Hyper-trend method for seasonal adjustment and trend-cycle decomposition of time series containing long-period cycles

Koki Kyo*

Faculty of Management and Information Science,
Niigata University of Management,
Kamo-shi, Niigata, Japan
Email: kyo@duck.niigataum.ac.jp
*Corresponding author

Genshiro Kitagawa

Mathematics and Informatics Center,
Tokyo University,
Bunkyo-ku, Tokyo, Japan
and
Meiji Institute for Advanced Study of Mathematical Sciences,
Meiji University,
Nakano-ku, Tokyo, Japan
Email: kitagawa@mi.u-tokyo.ac.jp

Abstract: In some instances, existing methods for decomposing a time series into several components cannot capture cyclical components that contain long-period cycles. We propose a systematic methodology to overcome this problem. In the proposed hyper-trend method, we assume that part of the cyclical variation is included in the estimate of the trend component. We then capture the remainder of the cyclical variation by re-decomposing the estimate of the trend component. The average coefficient of determination is introduced to evaluate the decomposed results. An overall procedure for applying the proposed approach is developed, and the performance of the proposed approach is demonstrated by analysing 20 commercial sales time series and 30 business cycle time series.

Keywords: seasonal adjustment; trend-cycle decomposition; hyper-trend method; state space model; Kalman filter; economic time series; commercial sales; business cycles; DECOMP; stationarity of a time series.

Reference to this paper should be made as follows: Kyo, K. and Kitagawa, G. (2021) 'Hyper-trend method for seasonal adjustment and trend-cycle decomposition of time series containing long-period cycles', *Asian J. Management Science and Applications*, Vol. 6, No. 2, pp.134–162.

Biographical notes: Koki Kyo is a Professor of Data Science at the Niigata University of Management, Japan. He received his BE and ME in Management Engineering from the Northeastern University, China, in 1982 and 1984, respectively, and PhD in Statistical Science from the Graduate University for Advanced Studies, Japan, in 1993. His current research interests include time series analysis, Bayesian statistical modelling, and identification of stochastic systems. He is a member of the Japan Statistical Society, and a member of the Asian Association of Management Science and Applications (AAMSA).

Genshiro Kitagawa received his BS and MS in Mathematics from the Tokyo University, Japan, in 1971 and 1973, respectively, and Doctor of Science in Mathematics from the Kyushu University, Japan, in 1983. He was the Director-General of the Institute of Statistical Mathematics, Japan in the period 2002–2011, and the President of Research Organization of Information and Systems, Japan in the period 2011–2017. He is currently a project professor at the University of Tokyo, Japan. His current research interests include time series analysis and statistical modelling. He is a Fellow of the American Statistical Association and a member of the International Statistical Institute, and many others.

1 Introduction

Generally, an economic time series consists of several components, such as trend, seasonal, and cyclical components. Specifically, for a monthly (or quarterly) time series, the trend component expresses the long-term variation, which varies gradually and smoothly. The seasonal component repeats a similar annual pattern. Moreover, the seasonal and cyclical components are similar because both have a periodic variation. Unlike seasonal variation, however, the cyclical component does not necessarily vary repeatedly in a specific pattern. Additionally, similar to the trend component, the seasonal component has a non-stationary mean, whereas the cyclical component can typically be assumed to be stationary. For economic time series, economic growth is represented by the trend component and the business cycles are measured by the cyclical component.

Seasonal adjustment is particularly important from a macro-econometric perspective because, for example, government economic policies are examined using seasonally adjusted economic indicators. The most widely used method for seasonal adjustment in current use is Census X-11, which was developed by the United States Census Bureau. This method uses a set of moving averages to produce seasonally adjusted data [see Shiskin et al., 1967; see also, e.g., Cleveland and Tiao (1976) for applications of the X-11 procedure]. Recent improvements to X-11 include X-12-ARIMA (Findley et al., 1998), TRAMO-SEATS (Gómez and Maravall, 1996), and X-13ARIMA-SEATS (Monsell, 2007).

However, as argued by Bell and Hillmer (2002), the statistical underpinnings of these methods are not always clear. Additionally, the complexity of the algorithm renders statistical processing difficult. In particular, the basic scheme of the X-12-ARIMA program is an autoregressive integrated moving average (ARIMA) approach (see Box and Jenkins, 1976), which is based on a differentiated series. As indicated by Granger and Joyeux (1980) and Sims et al. (1990), the differentiated series typically leads to some loss of information. Recognising the drawbacks inherent in X-11 and X-12-ARIMA, Akaike (1980) proposed a seasonal adjustment method based on Bayesian linear modelling, and developed the BAYSEA program (see Akaike and Ishiguro, 1980; Ishiguro, 1984).

Conventional seasonal adjustment programs decompose a time series into trend-cyclical, seasonal, and irregular components. For instance, Kitagawa (1981), Gersch and Kitagawa (1983), Kitagawa and Gersch (1984) and Kitagawa (2020) introduced the trend and cyclical parts into the model as separate components. This is very important for empirical analysis, particularly in economics, because the separate estimates for these two components are very useful for economic analysis. Based on Kitagawa's approach, the DECOMP program was developed (Kitagawa, 1985). In the DECOMP program, Bayesian smoothness priors are applied to the trend and seasonal components, and the cyclical component is represented by a stationary autoregressive (AR) model. A state space model is constructed as a representation of the Bayesian linear model. Thus, the Kalman filter can be applied to estimate the components, and a maximum likelihood method based on the Kalman filter is used for parameter estimation.

Despite this, it is sometimes difficult to capture the cyclical variation completely using the DECOMP program when the time series contains long-period cycles. In such a case, part of the cyclical variation may remain in the estimate of the trend component, so the estimate of the trend component may contain some unnecessary fluctuations. Thus, when processing the results of conventional seasonal adjustment, it is necessary to complete a trend-cycle decomposition process, that is, decompose the trend-cyclical component into trend and cyclical components. There are many approaches for trend-cycle decomposition, such as the Beveridge-Nelson approach (Beveridge and Nelson, 1981), unobservable component modelling (Watson, 1986; Clark, 1987), the Hodrick-Prescott filter (Hodrick and Prescott, 1997), the Baxter-King approach (Baxter and King, 1999), and the Morley-Nelson-Zivot approach (Morley et al., 2003). Although these approaches are useful for various applications, in many existing techniques for trend-cycle decomposition, the trend is expressed by a stochastic difference model in which the growth rate of the trend is a parameter. Moreover, in most existing methods, the cyclical component is expressed by a second-order AR model that lacks generality. Because many economic time series contain long-period cycles, it is difficult to capture the cyclical variation clearly using an AR model with a small order.

The objective of the present study is to develop an approach for decomposing a trend-cyclical series into trend and cyclical components. The proposed approach can be applied to perform trend-cycle decomposition for seasonal adjustment, and can be also used as a general tool for trend-cycle decomposition. This approach is called the *hyper-trend method* because, at its core, the modelling method is related to the concept of the hyper-trend. The main contribution of the present study is to develop a systematic program that can be applied to cases in which objective time series contain long-period cycles. Note that, for a similar purpose, Harvey et al. (2007) proposed a Bayesian modelling method, although their estimation method is somewhat complicated.

The remainder of this paper is organised as follows: in Section 2, we provide a full review of the DECOMP program, as this is a basic routine in the proposed method. In Section 3, we introduce the aim and requirements of the proposed method. Section 4 presents the proposed hyper-trend method, including several modelling techniques and the overall algorithm. In Section 5, we provide examples that demonstrate the performance of the newly proposed approach. Finally, we provide a summary and discussion in Section 6.

2 DECOMP program: a review

DECOMP is a program for seasonal adjustment that was originally developed by Kitagawa (1981) and Kitagawa and Gersch (1984). In this section, we provide a full review of the details of the DECOMP program, as this is used as a basic routine in the proposed hyper-trend approach.

2.1 Model

Although the DECOMP program contains a wide class of models, for the purpose of the present study, we focus on the following basic form:

$$y_n = t_n + s_n + c_n + w_n \qquad \text{(observation model)}, \tag{1}$$

$$t_n = 2t_{n-1} - t_{n-2} + v_{n1} \qquad \text{(trend component model)}, \tag{2}$$

$$s_n = -\sum_{j=1}^{p-1} s_{n-j} + v_{n2} \qquad \text{(seasonal component model)}, \tag{3}$$

$$c_n = \sum_{j=1}^{q} \alpha_j c_{n-j} + v_{n3} \qquad \text{(cyclical component model)}. \tag{4}$$

We call the above model set the *basic seasonal adjustment* (BSA) model. In equations (1)–(4), y_n is a monthly (or quarterly) time series that is observed on time points $n = 1, 2, \dots, N$, where N is the data span. t_n, s_n, c_n, and w_n represent the trend, seasonal, cyclical, and irregular components, respectively. p represents the period of seasonal variation; specifically, for monthly data $p = 12$ and for quarterly data $p = 4$. q is the order of the AR model for the cyclical component, and $\alpha_1, \alpha_2, \dots, \alpha_q$ are the AR coefficients. To enable statistical analysis, $w_n \sim \mathrm{N}(0, \sigma^2)$, $v_{n1} \sim \mathrm{N}(0, \tau_1^2)$, $v_{n2} \sim \mathrm{N}(0, \tau_2^2)$, and $v_{n3} \sim \mathrm{N}(0, \tau_3^2)$ are assumed to be white noise sequences that are independent of each other, where σ^2, τ_1^2, τ_2^2, and τ_3^2 are the variances. Typically, the AR coefficients $\boldsymbol{\alpha} = \{\alpha_1, \alpha_2, \dots, \alpha_q\}$ and the variances σ^2, τ_1^2, τ_2^2, τ_3^2 are regarded as unknown parameters.

The BSA model in equations (1)–(4) is essentially a Bayesian model in which the first equation expresses the structure of observations for the time series y_n and the other equations express the priors for each component to be estimated. To express the BSA model in a state space form, the following notation is used. First, the vectors $\boldsymbol{x}_{n1}$, $\boldsymbol{x}_{n2}$, and $\boldsymbol{x}_{n3}$ are defined based on the trend, seasonal, and cyclical components, respectively:

$$\begin{aligned} \boldsymbol{x}_{n1} &= (t_n, t_{n-1})^{\mathrm{t}}, \\ \boldsymbol{x}_{n2} &= (s_n, s_{n-1}, \dots, s_{n-p+2})^{\mathrm{t}}, \\ \boldsymbol{x}_{n3} &= (c_n, c_{n-1}, \dots, c_{n-q+1})^{\mathrm{t}}, \end{aligned} \tag{5}$$

and correspondingly,

$$\boldsymbol{F}_1 = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}, \ \boldsymbol{F}_2 = \begin{bmatrix} -1 & \cdots & \cdots & -1 \\ 1 & 0 & \cdots & 0 \\ & \ddots & \ddots & \vdots \\ 0 & & 1 & 0 \end{bmatrix}, \ \boldsymbol{F}_3 = \begin{bmatrix} \alpha_1 & \cdots & \cdots & \alpha_q \\ 1 & 0 & \cdots & 0 \\ & \ddots & \ddots & \vdots \\ 0 & & 1 & 0 \end{bmatrix},$$
$$\boldsymbol{G}_1 = \begin{bmatrix} 1 & 0 \end{bmatrix}^{\mathtt{t}}, \ \boldsymbol{G}_2 = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}^{\mathtt{t}}, \ \boldsymbol{G}_3 = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}^{\mathtt{t}},$$
$$\boldsymbol{H}_1 = \begin{bmatrix} 1 & 0 \end{bmatrix}, \ \boldsymbol{H}_2 = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}, \ \boldsymbol{H}_3 = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}. \tag{6}$$

Based on the above settings, a state vector $\boldsymbol{x}_n$ of dimension $p + q + 1$, and a system noise vector $\boldsymbol{v}_n$ of dimension 3 are defined, respectively, as

$$\boldsymbol{x}_n = (\boldsymbol{x}_{n1}^{\mathtt{t}}, \boldsymbol{x}_{n2}^{\mathtt{t}}, \boldsymbol{x}_{n3}^{\mathtt{t}})^{\mathtt{t}}, \quad \boldsymbol{v}_n = (v_{n1}, v_{n2}, v_{n3})^{\mathtt{t}}, \tag{7}$$

and the matrices

$$\boldsymbol{F} = \begin{bmatrix} \boldsymbol{F}_1 & \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{F}_2 & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{F}_3 \end{bmatrix}, \ \boldsymbol{G} = \begin{bmatrix} \boldsymbol{G}_1 & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{G}_2 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{G}_3 \end{bmatrix}, \ \boldsymbol{H} = \begin{bmatrix} \boldsymbol{H}_1 & \boldsymbol{H}_2 & \boldsymbol{H}_3 \end{bmatrix}, \tag{8}$$

are constructed, where $\boldsymbol{O}$ denotes the zero matrix and $\boldsymbol{0}$ denotes the zero vector with appropriate dimensions. Hence, the BSA model can be expressed in state space form as follows:

$$\boldsymbol{x}_n = \boldsymbol{F}\boldsymbol{x}_{n-1} + \boldsymbol{G}\boldsymbol{v}_n, \tag{9}$$
$$y_n = \boldsymbol{H}\boldsymbol{x}_n + w_n. \tag{10}$$

Following the assumptions for the BSA model, we have $\boldsymbol{v}_n \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{Q})$, $w_n \sim \mathrm{N}(0, R)$, where $\boldsymbol{Q} = \mathrm{diag}(\tau_1^2, \tau_2^2, \tau_3^2)$ and $R = \sigma^2$.

In the state space model given by equations (9) and (10), the components t_n, s_n, and c_n are involved in the state vector $\boldsymbol{x}_n$; thus, their estimates can be obtained from the estimate of $\boldsymbol{x}_n$ using the Kalman filter algorithm. Moreover, the unknown parameters $\boldsymbol{\alpha}$, σ^2, τ_1^2, τ_2^2, and τ_3^2 can be estimated using the maximum-likelihood method based on a numerical optimisation routine.

A special case of the BSA model is that in which the time series y_n is seasonally adjusted. In this case, we can set the seasonal component as $s_n = 0$, so the seasonal component model in equation (3) is not necessary. Thus, the terms $\boldsymbol{x}_{n2}$, $\boldsymbol{F}_2$, $\boldsymbol{G}_2$, $\boldsymbol{H}_2$, v_{n2}, and τ_2^2, which is related to the seasonal component, can be omitted from the above settings. We call this model the *trend-cycle decomposition* (TCD) model.

2.2 Estimating the components

Let $\boldsymbol{x}_0 \sim \mathrm{N}(\boldsymbol{x}_{0|0}, \boldsymbol{V}_{0|0})$ denote the state $\boldsymbol{x}_n$ at time point $n = 0$, $Y_{1:j} = \{y_1, y_2, \ldots, y_j\}$ be a set of observations of y_n up to time point $n = j$, and $f(\boldsymbol{x}_n|Y_{1:j})$ express the posterior probability density of $\boldsymbol{x}_n$ given the data $Y_{1:j}$. Then, for given values of $\boldsymbol{x}_{0|0}$ and $\boldsymbol{V}_{0|0}$, the distribution for the state $\boldsymbol{x}_n$ conditioned on $Y_{1:k}$ is Gaussian; that is, $f(\boldsymbol{x}_n|Y_{1:j})$ can be expressed as a Gaussian density with mean $\boldsymbol{x}_{n|j}$ and covariance matrix $\boldsymbol{V}_{n|j}$, so it is only necessary to obtain the mean $\boldsymbol{x}_{n|j}$ and the covariance matrix $\boldsymbol{V}_{n|j}$.

Given the values of the parameters $\boldsymbol{\alpha}$, $\sigma^2, \tau_1^2, \tau_2^2$, and τ_3^2, the initial conditions $\boldsymbol{x}_{0|0}$, $\boldsymbol{V}_{0|0}$, and the observations $Y_{1:N} = \{y_1, y_2, \ldots, y_N\}$, the means and the covariance matrices of the state $\boldsymbol{x}_n$ for $n = 1, 2, \ldots, N$ can be obtained using the following Kalman filter, which is composed of *one-step ahead prediction* and a *filter* (see, e.g., Harvey, 1989; Kitagawa 2020):

- *One-step ahead prediction*

$$\boldsymbol{x}_{n|n-1} = \boldsymbol{F}\boldsymbol{x}_{n-1|n-1},$$
$$\boldsymbol{V}_{n|n-1} = \boldsymbol{F}\boldsymbol{V}_{n-1n-1}\boldsymbol{F}^{\mathrm{t}} + \boldsymbol{G}\boldsymbol{Q}\boldsymbol{G}^{\mathrm{t}}.$$

- *Filter*

When y_n is observed,

$$\boldsymbol{K}_n = \boldsymbol{V}_{n|n-1}\boldsymbol{H}^{\mathrm{t}}(\boldsymbol{H}\boldsymbol{V}_{n|n-1}\boldsymbol{H}^{\mathrm{t}} + R)^{-1},$$
$$\boldsymbol{x}_{n|n} = \boldsymbol{x}_{n|n-1} + \boldsymbol{K}_n(y_n - \boldsymbol{H}\boldsymbol{x}_{n|n-1}),$$
$$\boldsymbol{V}_{n|n} = (\boldsymbol{I} - \boldsymbol{K}_n\boldsymbol{H})\boldsymbol{V}_{n|n-1}.$$

When y_n is a missing value, the above filter step is replaced by

$$\boldsymbol{x}_{n|n} = \boldsymbol{x}_{n|n-1}, \quad \boldsymbol{V}_{n|n} = \boldsymbol{V}_{n|n-1}.$$

Based on the results of the Kalman filter, as the final estimate of $\boldsymbol{x}_n$, its posterior distribution conditioned on all observations $Y_{1:N}$ can be obtained using the following *fixed-interval smoothing* for $n = N-1, N-2, \ldots, 1$:

- *Fixed-interval smoothing*

$$\boldsymbol{A}_n = \boldsymbol{V}_{n|n}\boldsymbol{F}^{\mathrm{t}}\boldsymbol{V}_{n+1|n}^{-1},$$
$$\boldsymbol{x}_{n|N} = \boldsymbol{x}_{n|n} + \boldsymbol{A}_n(\boldsymbol{x}_{n+1|N} - \boldsymbol{x}_{n+1|n}),$$
$$\boldsymbol{V}_{n|N} = \boldsymbol{V}_{n|n} + \boldsymbol{A}_n(\boldsymbol{V}_{n+1|N} - \boldsymbol{V}_{n+1|n})\boldsymbol{A}_n^{\mathrm{t}}.$$

Note that the result of fixed-interval smoothing for $n = N$ is contained in the results of the above Kalman filter.

Thus, the estimates for the components t_n, s_n, c_n can be obtained because the state space model described by equations (9) and (10) incorporates t_n, s_n, c_n in the state vector $\boldsymbol{x}_n$. The posterior distribution for the state vector $\boldsymbol{x}_n$, which is characterised by $\boldsymbol{x}_{n|N}$ and $\boldsymbol{V}_{n|N}$, is Gaussian, so estimates of t_n, s_n and c_n can be determined based on $\boldsymbol{x}_{n|N}$, and the estimate of the irregular component is obtained using equation (1) based on the estimates of t_n, s_n, and c_n. Note that, as shown in the filter step, the Kalman filter can be implemented without any problems for cases in which the time series contains missing values, which is a remarkable merit of the Kalman filter algorithm.

2.3 Parameter estimation

Given the time series data $Y_{1:N}$, the likelihood function for the parameters $\boldsymbol{\theta} = \{\boldsymbol{\alpha}, \sigma^2, \tau_1^2, \tau_2^2, \tau_3^2\}$ is given by (see Kitagawa, 2020)

$$L(\boldsymbol{\theta}|Y_{1:N}) = f(Y_{1:N}|\boldsymbol{\theta}) = \prod_{n=1}^{N} f_n(y_n|Y_{1:(n-1)}, \boldsymbol{\theta}), \tag{11}$$

where $f_n(y_n|Y_{1:(n-1)}, \boldsymbol{\theta})$ is the conditional density function of y_n, that is, a normal density given by

$$f_n(y_n|Y_{1:(n-1)}, \boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi\eta_{n|n-1}^2}} \exp\left\{-\frac{(y_n - \widehat{y}_{n|n-1})^2}{2\eta_{n|n-1}^2}\right\}, \tag{12}$$

where $\widehat{y}_{n|n-1}$ is the one-step ahead prediction for y_n and $\eta_{n|n-1}^2$ denotes the variance of the prediction error. $\widehat{y}_{n|n-1}$ and $\eta_{n|n-1}^2$ can be obtained as follows based on the results of the Kalman filter:

$$\begin{aligned} \widehat{y}_{n|n-1} &= \boldsymbol{H}\boldsymbol{x}_{n|n-1}, \\ \eta_{n|n-1}^2 &= \boldsymbol{H}\boldsymbol{V}_{n|n-1}\boldsymbol{H}^{\mathrm{t}} + R. \end{aligned}$$

By taking the logarithm of $L(\boldsymbol{\theta}|Y_{1:N})$ in equation (11), the log-likelihood is obtained as

$$\ell(\boldsymbol{\theta}|Y_{1:N}) = \log L(\boldsymbol{\theta}|Y_{1:N}) = \sum_{n=1}^{N} \log f_n(y_n|Y_{1:(n-1)}, \boldsymbol{\theta}). \tag{13}$$

Thus, generally, the estimates $\widehat{\boldsymbol{\theta}}$ of the parameters $\boldsymbol{\theta}$ can be obtained by maximising the log-likelihood $\ell(\boldsymbol{\theta}|Y_{1:N})$ in equation (13) together with equation (12) using a numerical method. Correspondingly, Akaike's information criterion (AIC) for the model can be calculated based on the definition

$$\mathrm{AIC}(q|Y_{1:N}) = -2\ell(\widehat{\boldsymbol{\theta}}|Y_{1:N}) + 2d(q),$$

where $d(q)$ denotes the dimension of $\boldsymbol{\theta}$. For the BSA model case, $d(q) = q + 4$, which depends on the value of the AR model order q. The minimum AIC method (see Akaike, 1974) allows the value of q to be determined by minimising the value of $\mathrm{AIC}(q|Y_{1:N})$ with respect to q.

Note that for the TCD model, the parameter τ_2^2 is removed, and so $d(q) = q + 4$. Moreover, for the special case of removing the cyclical component c_n from the BSA model (or the TCD model), q can be set to $q = 0$. In this case, the parameters $\boldsymbol{\theta}$ are composed of $\{\sigma^2, \tau_1^2, \tau_2^2\}$, and so the AIC value is given by $\mathrm{AIC}(0|Y_{1:N}) = -2\ell(\widehat{\boldsymbol{\theta}}|Y_{1:N}) + 2 \times 3$. When the value of $\mathrm{AIC}(0|Y_{1:N})$ is minimised among all values of $\mathrm{AIC}(q|Y_{1:N})$ for proper values of $q \geq 0$, it can be considered that $c_n = 0$; that is, the cyclical component cannot be identified from the viewpoint of the minimum AIC method.

2.4 Stationarity of the cyclical component

One approach for ensuring the stationarity of the cyclical component is to estimate the AR coefficients $\boldsymbol{\alpha} = \{\alpha_1, \alpha_2, \ldots, \alpha_q\}$ by controlling the partial autocorrelation (PARCOR) coefficients. For the AR model in equation (4), let $\rho_1, \rho_2, \ldots, \rho_q$ denote the PARCOR coefficients. The relation between the AR coefficients and the PARCOR coefficients is well known [see, e.g., Kitagawa, (2020), p.96]. Hence, the values of the AR coefficients can be controlled through the values of the PARCOR coefficients. Moreover, when the inequality $-1 < \rho_i < 1$ holds for $i = 1, 2, \ldots, q$, the AR model is stationary. Hence, under the setting

$$\rho_i = \Phi \frac{\exp(\xi_i) - 1}{\exp(\xi_i) + 1} \quad (i = 1, 2, \ldots .q), \tag{14}$$

for $\Phi > 0$, $-\Phi < \rho_i < \Phi$ is always true for $-\infty < \xi_i < \infty$. Note that equation (14) defines a set of one-to-one correspondences between $-\Phi < \rho_i < \Phi$ and $-\infty < \xi_i < \infty$ for $i = 1, 2, \ldots, q$.

Thus, a strategy that ensures the stationarity of the cyclical component is as follows: For any values of ξ_i $(i = 1, 2, \ldots, q)$, if $\Phi < 1$ is set, then it is possible to ensure that the AR model is stationary. Therefore, the AR coefficient can be estimated through the estimation of ξ_i $(i = 1, 2, \ldots, q)$ by setting a value of $\Phi < 1$. Hereafter, this is called the *stationarity threshold*.

3 Aim and requirements

3.1 Aim

As mentioned in Section 1, when a time series contains long-period cycles, it is difficult to capture the cyclical variation clearly using the BSA modelling method; hence, part of the cyclical variation may remain in the estimate of the trend component, and so the estimated trend may contain unnecessary fluctuations. Moreover, it is easy to imagine a scenario in which one wants to decompose a trend-cyclical time series into trend and cyclical components, but existing methods cannot provide satisfactory results.

Thus, our main aim is to develop an approach that can be applied to decompose a time series into several components, such as trend and cyclical components, in cases where the data contain long-period cycles. In the majority of the present work, we develop a comprehensive tool for decomposing a trend-cyclical component into hyper-trend and hyper-cyclical components when existing methods do not work well.

Before proceeding, we consider several subtle but important issues. As fundamentals of the new approach, we provide several requirements for constructing models and evaluating the decomposed results in the following subsection.

3.2 Requirements

For the present study, we are confronted with the difficulty that there are no established definitions for trend and cyclical components. However, examples of these components are ubiquitous in applicative experience in many fields, so we can specify the following requirements for implementing trend-cycle decomposition.

3.2.1 Cyclical component

To obtain a unique decomposition, we assume that the cyclical component is stationary in mean, i.e., its average in a given data span is zero. As mentioned in Subsection 2.4, when we use an AR model to express the cyclical component, we can ensure stationarity by controlling the values of the PARCOR coefficients.

We construct a model for observations at equal sampling intervals, and consider the stationarity of a cyclical component for different sampling intervals. Specifically, for a time series $\{c_1, c_2, \ldots\}$ of the cyclical component, we can measure its stationarity using the PARCOR coefficients for the model:

$$c_n = \sum_{j=1}^{q} \alpha_j c_{n-jk} + v_{n3}(k), \tag{15}$$

where k is an integer. To extract long-period cycles from a time series, we focus on the stationarity for the case in which $k > 1$.

Besides the stationarity requirement, the following additional requirements ensure that the cyclical component has several desirable properties:

- *Uniformity of dispersion*, which means that its variation is almost uniform over the whole data span.
- *Symmetry*, which means that it should generally vary symmetrically with respect to the zero level.
- *Conditionally large variation*, which means that its variation should be as large as possible so that a large part of the variation in the time series data can be absorbed by the cyclical component and the trend component becomes sufficiently smooth. Note that the word *conditionally* means that this requirement is predicated on the uniformity of dispersion and the symmetry requirements being satisfied.

As will be mentioned in the next subsection, the above additional requirements correspond to the requirements for the trend component.

3.2.2 Trend component

Regarding the properties of the trend component, Wu et al. (2007) provided the following definite requirements: The trend should reflect the intrinsic property of the time series. It is an integral part of the data and can be driven by the same mechanisms or part of the same mechanisms that generate the data. Hence, it requires that the method used in defining the trend is adaptive so that the trend extracted is derived from and based on the data.

We essentially agree with the viewpoint of Wu et al. (2007). In particular, we expect the trend to be smooth over time – if the trend exhibits a jagged variation, this should be assimilated by the cyclical or irregular components. In terms of smoothness, we recommend using the smoothness priors in equation (2). Corresponding to the consideration regarding the stationarity of the cyclical component, for a time series $\{t_1, t_2, \ldots\}$ of the trend component, we can also measure its smoothness using the variance of

$$v_{n1}(k) = t_n - 2t_{n-k} + t_{n-2k}, \tag{16}$$

where k is an integer (that is, the variance of $v_{n1}(k)$ expresses the smoothness of the trend component). When we expect the trend component to be smooth over a large interval, we can set the value of k to be relatively large. Note that, in our proposed approach, the correspondence between stationarity and smoothness means that the integer k in equation (16) is naturally the same as that in equation (15). The integer k is hereafter called the *length of the sampling interval* (LSI).

We also expect the trend to be adaptive to the time series, i.e., the trend component should fit long-term variations in the data. There may be a trade-off between the properties of smoothness and adaptivity. This trade-off relates to the properties of the cyclical component, i.e., if the corresponding cyclical component satisfies the uniformity of dispersion and the symmetry requirements, then the adaptivity of the trend component will be high, whereas when the large variation requirement is satisfied, the trend component will be smooth. Thus, the properties of the trend component can be evaluated in terms of whether they satisfy the requirements for the cyclical component.

3.3 Consideration about use of the requirements

Ideally, all the requirements should be considered in the process of modelling; however, this may lead to the model structure becoming very complicated and increase the cost of parameter estimation. In the present paper, we focus on constructing models so as to achieve a smooth trend component and ensure the stationarity of the cyclical component. The other requirements can be taken into consideration by evaluating the properties of the cyclical component after the process of modelling and estimation.

To obtain good estimates, we are required to make as many sets of estimates as possible so that a good set can be selected and highly satisfactory results can be obtained. This stimulates us to develop a systematic program based on several newly proposed modelling methods, and then to establish a criterion for evaluating the decomposed results.

4 Proposed approach

We develop a systematic program as a general tool for trend-cycle decomposition or for cases in which the BSA modelling method and other existing techniques do not work well. In Subsection 4.1, we construct several models for trend-cycle decomposition with estimation methods. In Subsection 4.2, we propose a criterion for determining the LSI. In Subsection 4.3, we summarise the overall procedure.

4.1 Models and estimation methods

4.1.1 Interval averaging trend-cycle decomposition model

We denote the estimates of the trend and cyclical components that are obtained based on the BSA model or the TCD model by $\widehat{t}_n^{\mathrm{F}}$ and $\widehat{c}_n^{\mathrm{F}}$; hereafter, these are referred to as the *first-stage estimates*.

For cases in which long-period cycles cannot be captured using the BSA model or the TCD model, we consider a model based on the smoothness of the trend component and the stationarity of the cyclical component for a large value of the LSI. Let $z_n = \widehat{t}_n^{\mathrm{F}}$.

Then, we denote the set of results for the first-stage estimate of the trend by $Z_{1:N} = \{z_1, z_2, \ldots, z_N\}$. Consider the possibility of decomposing the time series z_n into a trend component $\widetilde{t}_n$, a cyclical component $\widetilde{c}_n$, and an irregular component $\widetilde{w}_n$; that is, the structure of the time series z_n can be expressed as

$$z_n = \widetilde{t}_n + \widetilde{c}_n + \widetilde{w}_n. \tag{17}$$

We call $\widetilde{t}_n$ in equation (17) the *hyper-trend component* because it expresses a trend component in the first-stage estimate z_n of the trend component. Correspondingly, we call $\widetilde{c}_n$ the *hyper-cyclical component*.

Consider the smoothness of the trend $\widetilde{t}_n$ and the stationarity of $\widetilde{c}_n$ for $i \in \{1, 2, \ldots, k\}$ with a given value of $k \geq 2$. We construct a model for observations at equal sampling intervals. Specifically, we express the model as

$$z_m = \widetilde{t}_m + \widetilde{c}_m + \widetilde{w}_m, \tag{18}$$

$$\widetilde{t}_m = 2\widetilde{t}_{m-k} - \widetilde{t}_{m-2k} + \widetilde{v}_{m1}, \tag{19}$$

$$\widetilde{c}_m = \sum_{j=1}^{q} \alpha_j \widetilde{c}_{m-jk} + \widetilde{v}_{m3}, \tag{20}$$

where $m = i, i+k, i+2k, \ldots$, and the symbols used in equations (18)–(20) correspond to those used in the BSA model. Note that the integer k is the same as the LSI which appeared in equations (15) and (16).

Using the averages for every quantity in equations (18)–(20), we have

$$\bar{z}_m = \bar{t}_m + \bar{c}_m + \bar{w}_m, \quad \bar{w}_m \sim \mathrm{N}(0, \widetilde{\sigma}^2), \tag{21}$$

$$\bar{t}_m = 2\bar{t}_{m-k} - \bar{t}_{m-2k} + \bar{v}_{m1}, \quad \bar{v}_{m1} \sim \mathrm{N}(0, \widetilde{\tau}_1^2), \tag{22}$$

$$\bar{c}_m = \sum_{j=1}^{q} \alpha_j \bar{c}_{m-jk} + \bar{v}_{m3}, \quad \bar{v}_{m3} \sim \mathrm{N}(0, \widetilde{\tau}_3^2), \tag{23}$$

where $\bar{z}_m$ is the interval average for z_m up to time point $m-k+1$, that is,

$$\bar{z}_m = \frac{1}{k}(z_m + z_{m-1} + \cdots + z_{m-k+1}),$$

and the other quantities are similar interval averages that correspond to $\bar{z}_m$, e.g.,

$$\bar{t}_m = \frac{1}{k}(\widetilde{t}_m + \widetilde{t}_{m-1} + \cdots + \widetilde{t}_{m-k+1}), \tag{24}$$

and so on. If the LSI (k) is determined appropriately, then it is possible to capture long-period cycles. We call the model in equations (21)–(23) the *interval averaging trend-cycle decomposition* (IATCD) model.

Here, we compare the structure of the IATCD model with that of the TCD model. For $k > 1$, the trend model used in the IATCD model [as in equation (22)] controls the smoothness of the trend component in a wider interval than that used in the TCD model [as in equation (2)]. Correspondingly, the cyclical component model in the IATCD model [as in equation (23)] expresses the cyclical variation in a wider interval than that used in the TCD model [as in equation (4)]. Thus, using the IATCD model, we can extract long-term cycles and estimate the trend with a relatively high degree of smoothness.

However, the IATCD model has the same mathematical form as the TCD model. Thus, given a value of $i = 1, 2, \dots, k$ with a fixed value of $k \geq 2$, the IATCD model has the same state space form as the TCD model under the following settings.

$$\begin{aligned} y_\ell &= \bar{z}_{\ell k+i-1}, \quad t_\ell = \bar{t}_{\ell k+i-1}, \quad c_\ell = \bar{c}_{\ell k+i-1}, \\ w_\ell &= \bar{w}_{\ell k+i-1}, \quad \boldsymbol{v}_\ell = (\bar{v}_{(\ell k+i-1)1}, \bar{v}_{(\ell k+i-1)3})^{\mathrm{t}} \quad (\ell = 1, 2, \dots, N_i), \end{aligned}$$

with the relations $\boldsymbol{Q} = \mathrm{diag}(\widetilde{\tau}_1^2, \widetilde{\tau}_3^2)$ and $R = \widetilde{\sigma}^2$, where N_i denotes the integral part of $\frac{N-i+1}{k}$.

We assume that the parameters $\boldsymbol{\theta} = \{\boldsymbol{\alpha}, \widetilde{\sigma}^2, \widetilde{\tau}_1^2, \widetilde{\tau}_3^2\}$ are constants for all values of $i = 1, 2, \dots, k$, where $\boldsymbol{\alpha} = \{\alpha_1, \alpha_2, \dots, \alpha_q\}$. Let

$$\bar{Z}_{1:N_i}^{(i)} = \{\bar{z}_{k+i-1}, \bar{z}_{2k+i-1}, \dots, \bar{z}_{N_i k+i-1}\}$$

be a set of the interval average values (for the first-stage estimate of the trend component) under a given value of $i = 1, 2, \dots, k$ with a fixed value of $k \geq 2$. Then, we can calculate the likelihood for $\bar{Z}_{1:N_i}^{(i)}$ using

$$L(\boldsymbol{\theta}|\bar{Z}_{1:N_i}^{(i)}) = \prod_{\ell=1}^{N_i} f_\ell(\bar{z}_{\ell k-i+1}|\bar{Z}_{1:(\ell-1)}^{(i)}, \boldsymbol{\theta}),$$

with $\bar{Z}_{1:(\ell-1)}^{(i)} = \{\bar{z}_{k+i-1}, \bar{z}_{2k+i-1}, \dots, \bar{z}_{(\ell-1)k+i-1}\}$ and where $f_\ell(\bar{z}_{\ell k-i+1}|\bar{Z}_{1:(\ell-1)}^{(i)}, \boldsymbol{\theta})$ is the density function, which can be defined similarly to equation (12). Furthermore, based on the Bayesian model averaging approach (see Hoeting et al., 1999), by averaging the likelihood $L(\boldsymbol{\theta}|\bar{Z}_{1:N_i}^{(i)})$ on $i = 1, 2, \dots, k$, we calculate

$$\bar{L}(\boldsymbol{\theta}) = \frac{1}{k} \sum_{i=1}^{k} L(\boldsymbol{\theta}|\bar{Z}_{1:N_i}^{(i)}) \tag{25}$$

using a uniform prior, and obtain the estimates $\widehat{\boldsymbol{\theta}}$ for the parameters $\boldsymbol{\theta}$ by maximising the averaging likelihood $\bar{L}(\boldsymbol{\theta})$ in equation (25). Thus, we define the AIC for the IATCD model as

$$\mathrm{AIC}(q) = -2 \log \bar{L}(\widehat{\boldsymbol{\theta}}) + 2(q+3); \tag{26}$$

hence, we can determine the AR model order q by minimising the value of $\mathrm{AIC}(q)$ defined by equation (26). As mentioned in Subsection 4.1.1, when the inequality $\mathrm{AIC}(0) < \mathrm{AIC}(q)$ holds for all possible values of $q > 0$, we can set $\bar{c}_{\ell k+i-1} = 0$ for $\ell = 1, 2, \dots, N_i, i = 1, 2, \dots, k$, which implies that we can omit the long-period cyclical variation from the model. Note that we execute the strategy for the stationarity of $\bar{c}_{\ell k+i-1}$ in cases where $q > 0$. Finally, we estimate $\bar{t}_{\ell k+i-1}$ and $\bar{c}_{\ell k+i-1}$ using the Kalman filter and fixed-interval smoothing algorithms for $\ell = 1, 2, \dots, N_i$ under a given value of $i = 1, 2, \dots, k$ with a fixed value of $k \geq 2$. Let $\widehat{t}^*_{\ell k+i-1}$ denote the estimate of $\bar{t}_{\ell k+i-1}$. Then, we can obtain a set of estimates as

$$\widehat{T}^{(i)}_{1:N_i} = \{\widehat{t}^*_{k+i-1}, \widehat{t}^*_{2k+i-1}, \dots, \widehat{t}^*_{N_i k+i-1}\} \tag{27}$$

for a given value of $i = 1, 2, \dots, k$ with a fixed value of k.

Note that the AIC values for models with different values of the LSI (k) cannot be compared with each other because the initial conditions of the state vector for models with different values of k are very different; hence, it is difficult to determine the value of k using the minimum AIC method. We introduce a criterion for determining the value of the LSI (k) later.

4.1.2 Hyper-trend reconstruction model

Continuing the process of estimating the hyper-trend component, we propose a method for reconstructing the hyper-trend based on the estimates of the interval average (for the hyper-trend component) expressed by equation (27).

First, to avoid complexity in the expression, let u_n denote a time series for the estimate of the interval average $\bar{t}_n$, where n is the time index, which is equivalent with that in the original time series y_n. From equation (24), we have

$$\bar{t}_n = \frac{1}{k}(\widetilde{t}_n + \widetilde{t}_{n+1} + \cdots + \widetilde{t}_{n+k-1}).$$

Thus, to reconstruct the hyper-trend component, we construct a model as follows:

$$u_n = \bar{t} + \epsilon_n = \frac{1}{k}(\widetilde{t}_n + \widetilde{t}_{n+1} + \cdots + \widetilde{t}_{n+k-1}) + \epsilon_n, \tag{28}$$

where $\widetilde{t}_n$ is the hyper-trend in equation (17) and $\epsilon_n \sim \mathrm{N}(0, \psi^2)$ is the residual. Because the smoothness with $k > 1$ expressed by equation (19) has been applied in estimating the interval averaging trend $\bar{t}_m$, as a prior which expresses the smoothness with $k = 1$ for $\widetilde{t}_n$, we use the model

$$\widetilde{t}_n = 2\widetilde{t}_{n-1} - \widetilde{t}_{n-2} + \widetilde{v}_n, \quad \widetilde{v}_n \sim \mathrm{N}(0, \widetilde{\tau}^2). \tag{29}$$

We call the model in equations (28) and (29) the *hyper-trend reconstruction* (HTR) model.

Now, we define a k-dimensional state vector

$$\boldsymbol{x}_n = (\widetilde{t}_n, \widetilde{t}_{n-1}, \dots, \widetilde{t}_{n-k+1})^{\mathtt{t}}$$

and correspondingly set the matrices

$$\boldsymbol{F} = \begin{bmatrix} 2 & 1 & 0 & \cdots & 0 \\ 1 & 0 & \vdots & \ddots & \vdots \\ 0 & \ddots & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix}, \quad \boldsymbol{G} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{bmatrix}, \quad \boldsymbol{H} = \begin{bmatrix} \frac{1}{k} \\ \vdots \\ \vdots \\ \vdots \\ \frac{1}{k} \end{bmatrix}^{\mathtt{t}}.$$

We estimate the hyper-trend component based on the set $\widehat{T}^{(i)}_{1:N_i}$ defined in equation (27) for a given value of $i = 1, 2, \ldots, k$ with a fixed value of $k \geq 2$. For a value of $n = 1, 2, \ldots, N$, if there is an element, say $\widehat{t}^{*}_{m}$, in the set $\widehat{T}^{(i)}_{1:N_i}$ such that the equality $n = m$ holds, then we set $y_n = u_n = \widehat{t}^{*}_{m}$; otherwise, we regard $y_n = u_n$ as a missing value. Then, we can express the state space form of the HTR model using equations (9) and (10) based on the assumption that $\boldsymbol{v}_n = (\widetilde{v}_n) \sim \mathrm{N}(\mathbf{0}, \boldsymbol{Q})$, $w_n = \epsilon_n \sim \mathrm{N}(0, R)$ under the settings $\boldsymbol{Q} = (\widetilde{\tau}^2)$ and $R = \psi^2$. Moreover, if we set $\boldsymbol{\theta} = \{\psi^2, \widetilde{\tau}^2\}$, then we can apply the Kalman filter and calculate the log-likelihood $\ell(\boldsymbol{\theta}|\widehat{T}^{(i)}_{1:N_i})$ using equations (11)–(13). Hence, we obtain the estimates for the parameters $\boldsymbol{\theta}$ by maximising $\ell(\boldsymbol{\theta}|\widehat{T}^{(i)}_{1:N_i})$.

Finally, for a value of $i = 1, 2, \ldots, k$ with a fixed value of $k > 1$, we obtain the estimate of the state vector using the Kalman filter and fixed-interval smoothing algorithms; hence, the estimate $\widehat{t}^{\mathrm{S}}_{in}$ for the hyper-trend $\widetilde{t}_n$ can be obtained from the estimate of the state vector. Thus, we can obtain k sets of estimates for the hyper-trend component as $\{\widehat{t}^{\mathrm{S}}_{in}; n = 1, 2, \ldots, N\}$ for $i = 1, 2, \ldots, k$. Then, by taking the average on $i = 1, 2, \ldots, k$, we obtain

$$\widehat{t}^{\mathrm{S}}_{n} = \frac{1}{k}\sum_{i=1}^{k} \widehat{t}^{\mathrm{S}}_{in} \quad (n = 1, 2 \ldots, N),$$

and take $\widehat{t}^{\mathrm{S}}_{n}$ as the second-stage estimate of the trend component. Furthermore, $z_n - \widehat{t}^{\mathrm{S}}_{n}$ can be considered as an additional part of the estimate of the cyclical component. Hence, the second-stage estimates of the cyclical component are obtained as follows:

$$\widehat{c}^{\mathrm{S}}_{n} = \widehat{c}^{\mathrm{F}}_{n} + z_n - \widehat{t}^{\mathrm{S}}_{n} = \widehat{c}^{\mathrm{F}}_{n} + \widehat{t}^{\mathrm{F}}_{n} - \widehat{t}^{\mathrm{S}}_{n} \quad (n = 1, 2, \ldots, N). \tag{30}$$

4.2 Criterion for determining the LSI

The core of the proposed approach is the estimation method based on the IATCD and HTR models. However, to use these modelling methods, we must determine the value of the LSI (k). For this, we need to establish a suitable criterion. To illustrate the image of the criterion which will be constructed below, we present an example in which we decompose a trend-cycle series into trend and cyclical components with different values of k.

Figure 1 shows a time series for a set of example data with the data span $N = 492$. From this figure, we can see that the time series may be difficult to decompose because it is too uneven; in particular, the data reach a maximum level early in the series, then fall sharply to a low level around time point 200.

Figure 2 shows the time series of the second-stage estimates for the trend and cyclical components with $k = 2, 3, 4$. We can obtain some immediate conclusions from these results. For $k = 2$, a large part of the cyclical fluctuations remains in the estimated trend. When $k = 3$, the estimated results agree well with the requirements for the cyclical and trend components. For $k = 4$, the estimated trend is hardened, and to fit it to the data, the estimate of the cyclical component changes sharply at around time point 200.

Figure 1 Time series for example data

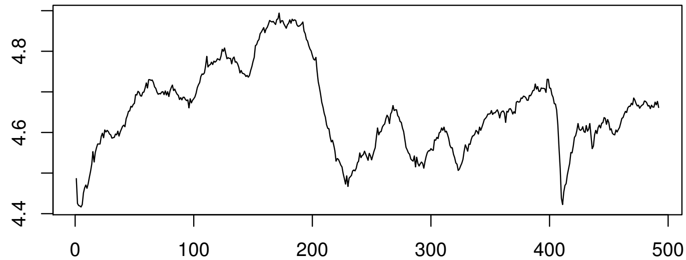
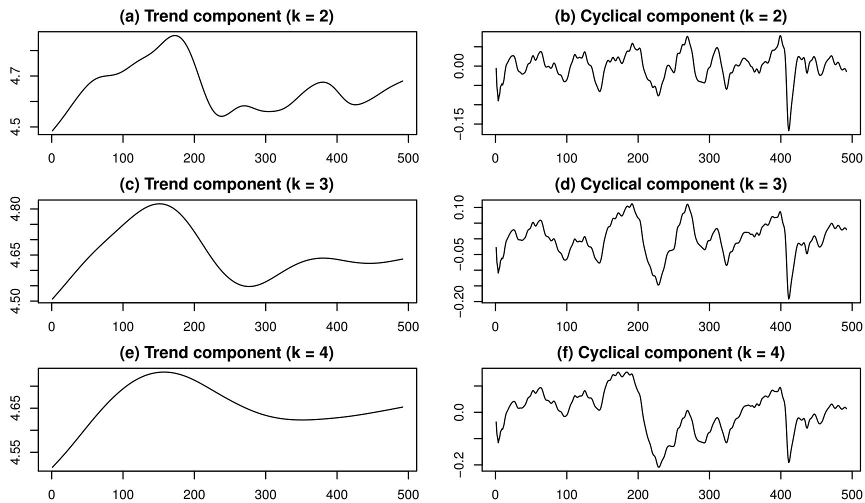


Figure 2 Second-stage estimates of trend and cyclical components



We now introduce a criterion for evaluating a set of decomposed results based on the additional requirements for the cyclical component mentioned in Subsection 3.2.1. We consider the time series for the second-stage estimate $\widehat{c}_n^{\mathrm{S}}(n = 1, 2 \ldots, N)$ as the objective, and simplify the notation by writing $\widehat{c}_n = \widehat{c}_n^{\mathrm{S}}$. We assume that no element in the set $\widehat{C}_{1:N} = \{\widehat{c}_n; n = 1, 2 \ldots, N\}$ is equal to zero, so we can divide $\widehat{C}_{1:N}$ into a subset of negative parts, in which all the elements are negative, and a subset of positive parts, in which all the elements are positive. Let the numbers of elements in these subsets be N_1 and N_2, respectively, so that $N = N_1 + N_2$. Denote the subset of negative parts by $\{\widehat{c}_i^-; i = 1, 2 \ldots, N_1\}$, and the subset of positive parts by $\{\widehat{c}_i^+; i = 1, 2 \ldots, N_2\}$, in which $\{\widehat{c}_1^-, \widehat{c}_2^-, \ldots, \widehat{c}_j^-\}$ are the first j negative elements in $\widehat{C}_{1:N}$ for $j = 1, 2, \ldots, N_1$, and $\{\widehat{c}_1^+, \widehat{c}_2^+, \ldots, \widehat{c}_j^+\}$ are the first j positive elements in $\widehat{C}_{1:N}$ for $j = 1, 2, \ldots, N_2$. Then, we generate the cumulative values of the negative part (CVNP) as

$$a_i = a_{i-1} + \widehat{c}_i^- \quad (i = 1, 2, \ldots, N_1),$$

and the cumulative values of the positive part (CVPP) as

$$b_i = b_{i-1} + \widehat{c}_i^+ \quad (i = 1, 2, \ldots, N_2)$$

with $a_0^+ = b_0^- = 0$.

For an estimate of the cyclical component, when the uniformity of dispersion requirement is satisfied, the series a_i decreases uniformly with i. Hence, the model

$$a_i = -\beta_1 i + e_{i1} \quad (i = 1, 2, \ldots, N_1) \tag{31}$$

can be constructed with a positive parameter β_1 and a small variance of the error term e_{i1}. Similarly, under the same condition, the series b_i increases uniformly with i, so the model

$$b_i = \beta_2 i + e_{i2} \quad (i = 1, 2, \ldots, N_2) \tag{32}$$

can be built with a positive parameter β_2 and a small variance of the error term e_{i2}. Furthermore, if the symmetry requirement is satisfied, then the values of $|\beta_1 - \beta_2|$ and $|N_1 - N_2|$ may be small, and so we can set $\beta_1 = \beta_2 = \beta$ with a large degree of confidence. Thus, the models in equations (31) and (32) can be expressed in a regression model as follows:

$$\begin{aligned} a_i &= -\beta i + e_{i1} \quad (i = 1, 2, \ldots, N_1), \\ b_i &= \beta i + e_{i2} \quad (i = 1, 2, \ldots, N_2). \end{aligned} \tag{33}$$

Moreover, when the conditionally large variation requirement is satisfied, the variances of a_i and b_i may be large relative to those of the error terms e_{i1} and e_{i2}. Thus, the correlation between the series $\{a_{N_1}, a_{N_1-1}, \ldots, a_1, b_1, b_2, \ldots, b_{N_2}\}$ and the series $\{-N_1, -(N_1 - 1), \ldots, -1, 1, 2, \ldots, N_2\}$ may be high, resulting in a large coefficient of determination (R^2) in the model of equation (33). That is, R^2 can be taken as an indicator for measuring the degree to which the additional requirements are satisfied. However, the value of R^2 may depend on the order of the elements in the set $\widehat{C}_{1:N}$, which determines the orders of the elements in the sets $\{\widehat{c}_i^-; i = 1, 2 \ldots, N_1\}$ and $\{\widehat{c}_i^+; i = 1, 2 \ldots, N_2\}$.

Thus, we reset the order of the elements in $\widehat{C}_{1:N}$. For example, initially, we reset $\widehat{C}_{1:N}$ as $\{\widehat{c}_N, \widehat{c}_1, \widehat{c}_2, \ldots, \widehat{c}_{N-1}\}$, and next, we reset $\widehat{C}_{1:N}$ as $\{\widehat{c}_{N-1}, \widehat{c}_N, \widehat{c}_1, \ldots, \widehat{c}_{N-2}\}$, and so on, so that we obtain N different sequences for the elements in $\widehat{C}_{1:N}$ including $\widehat{C}_{1:N}$ itself. For each resetting of $\widehat{C}_{1:N}$, we have a regression model given by equation (33). Let $R_j^2(k)$ be the value of R^2 for the j-th resetting with a given LSI value k. Then, we can define an average coefficient of determination (ACD) as

$$\mathrm{ACD}(k) = \frac{1}{N} \sum_{j=1}^{N} R_j^2(k). \tag{34}$$

Thus, the value of ACD(k) can be regarded as a function of the LSI (k); hence, it can be applied as a criterion for evaluating the estimation results and used to determine the value of k — a larger value of ACD(k) indicates better estimation results under the given value of the LSI (k).

We now demonstrate the performance of the ACD criterion for the example data in Figure 1. For $k = 2, 3, 4$, the values of the ACD are 0.9743, 0.9817, and 0.9433, respectively. Consequently, among these three cases, the ACD is maximised when $k =$

3, which is mainly because of the properties of the estimate for the cyclical component. Incidentally, the variances for the estimates of the cyclical components are 14.31, 39.46, and 83.80 and the average values of the coefficient β are 0.0258, 0.0436, and 0.0608 for $k = 2, 3, 4$, respectively.

4.3 Overall procedure

Based on the results obtained using the BSA model or the TCD model (which is a special case of the BSA model), we have established the IATCD and HTR models. For the proposed modelling methods, we assume that the objective data are time series that contain trend and cyclical components, and possibly a seasonal component. We call data that contain a seasonal component a *type-A time series*, and refer to data that do not contain a seasonal component as a *type-B time series*. Summarising the proposed modelling methods, we obtain the following overall procedure.

- First stage: original data decomposition

  For type-A and type-B time series, we use the BSA and TCD models, respectively, to obtain estimates of the components. We consider the possibility that, at this point, the first-stage estimate of the trend component contains part of the cyclical component. Thus, if the first-stage estimate of the trend component is monotonic or has only one extremum, we adopt the first-stage estimates and terminate the procedure; hence, we say that second-stage estimation is not necessary. Otherwise, we proceed to the second stage.

- Second stage: hyper-trend estimation

  We provide an appropriate upper limit K of the global value k. For each value of $k \in \{2, 3, \ldots, K\}$, we use the IATCD and HTR modelling methods to estimate the hyper-trend, and then obtain the second-stage estimates for the trend and cyclical components.

- Third stage: determining the value of the LSI (k)

  Based on the second-stage estimates, we compute the value of $\mathrm{ACD}(k)$ defined by equation (34) for $k \in \{2, 3, \ldots, K\}$. Additionally, we define $\widehat{k}$ as

  $$\mathrm{ACD}(\widehat{k}) = \mathrm{Max}\{\mathrm{ACD}(k); k = 2, 3, \ldots, K\}.$$

  Furthermore, as a reference, we compute $\mathrm{ACD}(1)$ based on the first-stage estimates. When

  $$\mathrm{ACD}(\widehat{k}) > \mathrm{ACD}(1), \tag{35}$$

  we consider $\widehat{k}$ as an estimate of the LSI for $k > 1$ and adopt the second-stage estimates corresponding to the value of $\widehat{k}$ as the final estimates. When equation (35) holds, we say that second-stage estimation is necessary. Otherwise, we adopt the first-stage estimates and consider the second-stage estimates to be unnecessary.

The proposed approach for seasonal adjustment and trend-cycle decomposition based on the above procedure is called the *hyper-trend method* because the hyper-trend estimation stage is the core of this approach.

5 Examples

5.1 Analysing time series of commercial sales in Japan

First, we applied the proposed approach to the seasonal adjustment of time series data related to commercial sales in Japan. The data were obtained from the website of the Ministry of Economy, Trade and Industry of Japan (https://www.meti.go.jp/statistics/tyo/syoudou/result2/index.html). We consider such data as the objective of the analysis because these time series may be influenced by business cycles and many other factors; hence, they may contain rich information about cyclical variation.

The data were publicised as monthly data with a seasonal component (i.e., type-A time series). To examine the performance of the proposed approach, we analysed all the series available on the website, except for those that contained missing values. We consider the 20 series listed in Table 1 as the objective for analysis.

Table 1 Analysed time series of commercial sales in Japan

Number	*Name of indicator*
S1	Total
S2	Wholesale
S3	General merchandise (in wholesale)
S4	Textiles
S5	Apparel and accessories
S6	Farm and aquatic products
S7	Food and beverages (in wholesale)
S8	Building materials
S9	Chemicals
S10	Minerals and metals
S11	Machinery and equipment (in wholesale)
S12	Furniture and house furnishings
S13	Medicines and toiletries
S14	Others in wholesale
S15	Retail
S16	General merchandise (in retail)
S17	Fabrics apparel and accessories
S18	Food and beverages (in retail)
S19	Motor vehicles
S20	Machinery equipment (in retail)

For each series, the data covered the period from January 1980 to December 2019; that is, the data span was $N = 480$ months. Note that, prior to analysis, we transformed all the data by taking the logarithm. We set the value of the stationarity threshold Φ to 0.95, and the upper limit K for the value of the LSI (k) was 12.

The results indicate that series S1, S6, S7, S8, S9, S10, S12, S13, S14, S17, S18, S19, and S20 did not require second-stage estimation. For the other seven series, we adopted the second-stage estimates; that is, for 35% of the time series, we used the proposed approach. Table 2 summarises the estimations of the series for which we used the proposed approach. Note that the parameter q represents the AR model order in the HTR model.

Table 2 Summary of estimations for the commercial sales data in Japan

Number	$\widehat{k}$*-value*	*ACD(1)*	*ACD(k)*	*q-value*
S2	$\widehat{k} = 4$	0.9797	0.9877	$q = 6$
S3	$\widehat{k} = 11$	0.9804	0.9866	$q = 4$
S4	$\widehat{k} = 6$	0.9861	0.9864	$q = 0$
S5	$\widehat{k} = 3$	0.9837	0.9874	$q = 5$
S11	$\widehat{k} = 6$	0.9815	0.9865	$q = 4$
S15	$\widehat{k} = 3$	0.9808	0.9909	$q = 8$
S16	$\widehat{k} = 10$	0.9898	0.9913	$q = 0$

We now discuss the estimation results of several typical series among those that required second-stage estimates. For series S2, S5, S11, S15, and S16, the estimation results are shown in Figures 3–7. In each figure, Figures 3(a)–7(a) show the logarithm of the data, Figures 3(b)–7(b) and 3(c)–7(c) are the first-stage estimates of the cyclical and trend components, respectively, Figures 3(d)–7(d) show the estimate of the hyper-cyclical component (the most important part), and Figures 3(e)–7(e) and Figures 3(f)–7(f) show the second-stage estimates of the cyclical and trend components, respectively.

Figure 3 Data and estimation results for series S2

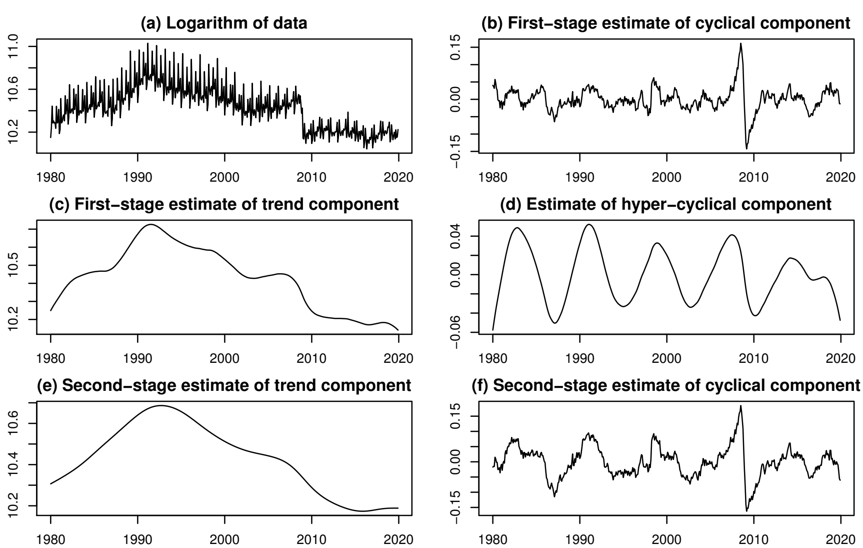


Figure 4 Data and estimation results for series S5

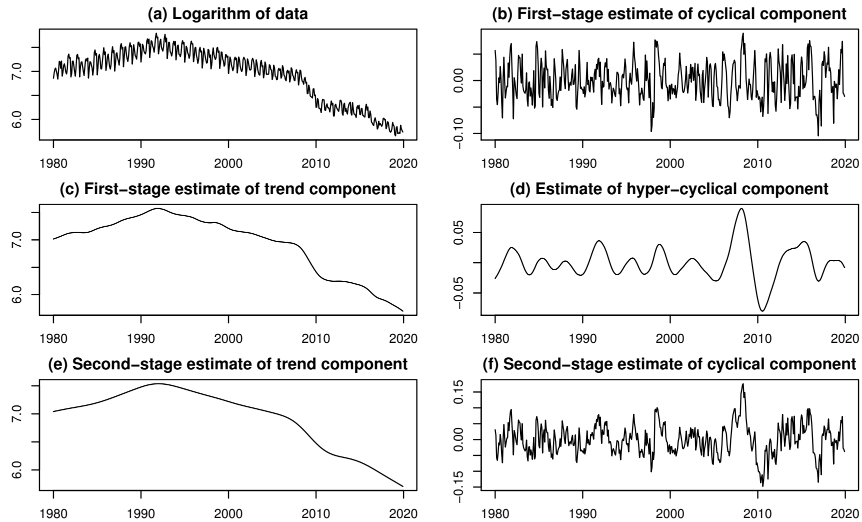
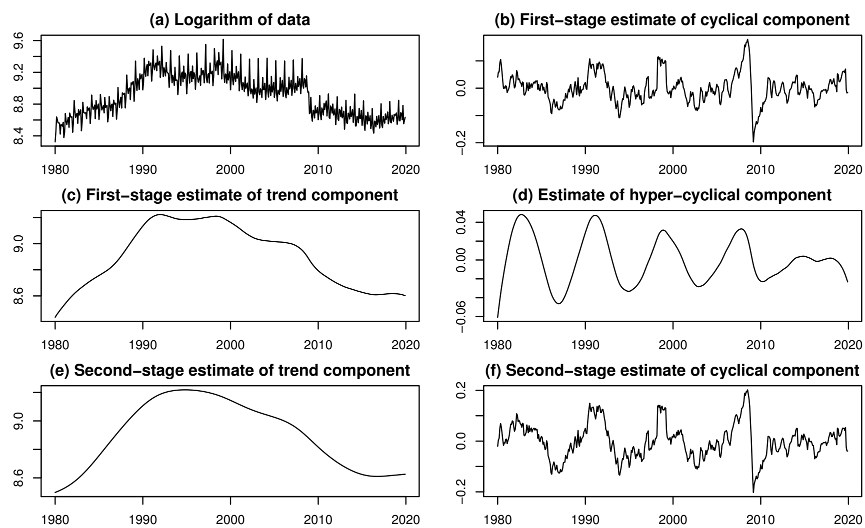


Figure 5 Data and estimation results for series S11



From Figures 3–7, we can draw the following conclusions. For almost all results, the second-stage estimates of the trends vary over time with a very high degree of smoothness and generally adapt to the long-term variation of the data. That is, the cyclical variation is clearly captured by the second-stage estimates of the cyclical component. Thus, the proposed approach performed very well. Note that sudden decreases appear in almost all estimates of the cyclical components. These reflect the influence of the Lehman shock from around 2008 to 2010.

Figure 6 Data and estimation results for series S15

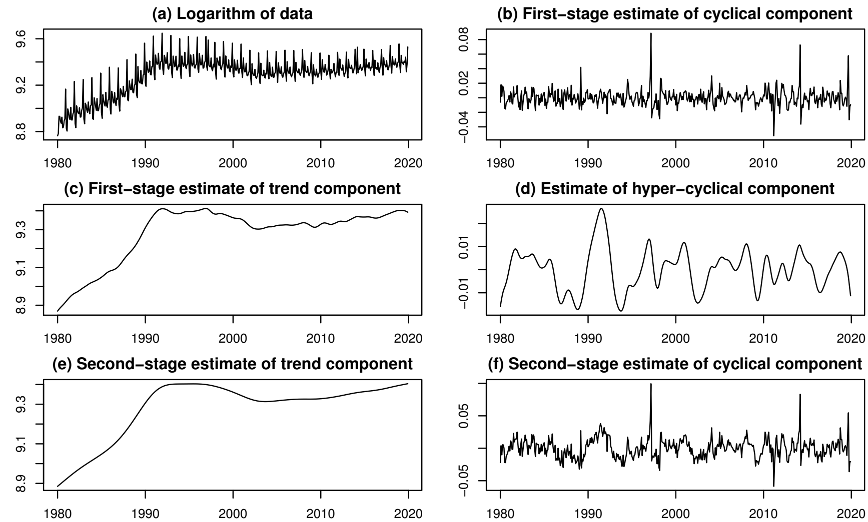
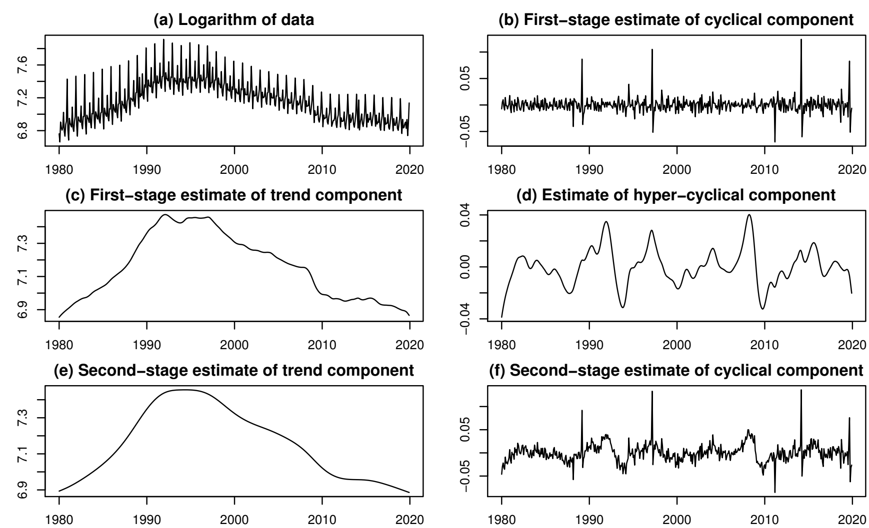


Figure 7 Data and estimation results for series S16



5.2 Analysing the time series of business cycle indicators in Japan

Next, we applied the proposed approach to the decomposition of time series data related to business cycle indicators in Japan. We obtained the data from the Japanese Cabinet Office website (https://www.esri.cao.go.jp/jp/stat/di/di.html). We considered these data because they are typical time series for economic analysis and are organised properly over a long period.

Table 3 Leading indicators for business cycle analysis in Japan

Number	*Name of indicator*
L1	Index of producer's inventory ratio of finished goods (final demand goods)
L2	Index of producer's inventory ratio of finished goods (producer goods for mining and manufacturing)
L3	New job offers (excluding new school graduates)
L4	Machinery orders at constant prices (manufacturing)
L5	Total floor area of new housing construction started
L6	Consumer confidence index
L7	Nikkei commodity price index (42 items)
L8	Money stock (M2, change from previous year)
L9	Stock prices (TOPIX)
L10	Index of investment climate (manufacturing)
L11	Sales forecast DI of small businesses

The time series of business cycle indicators were publicised as seasonally adjusted monthly data (i.e., type-B time series). We analysed all the series that were used for business cycle analysis in Japan. For each series, the data covered the period from January 1975 to December 2019; hence, the data span was $N = 540$ months. As in the previous subsection, we set the value of the stationarity threshold Φ to 0.95. The upper limit K for the value of the LSI (k) was set to 10. Moreover, the indicators for business cycle analysis were classified as either leading indicators, coincident indicators, or lagging indicators.

In total, there were 11 leading indicators; these are listed in Table 3. We transformed the data by taking the logarithm, except for series L8, L10, and L11, which contained some negative values. Based on the properties of the first-stage estimates and the values of the ACD, we determined that series L3, L4, L5, L8, and L11 did not require second-stage estimation. Hence, we applied the proposed approach to series L1, L2, L6, L7, L9, and L10, for which we adopted the second-stage estimates. Table 4 summarises these second-stage estimations.

Table 4 Summary of estimations for several leading indicators

Number	$\widehat{k}$*-value*	*ACD(1)*	*ACD(k)*	*q-value*
L1	$\widehat{k} = 8$	0.9740	0.9805	$q = 6$
L2	$\widehat{k} = 10$	0.9747	0.9794	$q = 7$
L6	$\widehat{k} = 2$	0.9755	0.9762	$q = 8$
L7	$\widehat{k} = 5$	0.9872	0.9885	$q = 0$
L9	$\widehat{k} = 5$	0.9815	0.9850	$q = 7$
L10	$\widehat{k} = 3$	0.9807	0.9845	$q = 7$

For three typical series among those for which second-stage estimates were required, the estimation results are shown in Figures 8–10. The results show that, for these series, the proposed approach worked well.

Figure 8 Data and estimation results for series L6

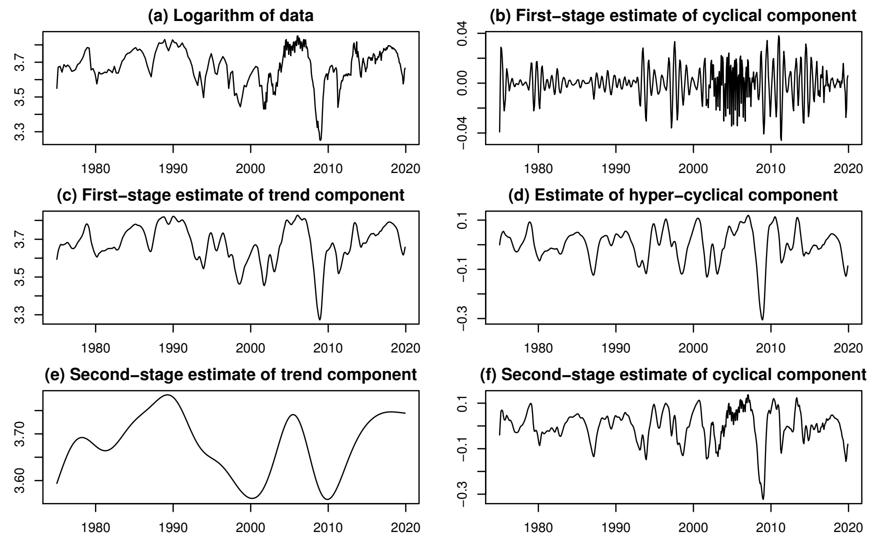
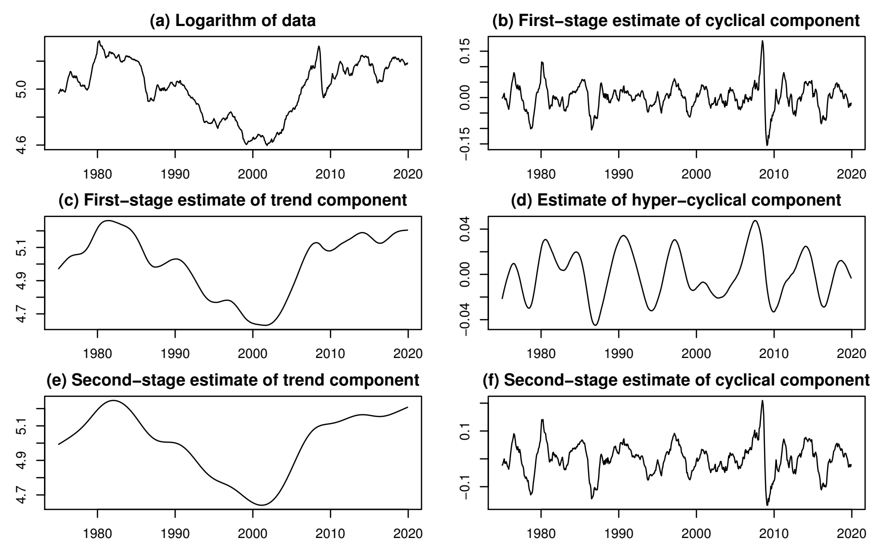


Figure 9 Data and estimation results for series L7



There were a total of ten coincident indicators; these are listed in Table 5. With the exception of series C6 and C7, which contained some negative values, we transformed the data for each series by taking the logarithm. Based on the properties of the first-stage estimates and the values of the ACD, we determined that series C1, C2, C4, C6, C7, C9, and C10 did not require second-stage estimation. Hence, we applied the proposed approach to series C3, C5, and C8, for which we adopted the second-stage estimates. Table 6 summarises the estimations for the series to which we applied the proposed approach.

Figure 10 Data and estimation results for series L9

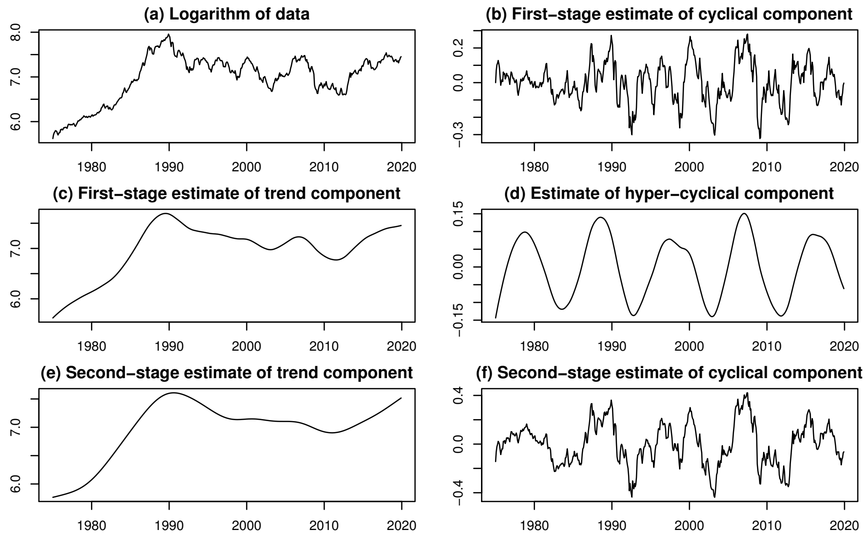


Table 5 Coincident indicators for business cycle analysis in Japan

Number	*Name of indicator*
C1	Index of industrial production (mining and manufacturing)
C2	Index of producers' shipments (producer goods for mining and manufacturing)
C3	Index of producers' shipments of durable consumer goods
C4	Index of labour input (industries covered)
C5	Index of producers' shipments (investment goods excluding transport equipments)
C6	Retail sales value (change from previous year)
C7	Wholesale sales value (change from previous year)
C8	Operating profits (all industries)
C9	Effective job offer rate (excluding new school graduates)
C10	Exports volume index

Table 6 Summary of estimations for several coincident indicators

Number	$\widehat{k}$*-value*	*ACD(1)*	*ACD(k)*	*q-value*
C3	$\widehat{k} = 4$	0.9762	0.9785	$q = 6$
C5	$\widehat{k} = 7$	0.0000	0.9847	$q = 6$
C8	$\widehat{k} = 10$	0.9599	0.9800	$q = 2$

The estimation results for series C5 and C8 are shown in Figures 11 and 12. From Figure 11, we can see that for series C5, the first-stage estimate of the cyclical component remained at the zero level [see Figure 11(b)] because we could not identify it using the minimum AIC method. Thus, some cyclical variation remained in the first-stage estimate of the trend component [see Figure 11(c)], and this was clearly captured using the second-stage estimation.

Figure 11 Data and estimation results for series C5

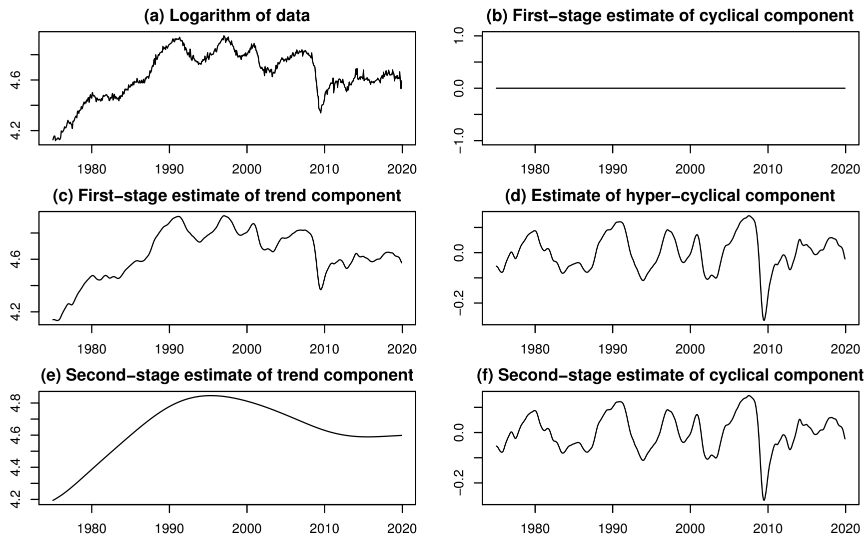
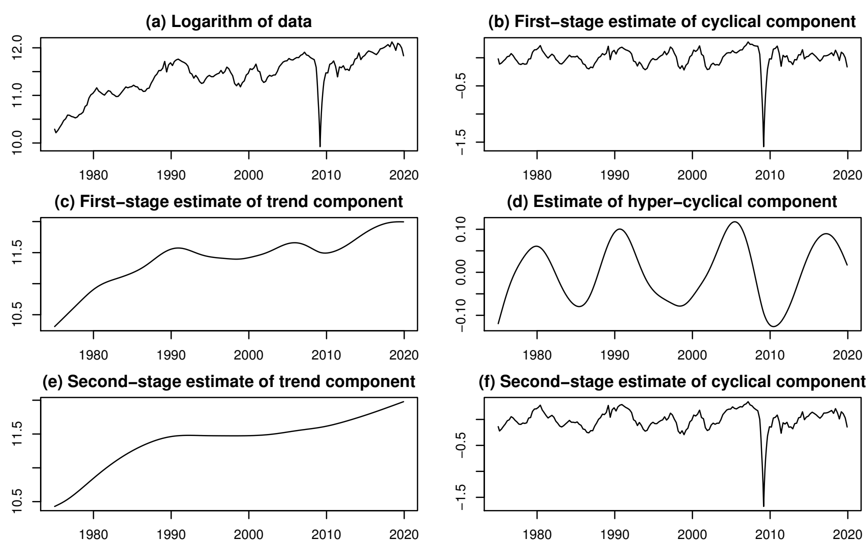


Figure 12 Data and estimation results for series C8



There were a total of nine lagging indicators; these are listed in Table 7. With the exception of series Lg2, Lg4, and Lg8, which contained some negative values, we transformed the data for each series by taking the logarithm. From the properties of the first-stage estimates and the values of the ACD, series Lg1, Lg3, Lg5, Lg6, and Lg8 did not require second-stage estimates. Hence, we applied the proposed approach to series Lg2, Lg4, Lg7, and Lg9, for which we adopted the second-stage estimates. The estimations for these series are summarised in Table 8.

Table 7 Lagging indicators for business cycle analysis in Japan

Number	*Name of indicator*
Lg1	Index of tertiary industry activity (business services)
Lg2	Index of regular workers employment (change from previous year)
Lg3	Business expenditures for new plant and equipment at constant prices (all industries)
Lg4	Living expenditure (change from previous year)
Lg5	Corporation tax revenue
Lg6	Unemployment rate
Lg7	Contractual cash earnings (manufacturing)
Lg8	Consumer price index (all items, less fresh food, change from previous year)
Lg9	Index of producer's inventory (final demand goods)

Table 8 Summary of estimations for several lagging indicators

Number	$\widehat{k}$*-value*	*ACD(1)*	*ACD(k)*	*q-value*
Lg2	$\widehat{k} = 2$	0.9881	0.9887	$q = 0$
Lg4	$\widehat{k} = 9$	0.9454	0.9754	$q = 6$
Lg7	$\widehat{k} = 2$	0.9863	0.9877	$q = 0$
Lg9	$\widehat{k} = 4$	0.9861	0.9865	$q = 0$

As typical examples, we present the estimation results for Lg2 and Lg7 in Figures 13 and 14. The results are generally satisfactory.

Figure 13 Data and estimation results for series Lg2

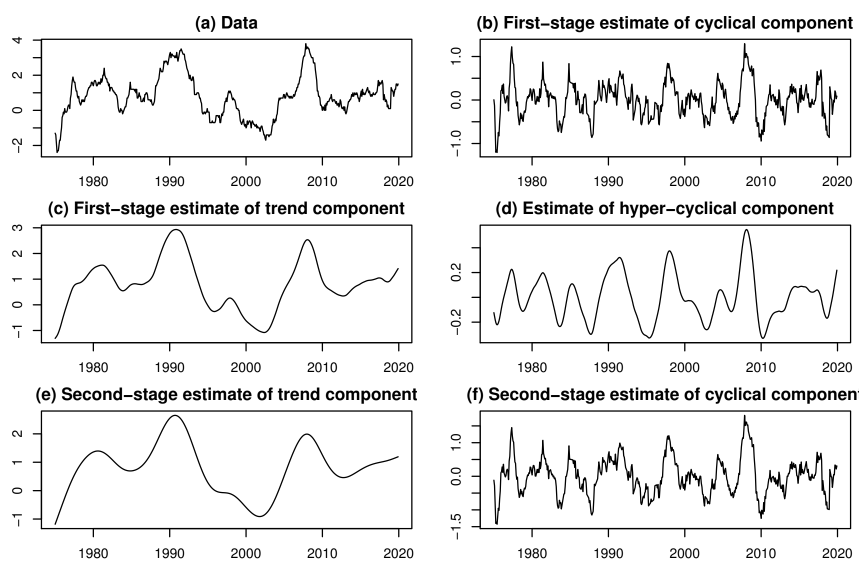
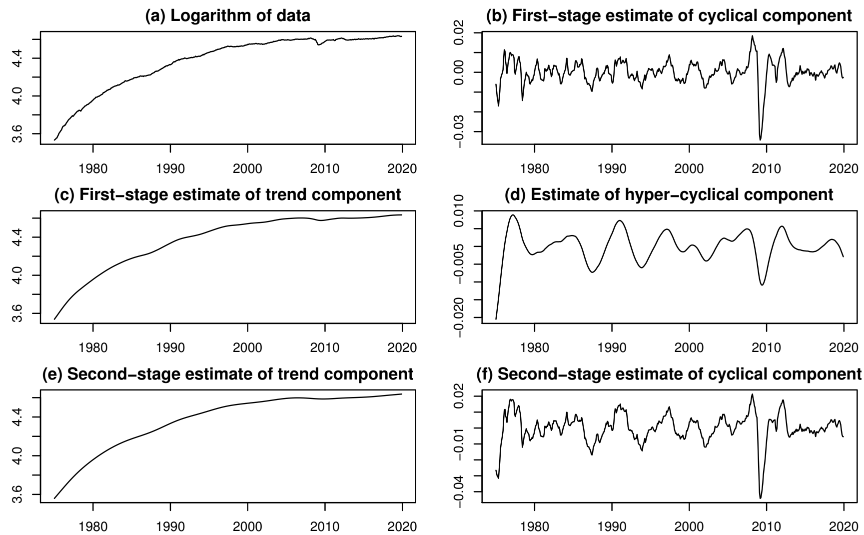


Figure 14 Data and estimation results for series Lg7



The results presented in this subsection can be summarised as follows. For the data related to commercial sales in Japan, we applied the second-stage estimates to seven of the 20 series (35%). Further, among the 30 series of business cycle indicators, we required the hyper-trend method for 13 series (about 45%). That is, for about 40% of the analysed series, we applied the proposed approach.

6 Summary and discussion

For economic analysis, it is very important to decompose a time series into several components. In the decomposed results, the trend and cyclical components are particularly important: the trend component indicates the long-period behaviour and the cyclical component indicates the cyclical variation, such as business cycles. Thus, the conclusion of any economic analysis may be influenced by the decomposed results. Sometimes, however, when the analysed time series contains long-period variations, existing methods cannot capture each component completely; hence, part of the cyclical variation may remain in the estimate of the trend component.

To overcome this difficulty, we proposed an approach called the hyper-trend method. In the newly proposed approach, we assume that part of the cyclical variation remains in the estimate of the trend component obtained using the existing decomposition method. We decomposed the estimated trend component into hyper-trend and hyper-cyclical components by introducing the IATCD and HTR modelling methods. In these modelling methods, we introduced the LSI as a parameter related to the pattern of variation of the hyper-cyclical component. To determine the value of the LSI, we further introduced the ACD criterion, which measures the properties of the estimate of the cyclical component. Finally, based on the proposed modelling methods, we provided an overall procedure for applying the proposed approach. The proposed approach is called the hyper-trend method because the procedure for estimating the hyper-trend components is the core technique.

To illustrate the performance of the proposed approach, we applied it to two types of time series data. The first type of time series related to commercial sales in Japan, and were monthly data containing a seasonal component. The second type of time series related to business cycle analysis indicators, and was seasonally adjusted monthly time series. Among the 20 series of the first type, the hyper-trend method was applied to seven series (35%). Moreover, among the 30 series of the second type, we required the hyper-trend method for 13 series (about 45%).

The results demonstrate the efficiency of the constructed ACD criterion. In this paper, to determine the value of the LSI, we used the method of maximising the ACD so that only one set of estimation results could be adopted. There may be multiple patterns for different cyclical components that correspond to different values of the ACD. Thus, the proposed approach can be extended to decompose a time series into multiple stationary components with different cycle periods by taking the ACD criterion as an inference.

Another possibility for the proposed approach is the application to the estimation of long-memory time series (see Granger and Joyeux, 1980). Although many methods for estimating long-memory time series have been proposed (see Robinson, 2003), it may be possible to develop a simple and easy method for analysing long-memory time series by extending the proposed approach.

Acknowledgements

This work was supported by a Grant-in-Aid for Scientific Research (B) (18H03210) and a Grant-in-Aid for Scientific Research (C) (19K01583) from the Japan Society for the Promotion of Science. We thank Stuart Jenkinson, PhD, from Edanz (https://jp.edanz.com/ac) for editing a draft of this manuscript.

References

Akaike, H. (1974) 'A new look at the statistical model identification', *IEEE Transactions on Automatic Control*, Vol. AC-19, No. 6, pp.716–723.

Akaike, H. (1980) 'Seasonal adjustment by a Bayesian modeling', *Journal of Times Series Analysis*, Vol. 1, No. 1, pp.1–13.

Akaike, H. and Ishiguro, M. (1980) 'BAYSEA, a Bayesian seasonal adjustment program', *Computer Science Monographs*, No. 13, The Institute of Statistical Mathematics, Tokyo.

Baxter, M. and King, R. (1999) 'Measuring business cycles: approximate band-pass filter for economic time series', *Review of Economics and Economic Statistics*, Vol. 81, No. 4, pp.575–593.

Bell, W.R. and Hillmer, S.C. (2002) 'Issues involved with the seasonal adjustment of economic time series', *Journal of Business and Economic Statistics*, Vol. 20, No. 1, pp.98–127.

Beveridge, S. and Nelson, C.R. (1981) 'A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the business cycle', *Journal of Monetary Economics*, Vol. 7, No. 2, pp.151–174.

Box, G.E.P. and Jenkins, G.M. (1976) *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco.

Clark, P.K. (1987) 'The cyclical component of U.S. economic activity', *The Quarterly Journal of Economics*, Vol. 102, No. 4, pp.797–814.

Cleveland, W.P. and Tiao, G.C. (1976) 'Decomposition of seasonal time series: a model for the census X-11 program', *Journal of the American Statistical Association*, Vol. 71, No. 355, pp.581–587.

Findley, D.F., Monsell, B.C., Bell, W.R., Otto, M.C. and Chen, B. (1998) 'New capabilities and methods of the X-12-ARIMA seasonal-adjustment program', *Journal of Business and Economic Statistics*, Vol. 16, No. 2, pp.127–177.

Gersch, W. and Kitagawa, G. (1983) 'The prediction of time series with trends and seasonalities', *Journal of Business and Economic Statistics*, Vol. 1, No. 3, pp.253–264.

Gómez, V. and Maravall, A. (1996) *Programs TRAMO (Time Series Regression with ARIMA Noise, Missing Observations, and Outliers) and SEATS (Signal Extraction in ARIMA Time Series): Instruction for the User*, Working paper, No. 9628, Research Department, Bank of Spain.

Granger, C.W.J. and Joyeux, R. (1980) 'An introduction to long-memory time series models and fractional differencing', *Journal of Time Series Analysis*, Vol. 1, No. 1, pp.15–29.

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

Harvey, A.C., Trimbur, T.M and van Dijk, H.K. (2007) 'Trends and cycles in economic time series: a Bayesian approach', *Journal of Econometrics*, Vol. 140, No. 2, pp.618–649.

Hoeting, J.A., Madigan, D., Raftery, A.E. and Volinsky, C.T. (1999) 'Bayesian model averaging: a tutorial', *Statistical Science*, Vol. 14, No. 4, pp.382–417.

Hodrick, R.J. and Prescott, E.C. (1997) 'Postwar U.S. business sycles: an empirical investigation', *Journal of Money, Credit and Banking*, Vol. 29, No. 1, pp.1–16.

Ishiguro, M. (1984) 'Computationally efficient implementation of a Bayesian seasonal adjustment procedure', *Journal of Times Series Analysis*, Vol. 5, No. 4, pp.245–253.

Kitagawa, G. (1981) 'A non-stationary time series model and its fitting by a recursive filter', *Journal of Time Series Analysis*, Vol. 2, No. 2, pp.103–116.

Kitagawa, G. (1985) *DECOMP: In TIMSAC-84 Part 2*, Computer Science Monographs, No. 23, The Institute of Statistical Mathematics, Tokyo.

Kitagawa, G. (2020) *Introduction to Time Series Modeling with Application in R*, 2nd ed., Chapman and Hall, New York.

Kitagawa, G. and Gersch, W. (1984) 'A smoothness priors-state space modeling of time series with trends and seasonality', *Journal of the American Statistical Association*, Vol. 79, No. 386, pp.378–389.

Monsell, B.C. (2007) 'The X-13A-S seasonal adjustment program', *Proceedings of the 2007 Federal Committee on Statistical Methodology Research Conference* [online] https://nces.ed.gov/FCSM/pdf/2007FCSM_Monsell-II-B.pdf.

Morley, J.C., Nelson, C.R. and Zivot, E. (2003) 'Why are the Beveridge-Nelson and unobserved-components decompositions of GDP so different?', *The Review of Economics and Statistics*, Vol. 85, No. 2, pp.235–243.

Robinson, P.M. (2003) *Time Series with Long Memory*, Oxford University Press, London.

Shiskin, J., Young, A.H. and Musgrave, J.C. (1967) *The X-11 Variant of the Census Method II Seasonal Adjustment Program*, Technical report, No. 15, US Bureau of the Census, Washington, DC.

Sims, C.A., Stock, J.H. and Watson, M.W. (1990) 'Inference in linear time series models with some unit roots', *Econometrica*, Vol. 58, No. 1, pp.113–144.

Watson, M.W. (1986) 'Univariate detrending methods with stochastic trends', *Journal of Monetary Economics*, Vol. 18, No. 1, pp.49–75.

Wu, Z., Huang, N.E., Long, S.R. and Peng, C.K. (2007) 'On the trend, detrending, and variability of nonlinear and nonstationary time series', *PNAS*, Vol. 104, No. 38, pp.14889–14894.