat{y}_t\}$ an estimation of $\{y_t\}$, i.e. the result of an asymmetric filter (when not all observations are available), and Γ_θ, ρ_θ and φ_θ the associated frequency response, gain and phase shift functions.

If we assume that $\{x_t\}$ is weakly stationary with a continuous spectral density h, the mean square revision error, $\mathbb{E}\left[(y_t - \hat{y}_t)^2\right]$, can be written as:

$$\begin{aligned} \mathbb{E}\left[(y_t - \hat{y}_t)^2\right] &= \frac{1}{2\pi} \int_{-\pi}^{\pi} |\Gamma_s(\omega) - \Gamma_\theta(\omega)|^2 h(\omega) \, d\omega \\ &= \frac{1}{2\pi} \times 2 \times \int_0^{\pi} |\Gamma_s(\omega) - \Gamma_\theta(\omega)|^2 h(\omega) \, d\omega \quad (1) \end{aligned}$$

This equality can also be generalized to non-stationary integrated process (for example imposing cointegration between both signals and using pseudo-spectral density, see Wildi and McElroy (2013)).

We have:

$$\begin{aligned} |\Gamma_s(\omega) - \Gamma_\theta(\omega)|^2 &= \rho_s(\omega)^2 + \rho_\theta(\omega)^2 + \\ & \quad 2\rho_s(\lambda)\rho_\theta(\lambda)\left(1 - \cos(\varphi_s(\omega) - \varphi_\theta(\omega))\right) \\ &= (\rho_s(\omega) - \rho_\theta(\omega))^2 + \\ & \quad 4\rho_s(\lambda)\rho_\theta(\lambda)\sin^2\left(\frac{\varphi_s(\omega) - \varphi_\theta(\omega)}{2}\right) \end{aligned}$$

The interval $[0, \pi]$ is then split in two: the pass-band $[0, \omega_1]$ (the frequency interval that contains the target signal) and the stop-band $[\omega_1, \pi]$.

The mean squared error defined in equation (1) can then be decomposed additively into four quantities, accuracy, timeliness, smoothness and residual:

$$\begin{aligned} A_w &= 2 \int_0^{\omega_1} (\rho_s(\omega) - \rho_\theta(\omega))^2 h(\omega) \, d\omega \\ T_w &= 8 \int_0^{\omega_1} \rho_s(\lambda)\rho_\theta(\lambda)\sin^2\left(\frac{\varphi_\theta(\omega)}{2}\right) h(\omega) \, d\omega \\ S_w &= 2 \int_{\omega_1}^{\pi} (\rho_s(\omega)^2 - \rho_\theta(\omega))^2 h(\omega) \, d\omega \\ R_w &= 8 \int_{\omega_1}^{\pi} \rho_s(\lambda)\rho_\theta(\lambda)\sin^2\left(\frac{\varphi_\theta(\omega)}{2}\right) h(\omega) \, d\omega \end{aligned}$$

In general, the residual R_w is small since $\rho_s(\omega)\rho_\theta(\omega)$ is close to 0 in the stop-band. Moreover, user priorities are rarely concerned about the time-shift properties of components in the stop-band. That's why, to derive linear filters the residual is not taken into account; and that's why the authors suggest minimizing a weighted sum of the first three indicators:

$$\mathcal{M}(\vartheta_1, \vartheta_2) = \vartheta_1 T_w(\theta) + \vartheta_2 S_w(\theta) + (1 - \vartheta_1 - \vartheta_2) A_w(\theta)$$

┌─ **Summary — Data-dependent filters** ─┐

Advantages:
- The values of the different criteria can be compared: the weights can be easily interpreted.

Drawbacks:
- Data-dependent filter: it depends on the symmetric filter, the data and the date of estimation.
- Some optimization problems might occur (several minimum, etc.).

ℝ **implementation**: package `MDFA` (https://github.com/wiaidp/MDFA).

For the moment, this method is not compared to the others: the filter is totally data-dependent, so it can be difficult to implement it in the seasonally adjusted method X-12ARIMA. Moreover, due to the high degree of freedom ("customization triangle" of $\mathcal{M}(\vartheta_1, \vartheta_2)$), it would need more studied to decide which parameter to choose.

5. Comparison of the different methods

5.1 Comparison with the FST approach

The FST approach provides a useful tool to validate a method of construction of linear filter. Indeed, a linear filter can be considered as suboptimal if we can find a set of weight for the FST approach that gives better results (in terms of fidelity, smoothness and timeliness) with the same (or higher) polynomial constraints.

For $h = 6$ (13-term symmetric filter), for the RKHS filters we find that:
- imposing asymmetric filters to preserve constants, the FST approach gives better results for all values of q for the filters computed with $b_{q,G}$ and $b_{q,\gamma}$, and for $q \leq 3$ for the one computed with $b_{q,\varphi}$.
- imposing asymmetric filters to preserve linear trends, the FST approach gives better results for $q \in [2,5]$ for the filters computed with $b_{q,G}$ and $b_{q,\gamma}$, and for $q \in [2,4]$ for the one computed with $b_{q,\varphi}$.

❯ In terms of quality criteria, FST approach seems to perform better than the RKHS approach to build asymmetric filters to apply to monthly regular data (i.e.: data for which we usually use $h = 6$).

For the local polynomial filters we find that:
- imposing asymmetric filters to preserve constants, the FST approach gives better results than the LC method for all values of q.
- imposing asymmetric filters to preserve linear trends, the FST approach gives better results for $q \in [2,5]$ for the LC and QL methods.

❯ In terms of quality criteria, local polynomial seems to perform better for real-time and one-period ahead filters ($q = 0, 1$).

Those comparisons can be seen in an online application available at https://aqlt.shinyapps.io/FSTfilters/.

5.2 Comparison in term of turning point detection

5.2.1 Methodology

In this section, we compare the filters in terms of delay to detect turning points. A turning point (upturn) is defined to occur at time t if $y_{t-k} \geq \cdots \geq y_{t-1} < y_t \leq y_{t+1} \leq \cdots y_{t+m}$. A downturn would be equivalently defined as a date t $y_{t-k} \leq \cdots \leq y_{t-1} > y_t \geq y_{t+1} \geq \cdots y_{t+m}$. Following Zellner et al. (1991), $k = 3$ and $m = 1$ are chosen given the smoothness of the trend-cycle data.

We used the calendar adjusted Eurostat database on monthly industrial production indices (`sts_inpr_m`, 2 404 series). The turning points are determined on the trend-cycle component computed with the seasonal adjustment method X-12ARIMA on the overall series. The time delay is defined as the number of months necessary to detect the correct turning point without any further revisions.

For each series and each date, the following methodology is used:

1. The series is seasonally adjusted with X-12ARIMA to extract: the linearized component, the length of the trend and seasonal filters, the decomposition scheme and the I-C ratio

2. The seasonal adjusted is done on the linearized component with the X-11 algorithm, using the same parameters as before but with a different trend-cycle filter and without using the forecasts of the pre-adjustment[8]. To control the impact of the estimates of the parameters of the Reg-ARIMA model, the linearized series used here is the one computed on the overall series.

3. For each estimate of the trend, we compute the downturns and the upturns.

Due to the high degree of freedom of the FST approach, for the moment we only compared the X-12ARIMA approach to local polynomial and RKHS filters[9]. For the RKHS filters, we only retain the ones that minimize the timeliness because in our case we want to minimize the time delay of turning point detection.

5.2.2 Results

The table 1 compares the different methods for the turning points detection during the Covid-19 crisis (i.e.: during the year 2020). The LC method (i.e.: Musgrave filters) seems to perform, in distribution, better than the other methods (and have similar performance than the X-12ARIMA method) and the less performant method seems to be the RKHS filters obtained by minimizing the timeliness. This could be ex-

---

[8]This can be done with the ℝ functions `rjdfilters::x11_lp()`, `rjdfilters::x11_rkhs()` and `rjdfilters::x11_fst()`.

[9]Further research will be done on the FST approach to see which set of weights performs better. One lead could be the study the set of weights that produces linear filters with better quality criteria than RKHS or local polynomial filters.

plained by the issue underlined in section 4.2: for the RKHS filters we can have several extrema in the optimization problem that lead to different asymmetric filters. The worst performance of QL and CQ filters compared to LC could be explained by the highest volatility of the estimates due to the strongest constraints on the polynomial trend conservation.

Figures 1 to 4 show two examples of trend-cycle estimates around the turning point of February 2020. For those series, the real time trend-cycle estimates produced by the methods QL, CQ and DAF seem unrealistic (too high rise or fall), and the ones produce by RKHS filters seems too noisy. X-12ARIMA and Musgrave filters (LC) seems to produce very similar results.

The same statistics can be computed in over periods, such as the global financial crisis of 2007–2008 (table 2). We observe the same hierarchy between the methods: the best results are obtained with the LC filters and X-12ARIMA. What can be surprising is that for almost all the series, the time delay to detect turning points without any further revisions is above 6 months, which is the number of months necessary to use the symmetric filter (the results are only displayed for the series that uses filters of order 13). This will be investigated in further studies, but it could be explained by the impact of outliers, detected during crisis that are corrected in the pre-adjustment and during the decomposition.

6. First conclusions

In this ongoing research, we describe and compare some modern approach to build asymmetric filters to minimize revision and time shift in the detection of turning points that could be used in the seasonal adjustment method X-12ARIMA.

From this first results, we can conclude that there is no need to build asymmetric filters which preserve polynomial trends of degree more than one (QL, CQ, DAF filters): it introduces more variance on the estimates (and thus more revisions), without any gain in term of detection of turning point. This implies that, in further simulations on the FST approach, we can focus on the one that preserves, at best, polynomial linear trends.

This study could be extended in many ways. For example, we could study other modern approaches, such as the Direct Filter Approach, or study the impact of atypical points (and thus new kinds of asymmetric filters based on robust methods). For the JSM presentation, we plan to investigate deeper the results presented in this draft manuscript.

References

Alexandrov, T., S. Bianconcini, E. Dagum, P. Maass, and T. Mcelroy (2012, 11). A review of some modern approaches to the problem of trend extraction. *Econometric Reviews 31*, 593–624.

Bee Dagum, E. and S. Bianconcini (2015). A new set of asymmetric filters for tracking the short-term trend in real-time. *Ann. Appl. Stat. 9*(3), 1433–1458.

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

Grun-Rehomme, M., F. Guggemos, and D. Ladiray (2018). Asymmetric moving averages minimizing phase shift. *Handbook on Seasonal Adjustment*.

Ladiray, D. (2018). Moving average based seasonal adjustment. *Handbook on Seasonal Adjustment*.

Musgrave, J. C. (1964). A set of end weights to end all end weights. *US Census Bureau [custodian]*.

Pierce, D. A. (1980). Data revisions with moving average seasonal adjustment procedures. *Journal of Econometrics 14*(1), 95–114.

Proietti, T. and A. Luati (2008). Real time estimation in local polynomial regression, with application to trend-cycle analysis. *Ann. Appl. Stat. 2*(4), 1523–1553.

Wildi, M. and T. McElroy (2013). Optimal real-time filters for linear prediction problems. *Journal of Time Series Econometrics 8*.

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

Zellner, A., C. Hong, and C.-k. Min (1991). Forecasting turning points in international output growth rates using bayesian exponentially weighted autoregression, time-varying parameter, and pooling techniques. *Journal of Econometrics 49*(1-2), 275–304.

Figure 1. Estimates of the trend-cycle component for the industrial production index in the manufacture of cement, lime and plaster (C235) in Germany (turning point in February 2020).

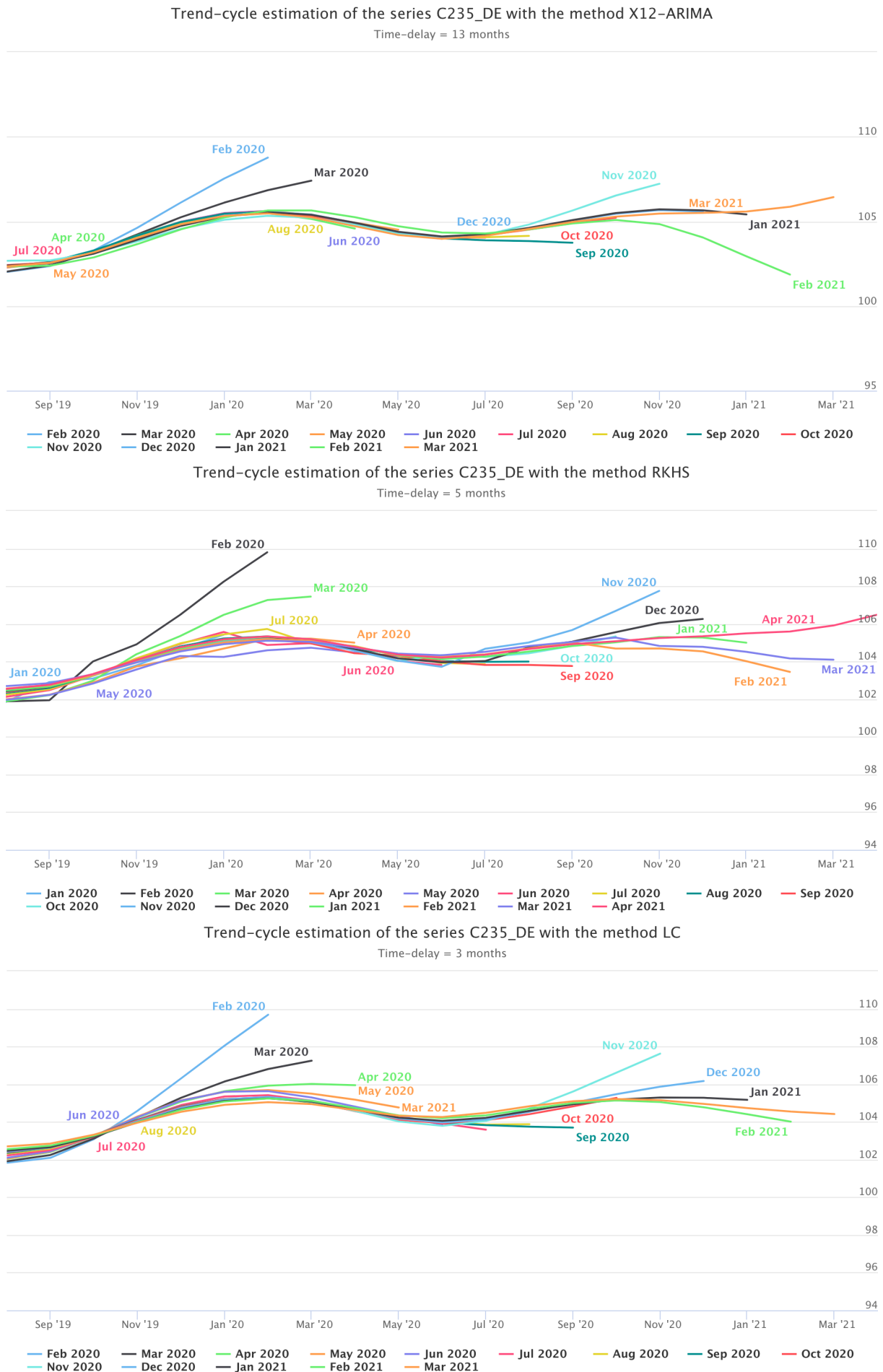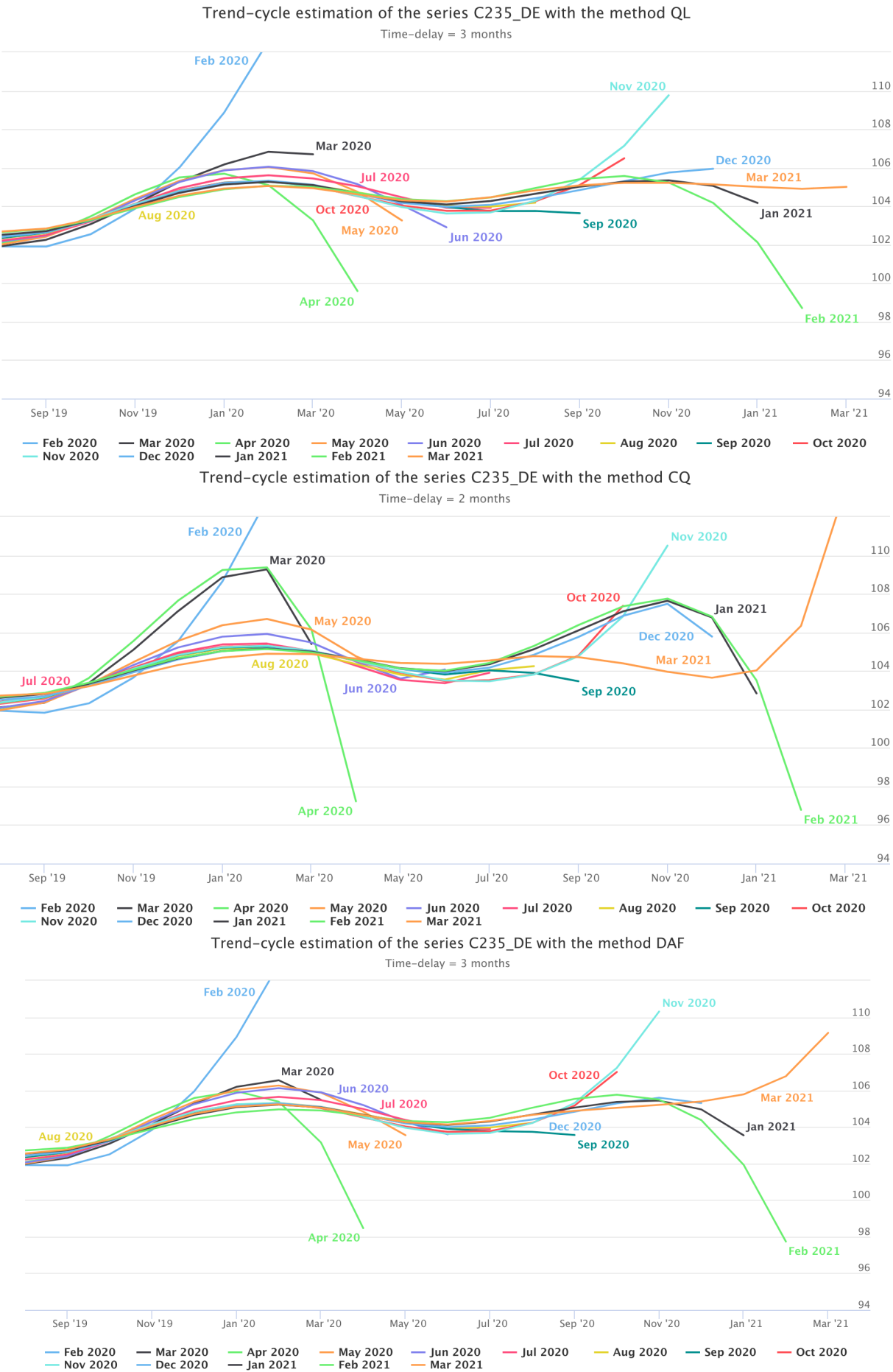**Figure 2.** Estimates of the trend-cycle component for the industrial production index in the manufacture of cement, lime and plaster (C235) in Germany (turning point in February 2020).

Trend–cycle estimation of the series C25_SE with the method X12–ARIMA

Time–delay = 5 months

Trend–cycle estimation of the series C25_SE with the method RKHS

Time–delay = 13 months

Trend–cycle estimation of the series C25_SE with the method LC

Time–delay = 13 months

Figure 3. Estimates of the trend-cycle component for the industrial production index in the manufacture of fabricated metal products, except machinery and equipment (C25) in Sweden (turning point in February 2020).

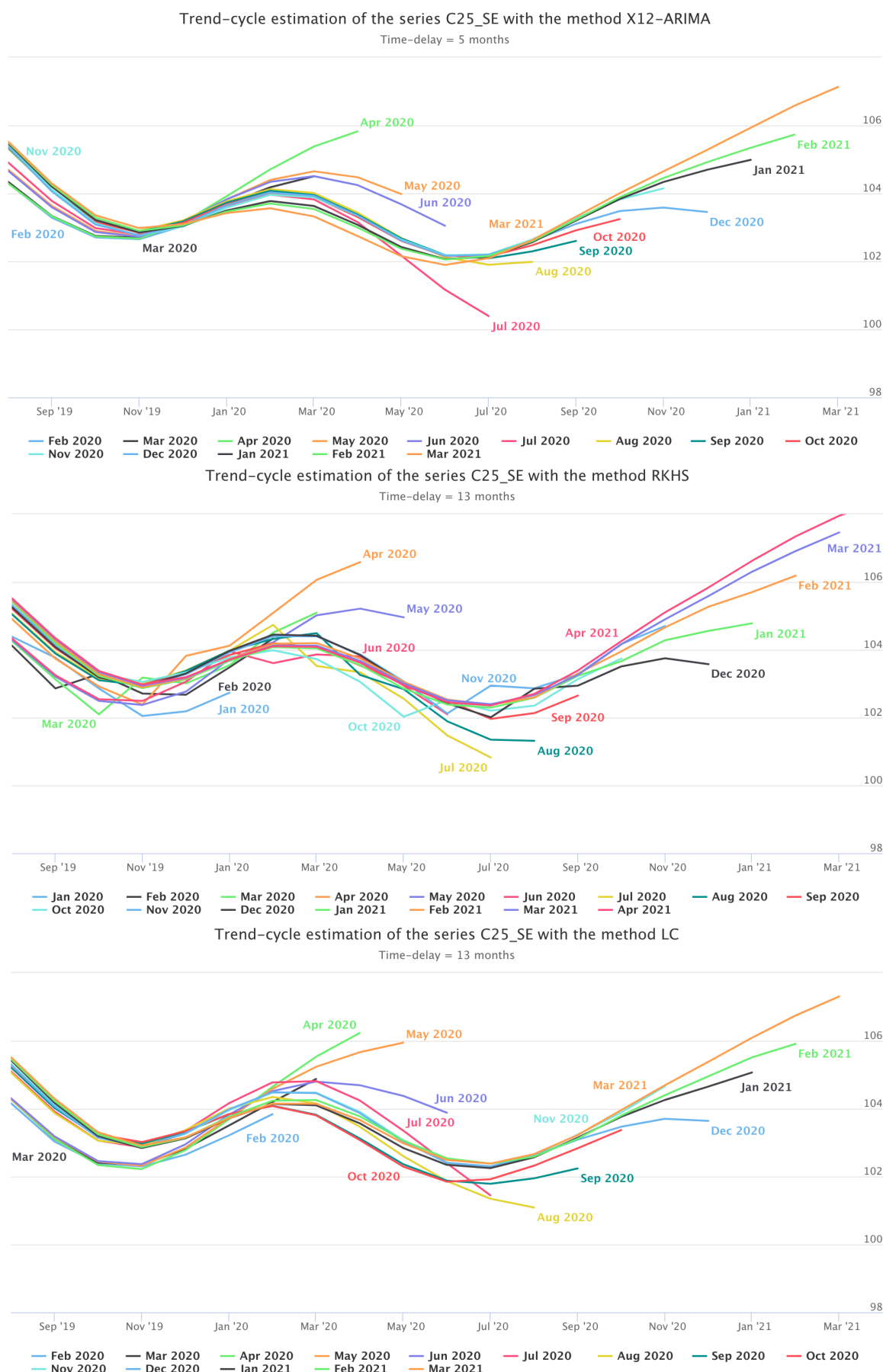**Figure 4.** Estimates of the trend-cycle component for the industrial production index in the manufacture of fabricated metal products, except machinery and equipment (C25) in Sweden (turning point in February 2020).
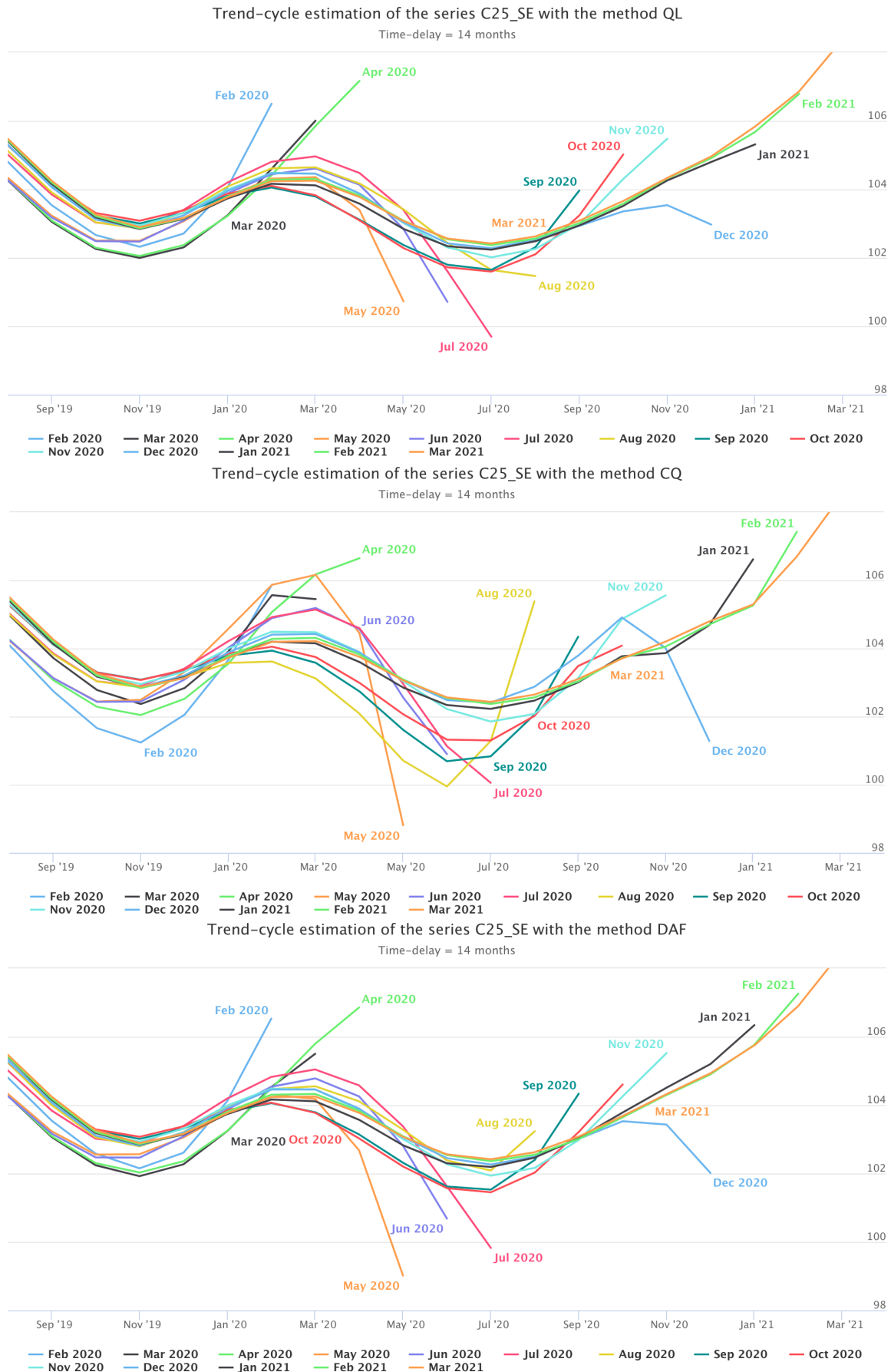
Table 1. Deciles of the time delay to detect turning points (downturn and upturn) in 2020 (900 observations).

	X12-ARIMA	RKHS	LC	QL	CQ	DAF
Min	2.0	2.00	2.00	2.00	2.00	2.00
D1	2.0	5.00	3.00	2.00	2.00	2.00
D2	3.0	5.00	3.00	3.00	2.00	2.00
D3	3.0	5.00	4.00	3.00	3.00	3.00
D4	3.0	5.00	4.00	3.00	5.00	3.00
Median	4.0	5.00	4.00	4.00	6.00	5.00
D6	4.0	7.00	4.00	5.00	6.00	6.00
D7	4.0	7.00	5.00	7.00	7.00	7.00
D8	5.0	7.00	5.00	7.00	7.00	7.00
D9	9.0	9.00	8.00	9.00	9.00	9.00
Max	14.0	14.00	14.00	14.00	14.00	14.00
Mean	4.4	6.29	4.69	4.97	5.32	5.09

Note:
Statistics computed on series for which the optimal trend-cycle filter is of length 13.
D2: 20% of the turning points are detected with less than 3 months for LC filter and less than 5 months for RKHS filter.

Table 2. Deciles of the time delay to detect turning points (downturn and upturn) between 2007 and 2009 (2 535 observations).

	X12-ARIMA	RKHS	LC	QL	CQ	DAF
Min	2.00	5.00	2.00	2.00	2.00	2.00
D1	6.00	7.00	5.00	7.00	8.00	7.00
D2	15.00	16.00	15.00	16.00	16.00	16.00
D3	21.00	22.00	21.00	21.00	22.00	21.00
D4	26.00	27.60	27.00	27.00	28.00	27.00
Median	31.00	34.00	32.00	33.00	33.00	33.00
D6	38.00	41.00	39.00	40.00	40.00	40.00
D7	49.80	52.00	50.00	51.00	51.00	51.00
D8	69.00	71.00	68.00	69.00	69.00	69.00
D9	126.20	129.00	127.00	127.00	127.00	127.00
Max	169.00	169.00	169.00	169.00	169.00	169.00
Mean	46.15	47.94	46.03	46.68	47.18	46.82

Note:
Only series for which the optimal trend-cycle filter is of length 13.
D2: 20% of the turning points are detected with less than 15 months for LC filter and X-12ARIMA and less than 16 months for other methods.