Luxembourg: Office for Official Publications of the European Communities, 2006

ISBN 92-79-03407-3

ISSN 1725-4825

Catalogue number: KS-DT-06-008-EN-N

A STRUCTURAL TIME SERIES MODEL FACILITATING FLEXIBLE SEASONALITY

Yoshinori Kawasaki *

The Institute of Statistical Mathematics,

Research Organization of Information and Systems,

4–6–7 Minami Azabu, Minato-ku, Tokyo, 106–8569, Japan

Abstract

We shed light on a class of models that increase the flexibility of the seasonal pattern within a framework of the structural time series model. The basic idea is to drive the seasonal summation model by a moving average process rather than by a white noise or an AR process. Generally, such an approach can be exhaustive in parameters, but the proposed model is parsimonious in the sense that we have only one extra parameter compared to the basic structural time series model. Because we stay at the linear Gaussian assumption, the estimation is quite easy and fast. The state space representation of the model is also given. An interpretation of moving average driven seasonal model is provided in terms of the offset effect on the pseudo-spectrum around the seasonal frequencies. The empirical analysis demonstrates that the proposed method is richly expressive in estimating the seasonal component, and is also supported by the minimum AIC procedure. A few cases where the proposed method is not working well provide us some useful information on the possible misspecification. Focusing attention on the two key quantities implied by the estimated models, we propose a graphical representation for the estimated models that help us to discover the unsuccessful cases and to confirm whether or not the alternative specification improves the modeling.

Subject Area: Seasonal adjustment techniques: comparison of alternative methods

1 Introduction

Time series with trend and seasonal components is an important generalization of nonstationary mean time series. Such time series occur for example in meteorological, oceanographic and economic studies. Trend and seasonal, given time index t, are the unobserved components to be squeezed out of the original observation. This ill-posed nature requires the introduction of reasonable stochastic constraints on the unobserved components, which leads to some Bayesian treatment. One of the earliest works of such Bayesian modeling is perhaps Harrison and Stevens (1971). Akaike (1980) presented a sophisticated methodology on Bayesian seasonal adjustment

*Associate Professor at Department of Statistical Modeling, The Institute of Statistical Mathematics. E-mail: kawasaki@ism.ac.jp

method which admits computationally straightforward penalized least squares methods. In estimating such a decomposition problem in a Bayesian framework, a Markovian representation via state space form is very useful. Kitagawa (1981), Gersch and Kitagawa (1983) and Kitagawa and Gersch (1984) extended the ideas in Akaike (1980) to state space formulations and incorporated seasonal decomposition features. Harvey and Todd (1983) and Harvey (1984, 1985) went side by side along with this line of research though a Bayesian point of view is not very much stressed. West *et al.* (1985) and West and Harrison (1986) developed a fully Bayesian framework while Akaike's methodology (and his follower's too) can be regarded as empirical- or quasi-Bayesian approach. In this paper we consider the modeling of nonstationary mean time series with trend and seasonal by state space methods.

The economic application of seasonal adjustment time series have provoked an extensive literature and a variety of software. Popular software products based on state space modeling are Kitagawa's DECOMP in TIMSAC 84 (Akaike *et al.* , 1985) and STAMP (Koopman *et al.* , 2000) initially developed by Harvey. Empirical researchers using such softwares sometimes complain that the estimated seasonal patterns are very steady considering the use of the stochastic seasonality. In a state space modeling, the most commonly used seasonal component model is the stochastic dummy seasonality, that is, the sum of seasonal factors within a period follows zero mean white noise. Frequently observed stiff seasonal patterns are rather due to this stochastic dummy specification, not to the state space formulation itself.

One way to increase the seasonal variability is to introduce a fat-tailed distribution for the innovation of seasonal component. In Kitagawa (1989), the spline based numerical integration developed in Kitagawa (1987) is applied for seasonal adjustment, while Kitagawa (1994) proposes a new implementation for the Gaussian-sum smoother. Monte Carlo filter presented by Kitagawa (1996) is also applicable to a non-Gaussian seasonal adjustment. A method proposed by Shephard and Pitt (1997) is also applicable though their presentation puts emphasis on non-Gaussian measurement. However, these techniques are more likely to be instrumental to depict the abrupt change in the seasonal pattern or to detect the outliers automatically, or when the measurement distribution is primarily non-Gaussian. In addition, there still remains a computationally intensive task, and often a user-friendly tool is not available for practitioners to perform it.

The aim of this article is to present a new model to increase the flexibility and variability of seasonal pattern within a framework of structural time series model. We still stay at the linear Gaussian assumption, hence the estimation is quite easy and fast. The empirical analysis performed in this article demonstrates that the proposed method is richly expressive in estimating the seasonal component, and that such a decomposition is supported in terms of one-step ahead prediction, too.

This paper is organized as follows. In section 2 we briefly review the framework of basic structural model for time series with trend and seasonality. A state space form for the BSM and its estimation algorithm are also reviewed in subsection 2.2 and 2.3. In section 3, a parsimonious modeling to produce flexible seasonality is presented. After introducing some existing methods in subsection 3.1, a model with seasonal summation driven by a finite order moving average process with a single unknown parameter is proposed in subsection 3.2. A state space form for the proposed model are described in section 3.4. Section 4 details the results of the empirical analyses on 11 economic time series of U.K. and Japan. Section 5 concludes this article.

2 Modeling Trend-Seasonality

2.1 Basic Structural Model

As a basis of the discussion of this article, this section explains a popular model called basic structural model. Modeling trend-seasonality with state-space form have been explored since the end of 1970's. Trend and seasonal are regarded as unobservable components, and for each unobservable component a stochastic model is assumed. One of the most popular specification is a set of the equations described as follows.

$$y_t = \mu_t + \gamma_t + \varepsilon_t \tag{1}$$

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t \tag{2}$$

$$\beta_t = \beta_{t-1} + \zeta_t \tag{3}$$

$$\gamma_t = -\sum_{j=1}^{s-1} \gamma_{t-j} + \omega_t \tag{4}$$

In above equations, we assume that each of ε_t, η_t, ζ_t, ω_t follows zero mean normal distributions but with different variance; σ_ε^2, σ_η^2, σ_ζ^2 and σ_ω^2 respectively. This set of equations is often referred to as Harvey's basic structural model (BSM hereafter), see Harvey (1989, p.47). Equation (1) is called observational (or measurement) equation. This reflects our observation that the salient features of economic time series are trend μ_t and seasonality γ_t, and the rest is regarded as irregular component ε_t.

Trend component consists of two latent variables μ_t and β_t, which is respectively referred to 'stochastic level' and 'stochastic slope'. The equation (2) plus (3) is called the local linear trend model. The name comes from the fact that the drift term β_t plays a role of a linear trend rather than a constant in (2). On the other hand, it is also possible to consider the following trend model in stead of (2) plus (3);

$$\mu_t = 2\mu_{t-1} - \mu_{t-2} + \eta_t. \tag{5}$$

If we rewrite (5) as $\mu_t = \mu_{t-1} + (\mu_{t-1} - \mu_{t-2}) + \eta_t$, it is easily understood that (5) is a special case of the local linear trend model in the sense that the stochastic slope β_t is also driven by the same process η_t rather than by a different process ζ_t. From now on, trend model is fixed to (5) in this article, and the seasonal adjustment model (1) together with (5) and (4) will be referred to the BSM again, as this will make no confusion here.

2.2 State Space Form

In order to facilitate the introduction of extended models in the next section, we sum up the basic outline of the state space representation of the BSM and its estimation. Due to the assumption of no correlation among innovation and noise process, the state space representation can be built up as a composition of small state space models for the individual components. To save space, we assume $s = 4$ just for the presentation purpose. From equation (5) and (4), it turns out that the essential quantity that determines the present distribution of μ_t and γ_t will be given by a vector

$$\alpha_{t-1} = (\mu_{t-1}, \mu_{t-2}, \gamma_{t-1}, \gamma_{t-2}, \gamma_{t-3})' \tag{6}$$

where the prime ($'$) denotes the transpose of a vector or a matrix. By setting submatrices as follows,

$$T_1 = \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix}, \; T_2 = \begin{bmatrix} -1 & -1 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \; R_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \; R_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix},$$

new matrices T and R are defined as

$$T = \begin{bmatrix} T_1 & O \\ O & T_2 \end{bmatrix}, \quad R = \begin{bmatrix} R_1 \\ R_2 \end{bmatrix}, \quad \eta_t = \begin{bmatrix} \eta_t \\ \omega_t \end{bmatrix}. \tag{7}$$

Then the transition of the state vector can be written in a matrix notation as

$$\alpha_t = T\alpha_{t-1} + R\eta_t. \tag{8}$$

As we observe that the measurement equation (1) just extracts and adds the components μ_t and γ_t, defining $z' = (1,0,1,0,0)$ yields the relation between the observation and the state as

$$y_t = z'\alpha_t + \varepsilon_t. \tag{9}$$

Now the BSM is put in a state space form by (8) and (9). Note that the specification of α_t, T, R and z' described above is not a unique one because the transformations of these vectors and matrices by a regular square matrix still give rise to the same state space model.

2.3 Model and State Estimation

Let a_{t-1} denote the minimum mean squared error (MMSE) estimator of α_{t-1} based on the observations up to time $t-1$. Let P_{t-1} denote the $m \times m$ covariance matrix of the estimation error, i.e.

$$P_{t-1} = \mathrm{E}[(\alpha_{t-1} - a_{t-1})(\alpha_{t-1} - a_{t-1})'].$$

Given a_{t-1} and P_{t-1}, the MMSE estimator of α_t and the covariance matrix of the estimation error is given by

$$\begin{aligned} a_{t|t-1} &= Ta_{t-1} \\ P_{t|t-1} &= TP_{t-1}T' + RQR' \end{aligned}$$

where $Q = \mathrm{diag}(\sigma_\eta^2, \sigma_\omega^2)$. These two equations are known as the *prediction equations*.

Once the new observation, y_t, becomes available, the estimator of α_t, $a_{t|t-1}$, can be updated. The *updating equations* are given by the following two equations,

$$\begin{aligned} a_t &= a_{t|t-1} + P_{t|t-1}z'f_t^{-1}(y_t - z'a_{t|t-1}) \\ P_t &= P_{t|t-1} - P_{t|t-1}z'f_t^{-1}zP_{t|t-1} \end{aligned}$$

where $f_t = z'P_{t|t-1}z + \sigma_\varepsilon^2$. Repetition of prediction and updating constitutes so-called the Kalman filter.

Unless $\sigma_\varepsilon^2 = 0$, the estimation problem of a state space model is double-folded. Given the unknown hyperparameters $\psi = (\sigma_\varepsilon^2, \sigma_\eta^2, \sigma_\omega^2)'$, running Kalman filter and fixed interval smoother yields the estimates of unobservable components $\{\hat{\mu}_t\}_{t=1}^T$, $\{\hat{\gamma}_t\}_{t=1}^T$ and hence $\{\hat{\varepsilon}_t\}_{t=1}^T$. The vector of unknown parameters, ψ, can be estimated by the maximum likelihood method. The likelihood function for a time series can be decomposed into the product of the density functions of one step ahead prediction error $v_t = y_t - z' a_{t|t-1}$. The variance of observation noise σ_ε^2 usually can be concentrated out of the likelihood function. Let $\psi^* = (\sigma_\eta^2, \sigma_\omega^2)'$, then

$$\log L_c(\psi^*) = -\frac{1}{2}\left\{T \log 2\pi\tilde{\sigma}^2(\psi^*) + \sum_{t=1}^{T} \log f_t + T\right\}$$

must be maximized with respect to the unknown parameters ψ^*, while $\tilde{\sigma}^2(\psi^*)$ is given by

$$\tilde{\sigma}^2(\psi^*) = \frac{1}{T}\sum_{t=1}^{T}\frac{v_t^2}{f_t}.$$

Model comparison will be done based on AIC (Akaike, 1973). As regards the initial state settings, we employ the 'large κ approximation' (Harvey, 1989, pp.121). The specific value for κ employed here will be stated in section 4 in conjunction with the scale of the time series. Once the unknown hyperparemeters are estimated, then the unobserved components are estimated by the fixed interval smoother. For the algorithm of the fixed interval smoother, see Anderson and Moore (1979, pp.187–190), Harvey (1989, pp.154) or Kitagawa and Gersch (1996, pp.58).

3 Parsimonious Modeling toward Flexible Seasonality

In this section, three seasonal component models which will be compared in the real data analysis section 4 are presented. In the subsection 3.1, the standard model (the BSM) and its existing modification are presented. In the section 3.2, a parsimonious modeling of MA driven seasonal summation will be introduced.

The basic motivation of this paper is to examine the appropriateness of a new seasonal model which increases the variability of the seasonal component. However, it is *not* because we believe that larger variance of the seasonal component is desirable. What is sought in this paper is, at first, to prepare the framework which allows us to estimate more flexible seasonal component than the BSM, and secondly, to investigate through empirical analysis whether such a model is really favored or not in terms of a model selection criterion. If we estimate a seasonal ARIMA model and try to decompose it to do seasonal adjustment, it is known that there is no unique solution for such a decomposition. Then we must place an arbitrary assumption on the allocation of the variance contributions among trend, seasonal and other components under consideration. For example, Box et al. (1978) asserts that the variance of the seasonal component should be minimized. It should be noted, however, that we do not have to employ such an arbitrary criterion because we only have to determine the hyperparameters by the maximum likelihood method and compare the candidate model by AIC statistic.

3.1 Driving Noise of Seasonal Summation

Let s be the number of seasons observed in a period. For economic time series, the length of a period is usually one year and the cases of $s = 4$ and 12 draw great deal of attention. Now, define the seasonal summation operator $S(L)$ by

$$S(L) = 1 + L + \cdots + L^{s-1}. \tag{10}$$

Then a concise expression $S(L)\gamma_t = \omega_t$ can be given to (4). This model is often referred to as 'dummy seasonality', see Harvey (1989). To avoid a possible confusion with deterministic dummy seasonality, we call this model the stochastic dummy seasonality. The model (4) can be regarded as a stochastic constraint on seasonal component such that the sum of s-consecutive seasonal factors will follow zero mean independent random variable. Though the seasonal component γ_t can vary as time evolves, the seasonal pattern cannot change very much if the estimated dispersion parameter $\hat{\sigma}_\omega^2$ is very small. As is already defined in section 2.1, the second order difference equation (5) plus (10) will be referred to as the BSM in the data analysis section 4.

There have been several researches to allow more flexibility for the seasonal component. One idea is to employ trigonometric seasonal specification, see Hannan, Terrell and Tuckwell (1970) and Harvey (1989, p.41–42). But a scepticism may be cast on this model that a more emphasis is put on the evolution of separate seasons than on the serial association of consecutive seasons. Another obvious drawback is that this approach requires many additional hyperparameters while we cannot always expect the gain in fitting accuracy. In such a case, though it depends on the situation, the number of hyperparameters may be reduced by the equality/zero constraints on some of the dispersions of seasonal components.

Kitagawa and Gersch (1984, p.386) introduce a higher order seasonal polynomial such that $S^2(L)\gamma_t = \omega_t$. This type of modeling can cope with gradual change in seasonal pattern while its difficulty is that the state dimension of the model becomes larger. This extension has been already implemented in the software DECOMP in TIMSAC-84, and is also available on Web-Decomp. (As regards Web-Decomp, see Sato (1997) and the web site he maintains; `http://www.ism.ac.jp/~sato/` or `http://ssnt.ism.ac.jp/inets2/title.html`)

Seasonal model (4) can be viewed as the seasonal summation is driven by a white noise process. One idea to bring more variability to the stochastic dummy seasonality is to replace the white noise by the 'colored' (i.e., autocorrelated) noise,

$$(1 - \Phi L)S(L)\gamma_t = \omega_t. \tag{11}$$

Provided that the variance σ_ω^2 is the same, the unconditional variance of the seasonal summation, $\sigma_\omega^2/(1-\Phi)^2$ is greater than σ_ω^2 as long as $\Phi < 1$. Ozaki and Thomson (1992) call (11) the pink-noise driven seasonal component model because the innovation process $\omega_t/(1-\Phi L)$ has much power at lower frequencies that reminds us of the infrared ray. Throughout this article, the second order stochastic trend (5) plus (11) will be referred to as the BSM-AR.

3.2 Seasonal Summation Driven by MA

To increase the seasonal summation variability, it appears more direct to introduce a finite order MA process on the right hand side of (10). Let $\Theta(L)$ denote a certain form of polynomial of the backward shift operator L, then the seasonal component model can be written as $S(L)\gamma_t = \Theta(L)\omega_t$. What we concern about is not a formal extension of models but a practical guideline to specify $\Theta(L)$.

A motivation for MA-driven seasonal model is found, for example, in the past effort to build a time series model which is expected to play the same role as a conventional seasonal adjustment procedure. From 1970's to early 80's, many researchers sought unobserved components models that approximate Census X-11 seasonal adjustment procedure. See Cleveland and Tiao (1976), Wallis (1982), Burridge and Wallis (1984), for example. Some authors proposed decomposition methods based on seasonal ARIMA model, Burman (1980) and Hillmer and Tiao

(1982) to name a few. As Burridge and Wallis (1984) pointed out, models with a predominantly autoregressive specification generate long signal-extraction filters that are not able to approximate the relatively rapid decline of the X-11 filter coefficients. Because the seasonal model in the BSM involves only AR polynomial in (10), the corresponding filter coefficients are not ignorable even in the remote lags. In this context, there is a clear preference for moving average dominated specifications in modeling seasonality. In general, many articles in this period recommend the inclusion of a seasonal moving average model. For example, Burridge and Wallis (1984) proposed the following component model (for monthly economic time series) which will correspond to the symmetric filter in Census X-11 procedure,

$$S(L)\gamma_t = (1+0.71L^{12}+1.00L^{24})\omega_t.$$

Another motivation emanates from accumulated experiences on seasonal ARIMA model fitting to seasonal time series. Since Box and Jenkins (1976), there have been many empirical works to show that ARIMA(0,1,1)$\times$(0,1,1)$_s$ accounts for a wide variety of economic time series with seasonality. This econometric folklore tells us that, for a detrended series $\bar{y}_t$,

$$(1-L^s)\bar{y}_t = (1-\Theta L^s)\varepsilon_t,\ |\Theta|<1 \tag{12}$$

fits reasonably. Seasonal differencing operator $1-L^s = (1-L)(1+L+\cdots+L^{s-1})$ involves the usual differencing operator $1-L$. To avoid common factor between trend and seasonal component model, only the summation type operator is considered when modeling seasonality in a structural time series model. Hence the equation (12) suggests the following seasonal component model,

$$S(L)\gamma_t = (1+\Theta L+\Theta^2L^2+\cdots+\Theta^{s-1}L^{s-1})\omega_t. \tag{13}$$

where $|\Theta|<1$. Provided that the variance σ_ω^2 is the same, the unconditional variance of the seasonal summation, $(1+\Theta+\Theta^2+\cdots+\Theta^{s-1})\sigma_\omega^2$ is greater than σ_ω^2 unless $\Theta=0$. Throughout this article, the second order stochastic trend (5) plus (13) will be referred to as the BSM-MA. Originally, the words BSM, BSM-AR and BSM-MA refer to a set of equations. From now, these words will be used as if they point only to the seasonal component models in some cases , but this will not invite misunderstanding.

In stead of (13), we could start with a more general model such as

$$S(L)\gamma_t = (1+\theta_1L+\theta_2L^2+\cdots+\theta_{s-1}L^{s-1})\omega_t. \tag{14}$$

There are a couple of reasons why we do not employ this general form. Firstly, (14) is exhaustive in the number of parameters. Second, even if we employ the 'general-to-specific' modeling strategy, there is another parametrization to take over (14) that gives us a more intuitively natural interpretation of the parameters, and that shows a more reasonable way to put restrictions on the parameters. Finally, such a specification enables us to understand the offset effect of the BSM-MA, which will be clarified in the next subsection.

3.3 Pseudo-Spectrum Offset around Seasonal Frequencies

In this subsection, we shed another light on the role of MA term in seasonal component model. It is known that the simultaneous use of AR and MA operators can mimic a 'line spectrum' when the zeros of both operators have a common argument. Whether it gives rise to a peak or a trough depends on the magnitude relation of the modulus of roots. If the modulus of AR

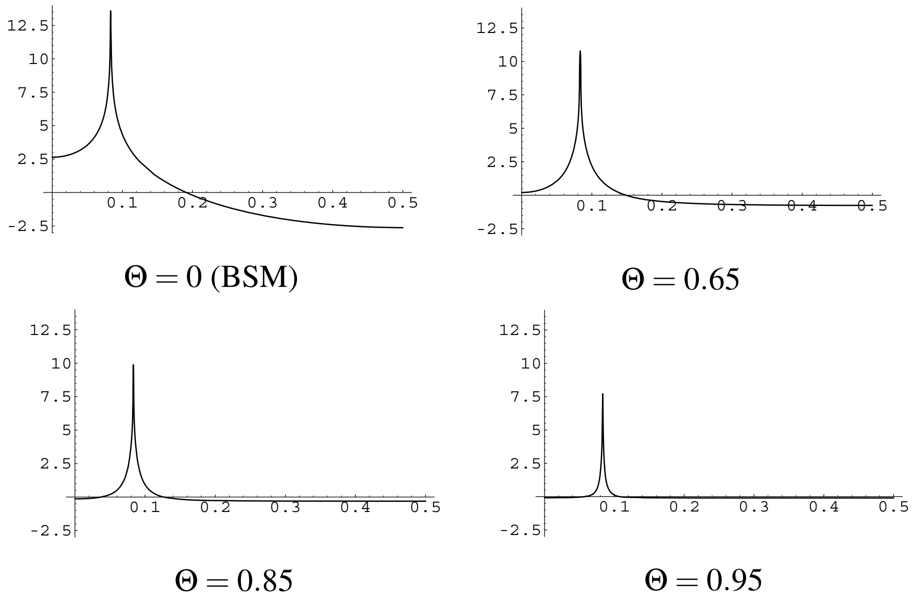

Figure 1: Log-spectrum shape of ARMA(2,2) processes.

roots are greater/smaller than those of MA roots, then the spectrum has peaks/troughs. Let us consider the power spectrum $f(\lambda)$ ($0 \leq \lambda \leq \pi$) of ARMA(2,2) process given by

$$f(\lambda) \propto \left|1-\sum_{j=1}^{2}\Theta_j e^{-2\pi ij\lambda}\right|^2 \Big/ \left|1-\sum_{j=1}^{2}\Phi_j e^{-2\pi ij\lambda}\right|^2,$$

where Φ_1 and Φ_2 are fixed to $0.99\sqrt{2}$ and $-(0.99)^2$ respectively, and $\Theta_1 = \Theta\sqrt{2}$ and $\Theta_2 = -\Theta^2$ vary dependent on the parameter Θ. Because we only consider $|\Theta| < 1$ cases, the power spectrum $f(\lambda)$ has its peak at $\lambda^* = 0.083$.

Four panels in Figure 1 shows how the shape of $\log f(\lambda)$ changes as Θ tend to unity. The left-upper panel ($\Theta = 0$) corresponds to the BSM, in which case the power spectrum is widely spread around $\lambda^* = 0.083$. It is easily seen that the peak becomes sharper and more concentrated around λ as Θ tends to unity. In addition to that, the level of power spectrum except for the peak frequency gets flatter as $\Theta \to 1$. In other words, simultaneous use of AR and MA polynomials with common argument offsets the power spectrum at all frequencies but the common argument (in this example λ^*=0.083). As a result, the shape of log-power spectrum closely resembles that of a line spectrum apart from constant.

If we turn to the seasonal component models (4), (11) and (13), the power spectrum cannot be defined any more because the process is not stationary due to the unit roots contained in $S(L) = 0$. Even for such a case, a formally defined spectrum called pseudo-spectrum is often considered to characterize a time series. See Harvey (1989, p.64). As an example, the log of pseudo-spectra of seasonal components for both BSM and BSM-MA with $\Theta = 0.9$ are drawn in the left panel of Figure 2. For simplicity, we assign the same value to σ_ω^2 in both models, and the root of AR polynomial is slightly pitched outside the unit circle just for drawing this figure. (Theoretically, two peaks at $\lambda = 0.25, 0.5$ should be infinite.) We observe that the seasonal component of the BSM accounts for not only the power at seasonal frequencies but also the substantial portion of the power at the neighboring frequencies. Considering the offset effect mentioned above, there is a possibility that BSM-MA may prevent the BSM from excessively

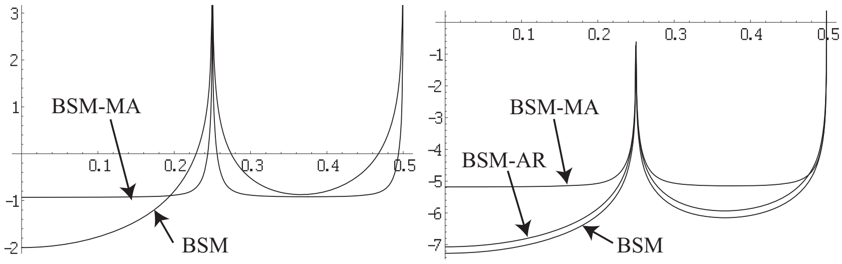

Figure 2: Left: Log of pseudo-spectra of the BSM and BSM-MA (with $\Theta = 0.9$), Right: Estimated pseudo-spectra of seasonal component for private sector consumption in Japan (PCSMP).

removing the frequency components neighboring the seasonal frequencies. In addition, the seasonal component of BSM-MA with $\Theta = 0.9$ has more power at lower frequencies than the BSM, which is also due to the offset effect. Meanwhile, the stochastic dummy model loses its power at lower frequencies and enhances the power at higher frequencies.

In terms of the smoother mean, however, the most crucial factor is the level of the power spectrum, σ_ω^2. Technically, $\sigma_\omega^2 \approx 0$ is also required for $f(\lambda)$ to be close to a line spectrum. So it is legitimate to say that the simultaneous use of AR and MA terms with a common argument will produce an almost deterministic periodicity contaminated with a white noise of which level completely depends on the time series in question. If $\Theta \to 1$ and $\sigma_\omega^2 \to 0$ simultaneously, the estimated seasonal factor should be almost deterministic. This is the case of 'cancellation', see Box and Jenkins (1976, pp. 248). Then, removing the seasonality in advance by deterministic dummy variables would be appropriate. If $\Theta \to 1$ but $\sigma_\omega^2 >> 0$, then the seasonal factor will be interpreted as the deterministic dummy variables on which a white noise process is superimposed. The right panel of Figure 2 shows the log of the estimated power spectra of the seasonal components of BSM, BSM-AR and BSM-MA for Japanese private sector total consumption. As is common with many series analyzed in the section 4, the power level of the seasonal component is much increased and flattened by employing BSM-MA.

ARMA modeling with common argument and different modulus can be introduced to all the seasonal frequencies of the stochastic dummy model. Let us denote the fundamental seasonal frequency and its harmonics by $\lambda_j = 2\pi j/s$ for $j = 1, \ldots, [s/2]$ where

$$[s/2] = \begin{cases} s/2 & \text{for } s \text{ even} \\ (s-1)/2 & \text{for } s \text{ odd} \end{cases}$$

The seasonal summation operator can be written as the product of the full complement of trigonometric operators, i.e.,

$$S(L) = \prod_{j=1}^{[s/2]} \gamma_j(L)$$

where

$$\gamma_j(L) = 1 - (2\cos\lambda_j)L + L^2, \quad j = 1, \ldots, [s/2]$$

when s is odd. When s is even, $\gamma_j(L)$ is defined as above for $j = 1, \ldots, s/2 - 1$, while for $j = s/2$ it is

$$\gamma_{s/2}(L) = 1 + L.$$

For monthly data the seasonal summation operator can be factorized into the six trigonometric operators, see Harvey (1989, p.22) for example. Let us take a look at one of the trigonometric

AR polynomials, $1-\sqrt{3}L+L^2$, which corresponds to 12 months period. Hence introducing MA term polynomial $1-\sqrt{3}\Theta_1 L+\Theta_1^2 L^2$ with $-1<\Theta_1<1$ leads to a sharp peak at 'once-in-a-year' frequency, $\lambda_1=\pi/6$. Allowing MA term at all the seasonal frequencies, the stochastic dummy seasonal component model can be extended to take the following form,

$$\begin{aligned} S(L)\gamma_t &= (1-\sqrt{3}\Theta_1 L+\Theta_1^2 L^2)(1-\Theta_2 L+\Theta_2^2 L^2)(1+\Theta_3^2 L^2) \\ &\times\ (1+\Theta_4 L+\Theta_4^2 L^2)(1+\sqrt{3}\Theta_5 L+\Theta_5^2 L^2)(1+\Theta_6 L)\omega_t \end{aligned}$$

where the $|\Theta_j|$'s are all expected to be less than unity. If the six parameters $\Theta_1,\ldots,\Theta_6$ are estimated freely, it means that peak properties can differ by the seasonal frequencies. But it may cause too much flexibility to obtain a slight gain in accounting for the process variation. Thus in this article we assume $\Theta_1=\cdots=\Theta_6=\Theta$ which reduces to BSM-MA, (13). This equality constraint can be rephrased that the offset effect is expected all alike for the seasonal frequencies, and the roots of the MA polynomial are pitched outside the unit circle at an equal distance.

3.4 State Space Representation for BSM-MA

We close this section with a remark on a state space representation of BSM-MA model. For the BSM, the state vector (6) consists of the unobserved components and their lagged variables. As regards the transition matrix for the seasonal component (the T_2 block of matrix T in (7)), the first row essentially corresponds to the seasonal component model and other rows merely shift the time index. The same holds for BSM-AR, too. However, such a simple construction cannot be extended straightforward if the moving average terms are incorporated in the model.

A state space representation for (13) will be given as follows. Let $\tilde{\gamma}_{t+i|t-1}$ be a predictor of γ_{t+i} based on the observation up to time $t-1$, and on the innovations up to t, namely,

$$\tilde{\gamma}_{t+i|t-1} = -\sum_{j=i+1}^{s-1}\gamma_{t+i-j} - \sum_{j=i}^{s-1}\Theta^j\omega_{t+i-j}.$$

It can be readily verified that the following recursive relationship holds,

$$\gamma_t = \gamma_{t-1}+\tilde{\gamma}_{t|t-2}+\omega_t \tag{15}$$

$$\tilde{\gamma}_{t+i|t-1} = -\gamma_{t-1}+\tilde{\gamma}_{t+i|t-2}-\Theta^i\omega_t,\ i=1,\ldots,s-1. \tag{16}$$

Let us define the state vector as

$$\alpha_t = (\gamma_t,\tilde{\gamma}_{t+1|t-1},\cdots,\tilde{\gamma}_{t+s-1|t-1})'$$

where s denotes the number of seasons in a period. Then a set of s-equations given by (15) and (16) constitute a state space representation together with the state α_t, and the submatrices T_2 and R_2 in (7) should be replaced by the followings,

$$\tilde{T}_2 = \begin{bmatrix} -1 & 1 & & & \\ -1 & & 1 & & \\ \vdots & & & \ddots & \\ -1 & & & & 1 \\ 0 & & & & 0 \end{bmatrix}, \quad \tilde{R}_2 = \begin{bmatrix} 1 \\ -\Theta \\ -\Theta^2 \\ \vdots \\ -\Theta^{s-1} \end{bmatrix}.$$

Note that the dimension of the state for the usual stochastic dummy seasonal model (4) and the seasonal summation driven by AR (11) is $s-1$ while it increases just by one for the model (13). This predictor-based state space representation will be attributed to Akaike (1974).

4 Real Data Analysis

In this section, we analyze 11 time series with trend and seasonality. First 6 series are Japanese economic time series; consumption in private sector (to be abbreviated to PCSMP, and its span analyzed is from 1980Q1 to 2002Q4), machinery order (MORDER, 1987:04–2002:12), money supply (M2CD, 1980:01–2002:12), new car registration (NEWCAR, 1980:01–2002:12), industrial production (IIP, 1980:01–2002:12), and Tokyo district sales of department stores (TDS, 1980:01–2002:12). The remaining 5 series are taken from the textbook of Harvey (1989); coal, gas and electricity demand of other final users (UKCOAL, UKGAS, UKELEC, 1960Q1–1986Q4), car drivers killed or seriously injured (CDKSI, 1969:01–1982:12), and international airline passengers (AIRLINE, 1949:01–1960:12) of which original source is Box and Jenkins (1976). All the series are log-transformed. In the preliminary analysis, it is verified that AIC statistic corrected by the determinant of Jacobian matrix supports the log-transformation for each series. Because of the log transformation, the scale of the original series sticks around from 5 to 15. Hence, as for the initialization of the Kalman filter, we assume the diagonal matrix for P_0 of which elements are all set to 10^4 in all the models. The first element of the initial state mean is replaced by the sample mean which is computed using the first quarter of the time series. The rest of the initial state element are assumed to be 0.

4.1 Preprocessing

As regards PCSMP and TDS, the effects of the introduction of consumption tax (1989:04) and its rise (1997:04) are removed prior to the model comparison. The unusual increase in March is the consequence of spending rush ahead of the consumption tax introduction or its hike while the atypical depress in April is the counteraction to the March's rush. From the visual inspection of seasonal factors obtained from the BSM, there seems to exist two outliers for PCSMP (1997Q1, 1997Q2) and four outliers for TDS (1989:03, 1989:04, 1997:03, 1997:04). Just for confirmation, each series was analyzed by Web-decomp with the level-2 outlier detection option. Though only 1997Q2 of PCSMP is judged to be outlier-free, we treat the all six observations as outliers.

After the locations of outlier are identified, the preadjustment will be performed in the following manner. Firstly, we create the season-wise series from the original time series. In other words, the observations for specific season (for example, only Q1, only January, etc.) are collected to form another time series. Secondly, for this season-wise series, the data judged as outliers are treated as missing values. We fit first or second order trend model to the season-wise series, and the obtained smoother mean for the missing values replace the outliers. A list of corrections for original data follows. For PCSMP, 72000.8 → 68552.8 (1997Q1), 67146.6 → 68360.0 (1997Q2). As for TDS, 287.883 → 227.255 (1989:03), 175.539 → 203.041 (1989:04), 271.883 → 210.733 (1997:03) and 166.127 → 182.729 (1997:04).

4.2 Overview of Results

The estimation results are summarized in Table 1 and Table 2. The BSM-MA attains the minimum AIC for all series but UKCOAL, and the improvements in AIC are sometimes substantial. In every case, the estimated innovation variance of the seasonal component, σ_ω^2, is larger than those estimated in the BSM and the BSM-MA. It can be interpreted that MA-driven seasonal

Table 1: Estimation results for Japanese macroeconomic data

PCSMP	$\hat{\sigma}^2_\eta$	$\hat{\sigma}^2_\omega$	$\hat{\sigma}^2_\varepsilon$	Φ or Θ	AIC
BSM	0.16×10^{-5}	0.86×10^{-5}	0.35×10^{-4}	—	-508.60
BSM-AR	0.16×10^{-5}	0.53×10^{-5}	0.39×10^{-4}	0.83×10^{-2}	-509.70
BSM-MA	0.15×10^{-5}	0.68×10^{-4}	0.31×10^{-6}	0.84	-523.88
MORDER	$\hat{\sigma}^2_\eta$	$\hat{\sigma}^2_\omega$	$\hat{\sigma}^2_\varepsilon$	Φ or Θ	AIC
BSM	0.31×10^{-4}	0.28×10^{-2}	0.29×10^{-5}	—	-306.73
BSM-AR	0.30×10^{-4}	0.28×10^{-2}	0.29×10^{-5}	0.63×10^{-2}	-303.16
BSM-MA	0.20×10^{-4}	0.40×10^{-2}	0.31×10^{-5}	0.89	-329.02
M2CD	$\hat{\sigma}^2_\eta$	$\hat{\sigma}^2_\omega$	$\hat{\sigma}^2_\varepsilon$	Φ or Θ	AIC
BSM	0.24×10^{-5}	0.88×10^{-6}	0.10×10^{-5}	—	-2144.16
BSM-AR	0.24×10^{-5}	0.85×10^{-6}	0.99×10^{-6}	0.18×10^{-1}	-2154.06
BSM-MA	0.14×10^{-5}	0.28×10^{-5}	0.85×10^{-6}	0.85	-2166.41
NEWCAR	$\hat{\sigma}^2_\eta$	$\hat{\sigma}^2_\omega$	$\hat{\sigma}^2_\varepsilon$	Φ or Θ	AIC
BSM	0.44×10^{-5}	0.31×10^{-3}	0.87×10^{-3}	—	-706.40
BSM-AR	0.44×10^{-5}	0.31×10^{-3}	0.87×10^{-3}	0.23×10^{-6}	-702.40
BSM-MA	0.40×10^{-5}	0.16×10^{-2}	0.26×10^{-5}	0.92	-802.69
IIP	$\hat{\sigma}^2_\eta$	$\hat{\sigma}^2_\omega$	$\hat{\sigma}^2_\varepsilon$	Φ or Θ	AIC
BSM	0.12×10^{-4}	0.85×10^{-4}	0.12×10^{-6}	—	-1324.49
BSM-AR	0.12×10^{-4}	0.86×10^{-4}	0.12×10^{-6}	0.36×10^{-2}	-1320.90
BSM-MA	0.90×10^{-5}	0.11×10^{-3}	0.11×10^{-6}	0.28	-1325.01
TDS	$\hat{\sigma}^2_\eta$	$\hat{\sigma}^2_\omega$	$\hat{\sigma}^2_\varepsilon$	Φ or Θ	AIC
BSM	0.14×10^{-5}	0.11×10^{-3}	0.14×10^{-3}	—	-1113.81
BSM-AR	0.14×10^{-5}	0.86×10^{-4}	0.15×10^{-3}	0.66×10^{-2}	-1126.18
BSM-MA	0.12×10^{-5}	0.38×10^{-3}	0.44×10^{-5}	0.89	-1168.32

summation has successfully brought more flexibility than the BSM and the BSM-AR. Moreover, the BSM-MA is also supported from a predictive point of view, i.e., in terms of minimum AIC.

The estimated AR paremeters of BSM-AR, $\hat{\Phi}$'s, are generally small. At least within the worked examples in this paper, they rarely exceed 0.01. This even helps to increase the power of the seasonal component. A typical result of BSM-AR is seen in the case of PCSMP for example. In the right panel of Figure 2, the log-spectrum of BSM-AR is drawn slightly above the BSM, and the estimated seasonal innovation variance increased from 0.53×10^{-5} to 0.86×10^{-5}. As a result, the plots of their smoother mean are almost indistinguishable from each other. Ozaki (1997) also reports similar results on BSM-AR which he refers to 'dynamic BAYSEA' model.

On the other hand, the estimated MA parameter $\hat{\Theta}$ for PCSMP is 0.84, and the estimated seasonal innovation variance is 0.68×10^{-4}. Accordingly, the level of log-power spectrum in the right panel of Figure 2 is flattened and pushed upward in comparison with those of the BSM and the BSM-AR.

Two panels in the bottom of Figure 3 show the seasonal components estimated by the BSM and the BSM-MA, and the right-upper panel shows their difference. Four panels in Figure 4 show the annual plot of every quarter. We observe that these annual plots of BSM-MA swings around those of the BSM and BSM-AR. Figure 4 clearly exhibits the difference between the

Table 2: Estimation results for UK data in Harvey (1989)

UKCOAL	$\hat{\sigma}^2_\eta$	$\hat{\sigma}^2_\omega$	$\hat{\sigma}^2_\varepsilon$	Φ or Θ	AIC
BSM	0.72×10^{-5}	0.67×10^{-9}	0.17×10^{-1}	—	-35.81
BSM-AR	0.72×10^{-5}	0.33×10^{-7}	0.17×10^{-1}	0.57×10^{-7}	-31.81
BSM-MA	0.74×10^{-5}	0.73×10^{-4}	0.17×10^{-1}	0.98	-31.83
UKGAS	$\hat{\sigma}^2_\eta$	$\hat{\sigma}^2_\omega$	$\hat{\sigma}^2_\varepsilon$	Φ or Θ	AIC
BSM	0.84×10^{-5}	0.41×10^{-2}	0.95×10^{-3}	—	-134.07
BSM-AR	0.85×10^{-5}	0.41×10^{-2}	0.92×10^{-3}	0.17×10^{-2}	-130.11
BSM-MA	0.64×10^{-5}	0.75×10^{-2}	0.20×10^{-5}	0.64	-157.36
UKELEC	$\hat{\sigma}^2_\eta$	$\hat{\sigma}^2_\omega$	$\hat{\sigma}^2_\varepsilon$	Φ or Θ	AIC
BSM	0.88×10^{-5}	0.48×10^{-3}	0.12×10^{-2}	—	-235.75
BSM-AR	0.89×10^{-5}	0.49×10^{-3}	0.12×10^{-2}	0.43×10^{-2}	-232.21
BSM-MA	0.85×10^{-5}	0.24×10^{-2}	0.73×10^{-5}	0.78	-258.52
CDKSI	$\hat{\sigma}^2_\eta$	$\hat{\sigma}^2_\omega$	$\hat{\sigma}^2_\varepsilon$	Φ or Θ	AIC
BSM	0.22×10^{-5}	0.24×10^{-8}	0.44×10^{-2}	—	-213.44
BSM-AR	0.19×10^{-5}	0.11×10^{-3}	0.41×10^{-2}	0.21×10^{-8}	-204.67
BSM-MA	0.21×10^{-5}	0.48×10^{-2}	0.12×10^{-5}	0.99	-302.70
AIRLINE	$\hat{\sigma}^2_\eta$	$\hat{\sigma}^2_\omega$	$\hat{\sigma}^2_\varepsilon$	Φ or Θ	AIC
BSM	0.29×10^{-3}	0.28×10^{-3}	0.14×10^{-5}	—	-391.64
BSM-AR	0.20×10^{-3}	0.29×10^{-3}	0.32×10^{-3}	0.15×10^{-1}	-348.26
BSM-MA	0.88×10^{-5}	0.94×10^{-3}	0.13×10^{-5}	0.94	-445.99

seasonal component of the BSM (thin solid line with symbol +) and of BSM-MA (thick solid line). The BSM and BSM-AR (thin solid line without symbol in Figure 4) produce quite similar results in both figures. Other successful cases exhibit the similar features, but the graphs are omitted for the reason of space.

4.3 Noteworthy Exceptions

Table 1 and 2 show that for most of time series considered here the seasonal summation driven by MA model improves the simple modeling by the BSM, except UKCOAL. What is striking is the decrease in the AIC statistic in the case of CDKSI. From Table 2, AIC of BSM-MA, -302.70 is much smaller than that of the BSM, -213.44. In terms of information criterion, BSM-MA is overwhelmingly superior to the BSM. Nonetheless, once we give a glance at over the right-lower panel of Figure 6, a doubt comes up if we should accept the difference of AIC at its face value. To put it plainly, the seasonal component of BSM-MA appears to be just the subtraction of trend component from the original series. On the other hand, considering the unstable seasonality in the original time series, the seasonal pattern estimated by the BSM looks too regular to be plausible. When it comes to seasonal adjustment, neither the BSM nor BSM-MA gives a satisfactory solution.

Let s_t and $\bar{s}_t$ be the seasonal component derived from the BSM and the BSM-MA respectively. Then $\sigma^* = \sqrt{T^{-1}\sum_{t=1}^{T}(s_t-\bar{s}_t)^2}$ is the standard deviation of the perturbation introduced by MA term. The key feature of the CDKSI case is that σ^* is very large in comparison with the maximum amplitude of the seasonal pattern of the BSM, namely $\max s_t - \min s_t$.

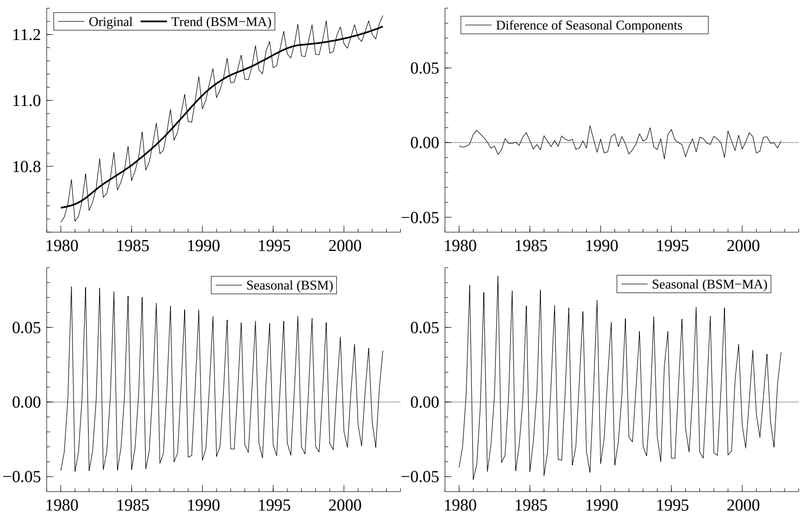

Figure 3: Original plus trend (upper-left), seasonal factors (bottom) and the difference of two seasonal factors (upper-right) for PCSMP.

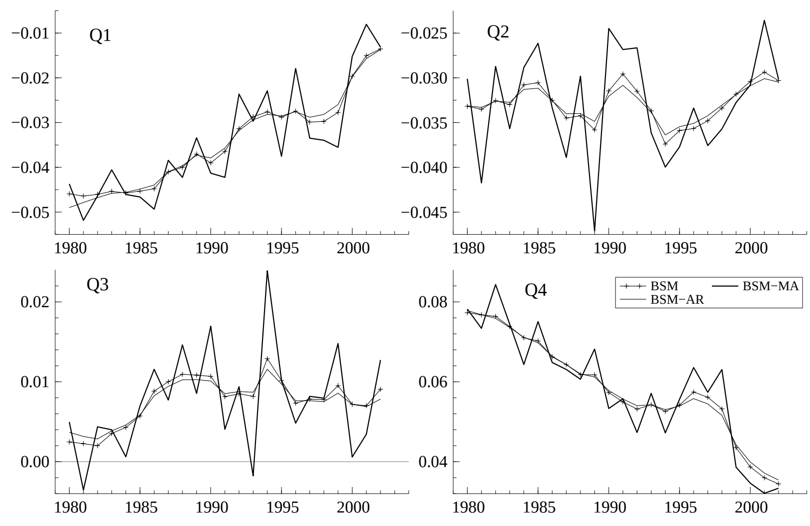

Figure 4: Annual plot of every quarters of the estimated seasonal components in the PCSMP case.

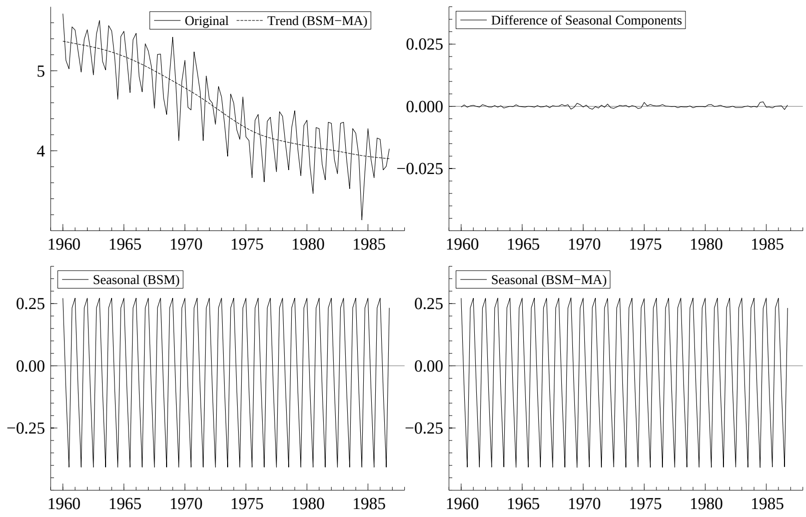

Figure 5: Original plus trend (upper-left), seasonal factors and the difference of two seasonal factors (upper-right) for UKCOAL.

Turning to the case of UKCOAL, we know from Table 2 that the BSM is better than the BSM-MA by the minimum AIC criterion. Figure 5 show the results of the UKCOAL case. The lower two panels exhibit that the seasonal components are indistinguishable from one another , and almost deterministic whichever model to be employed as the seasonal component. The difference of seasonal factors of the BSM and the BSM-MA (the right-upper panel of Figure 5) manifests that the BSM-MA could not introduce any additional variability into the seasonal component. To conclude, the key feature of the UKCOAL case is, σ^* is too small relative to $\max s_t - \min s_t$.

Another significant feature shared by both CDKSI and UKCOAL cases is that the estimated seasonal MA parameter, $\hat{\Theta}$, is extremely close to 1. But from the right-lower panel of Figure 6, we cannot say this is the case of polynomial cancellation because the estimated seasonal pattern is far from deterministic. Looking at Figure 6, an idea easily comes up with us that the wild seasonal component may be regarded as the nearly periodic seasonal pattern laid over the zero mean stationary process. As for UKCOAL, the estimated seasonal patterns are nearly deterministic nevertheless the original series does not exhibit such a steady seasonality. Hence in any case, it is suspected that some important component may be lacking in the model specification.

4.4 Including a Cyclical Component

Upon the observations in the previous subsection, we add a cyclical component to the BSM-MA model and see its impact on the AIC values, on the seasonal moving average parameter (Θ) and on the seasonal pattern. We assume that the cyclical component can be expressed by a finite order (up to 4-th order here) stationary autoregressive process,

$$\psi_t = \sum_{i=j}^{m} \rho_j \psi_{t-j} + \kappa_t,$$

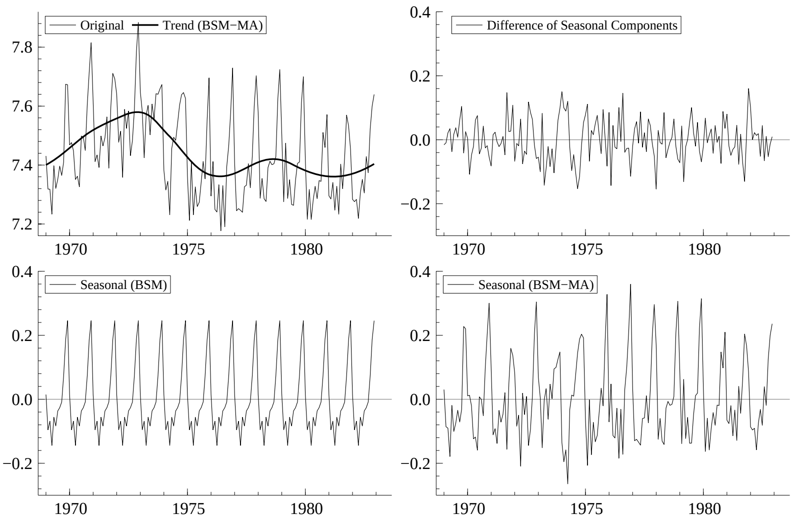

Figure 6: Original plus trend (upper-left), seasonal factors and the difference of two seasonal factors (upper-right) for CDKSI.

and that the time series can be decomposed as

$$y_t = \mu_t + \gamma_t + \psi_t + \varepsilon_t.$$

As the state space representation for this decomposition is obvious, the reader is invited to visit the textbook like Harvey (1989), Kitagawa and Gersch (1996).

For UKCOAL, it turned out that the BSM-MA with the first order stationary AR component attains the minimum AIC, −62.68. This is much smaller than the AIC value for the BSM-MA, −31.83. The estimated trend, seasonal and cyclical component are plotted in the left column of Figure 7. The seasonal MA parameter $\hat{\Theta}$ is 0.99, which is extremely close to unity. As we recall, however, the BSM attained the minimum AIC at least in the analysis without a cyclical component, so we had better compare the BSM and the BSM with a cyclical component model. The AIC of the latter is −66.68 which is also much smaller than the AIC of the BSM, −35.81. To conclude, the best model for UKCOAL is the BSM with a cyclical component expressed by AR(1). If we fit the MA driven seasonal model (13), the estimated $\hat{\Theta}$ is close to 1. This suggests the seasonality in UKCOAL is almost deterministic, and the simple seasonal summation (4) with very small σ_ω^2 suffices.

Now we turn to the CDKSI case. After the model estimation and selection, we find including a second order statinary AR to the BSM-MA improves the AIC value, −302.70 → −304.46. The seasonal MA parameter for CDKSI is substantially diminished from 0.99 to 0.28. We doubt if the MA parameter is really needed, hence we fit the BSM including a AR(2) component. Contrary to our expectation, the AIC statistic of the model is −269.39, which is inferior to the BSM-MA with a cyclical component. Thus, the best model among the models we tried here is the BSM-MA with a cyclical component model expressed by a statinary AR of order 2. Three panels in the right column of Figure 7 show the estimated components for CDKSI. Paralleling the AR component with the difference of seasonal components in Figure 6, the fluctuations brought by the MA process are almost captured by the cyclical component.

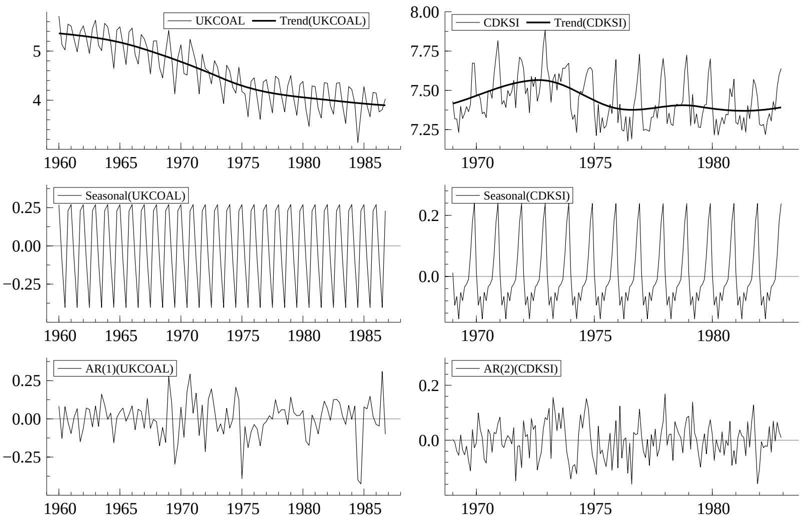

Figure 7: Original plus trend (upper), seasonal (middle) and AR component (bottom) for UK-COAL (left panels) and CDKSI (right panels).

4.5 A Graphical Representation

In the subsection 4.3, it is remarked that the appropriateness of the flexibility brought by MA term should be determined in connection with the range of seasonal pattern of the time series. Thus we introduce a simple measure on the pertinence of the seasonal component model, and propose a graphical representation. Let $R = \max s_t - \min s_t$. What appears to be essential is the ratio, R/σ^*. Considering that R is a sort of range and σ^* is the standard deviation. It seems more natural to consider the length of interval, such as $\pm 2\sigma^*$ or $\pm 3\sigma^*$ for example. Here we adopt $\pm 3\sigma^*$ interval, and define the following quantity to measure the impact of the fluctuation brought by MA term on the seasonal pattern,

$$M = \log_{10}\{R/6\sigma^*\}.$$

If $\bar{s}_t$ is very wild, then the argument of the log function will be close to unity, so M is close to 0. If s_t and $\bar{s}_t$ are alike, then small value of σ^* will lead to large M. (If σ^* happens to be zero, then we discard the measure M. Such a case does not interest us at all because the MA term is not effectively working and the BSM is obviously better.) As is already pointed out in section 4.3, another key quantity is $\hat{\Theta}$, the estimated MA parameter. Therefore, the 2-dimensional plot of $(M, \hat{\Theta})$ is expected to give some information on the modeling of the time series of interest.

Figure 8 shows the graphical layout of the BSM-MA models applied to the 11 time series in this paper. The horizontal axis denotes the measure M defined above, and the vertical axis indicates the seasonal MA parameter, Θ. Two points connected with an arrow mean that the AIC statistic is improved by including a cyclical component, and subsequently the graphical layout of (M,Θ) is changed. It is striking that $\hat{\Theta}$'s are diminished and M's are centered around 1.00 ± 0.25 after the cyclical component is added to the model.

We observe that the two exceptional cases (CDKSI and UKCOAL) are located at the upper-left and the upper-right in this model map. $M \approx 0$ means the seasonal variability increased by the BSM-MA model is almost comparable with the seasonal range of original time series. On the other hand, large M indicates that the BSM-MA cannot introduce any flexible seasonal

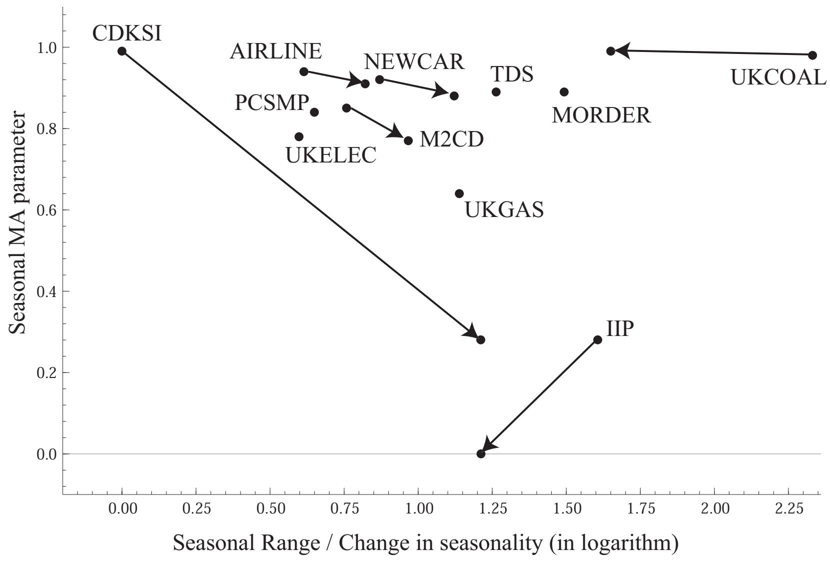

Figure 8: Graphical layout of the estimated models.

pattern compared to the BSM. In some cases this suggests that the seasonality for the time series of interest is almost deterministic, which sounds reasonable for the UKCOAL case. From our empirical analysis, it is inferred that $\hat{\Theta} \approx 1$ suggests the 'cancellation' is occurring on the seasonal component model (e.g. UKCOAL), or the possible misspecification as in the CDKSI case.

5 Conclusion

This article proposed a parsimonious modeling of flexible seasonality within a framework of structural time series models. The basic idea is to drive the seasonal summation by a moving average process with just one parameter, which has been referred to the BSM-MA throughout this article. A state space representation for the model is also given. Compared to the simple seasonal summation (BSM) and the AR-driven seasonal summation (BSM-AR), the BSM-MA attains the minimum AIC in 10 out of 11 cases. In all successful cases, the estimated seasonal innovation variance is larger than those estimated by the BSM and BSM-AR, which leads to the increase of the power spectrum of the seasonal component. Annual plot of every quarters or months reveals the wiggly movement introduced by moving average terms. A close examination of the UKCOAL and CDKSI cases provides us some information. Adequacy of the additional seasonal perturbation brought by the BSM-MA depends on how big it is relative to the range of seasonal pattern. Hence the log of the ratio of the maximum amplitude of seasonal pattern to the interval length of the ± 3-standard deviations obtained by the seasonal difference between the BSM and the BSM-MA is introduced as a measure, M. Both too small and too large M suggests the possible misspecification. Using M and Θ, a graphical representation for the estimated models is also proposed, which serves to mark out the seemingly unsuccessful cases. Even for such cases, the decomposition including a cyclical component is proved to amend the existing models from as is shown in our empirical analysis.

Acknowledgement
This work was carried out partly under Grant-in-Aid for Young Scientists (B) (No. 14780177) sponsored by the Ministry of Education, Culture, Sports, Science and Technology, and also under Grant-in-Aid for Scientific Research (C) (No. 17500189) sponsored by Japan Society for the Promotion of Science.

References

[1] Akaike, H. (1973). Information theory and extension of the maximum likelihood principle. In *Second International Symposium of Information Theory*, N. B. Petrov and F. Czaki (eds.), 267–281, Budapest: Akademiai Kiado.

[2] Akaike, H. (1974). Markovian representation of stochastic processes and its application to the analysis of autoregressive moving average processes, *Ann. Inst. Statist. Math.*, Vol. 26, 363–387.

[3] Akaike, H.(1980). Seasonal adjustment by a Bayesian modeling, *J. Time Series Anal.*, Vol. 1, 1–13.

[4] Akaike, H., Ozaki, T., Ishiguro, M., Ogata, Y., Kitagawa, G., Tamura, Y.-H., Arahata, E., Katsura, K. and Tamura, Y. (1985). TIMSAC-84 Part 1, *Computer Science Monographs No. 23*, The Institute of Statistical Mathematics, Tokyo.

[5] Anderson, B. D. O. and Moore, J. B. (1979). *Optimal filtering*, Prentice-Hall, New Jersey.

[6] Box, G. E. P., Hillmer, S. C. and Tiao, G. C. (1979). Analysis and modelling of seasonal time series, *NBER-Census Conference on Seasonal Analysis of Economic Time Series*, ed. Arnold Zellner, Washington D.C., 309–334.

[7] Box, G. E. P. and Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control, Revised ed.*, Holden-Day.

[8] Burman, J. P. (1980). Seasonal adjustment by signal extraction, *Journal of the Royal Statistical Society, Series A*, Vol. 143, 321–327.

[9] Burridge, P. and Wallis, K. F. (1984). Unobserved-components models for seasonal adjustment filters, *J. Bus. Econ. Stat.*, Vol. 2, 350–359.

[10] Cleveland, W. P. and Tiao, G. C. (1976). Decomposition of seasonal time series: a model for the X-11 program, *J. Amer. Stat. Assoc.*, Vol. 71, 581–587.

[11] Gersch, W. M. and Kitagawa, G. (1983). The prediction of time series with trends and seasonalities, *J. Bus. Econ. Stat.*, Vol. 1, 253–264.

[12] Hannan, E. J., Terrell, R. D. and Tuckwell, N. (1970). The seasonal adjustment of economic time series, *International Economic Review*, Vol. 11, 24–52.

[13] Harrison, P. J. and Stevens, S. C. (1971). A Bayesian approach to short-term forecasting, *Operational Research Quarterly*, Vol. 22, 341–362.

[14] Harvey, A. C. (1984). A unified view of statistical forecasting procedures (with discussion), *Journal of Forecasting*, Vol. 3, 245–283.

[15] Harvey, A. C. (1985). Trend and cycles in macroeconomic time series, *Journal of Business and Economic Statistics*, Vol. 3, 216–227.

[16] Harvey, A. C. (1989). *Forecasting, structural time series models and the Kalman filter*, Cambridge University Press, Victoria, Australia.

[17] Harvey, A. C. and Todd, P. H. J. (1983). Forecasting economic time series with structural and Box-Jenkins models (with discussion), *Journal of Business and Economic Statistics*, Vol. 1, 299–315.

[18] Hillmer S. C. and Tiao, G. C. (1982). An ARIMA-model-based approach to seasonal adjustment, *J. Amer. Stat. Assoc.*, Vol. 77, 63–70.

[19] Kitagawa, G. (1981). A nonstationary time series model and its fitting by a recursive filter, *J. Time Series Anal.*, Vol. 2, 103–116.

[20] Kitagawa, G. (1987). Non-Gaussian state space modeling of non stationary time series (with discussion), *J. Amer. Stat. Assoc.*, Vol. 82, 1032–1063.

[21] Kitagawa, G. (1989). Non-Gaussian seasonal adjustment, *Computers & Mathematics with Applications*, Vol. 18, 503–514.

[22] Kitagawa, G. (1994). The two-filter formula for smoothing and an implementation of the Gaussian sum smoother, *Ann. Inst. Statist. Math.*, Vol. 46, 605–623.

[23] Kitagawa, G. (1996). Monte carlo filter and smoother for non-Gaussian nonlinear state space models, *Journal of Computational and Graphical Statistics*, Vol. 5, 1–25.

[24] Kitagawa, G. and Gersch, W. M. (1984). A smoothness-prior state space modeling of time series with trend and seasonality, *J. Amer. Stat. Assoc.*, Vol. 79, 378–389.

[25] Kitagawa, G. and Gersch, W. M. (1996). *Smoothness priors analysis of time series, Lecture Notes in Statistics 116*, Springer-Verlag, New York.

[26] Koopman, S. J., Harvey, A. C., Doornik, J. A. and Shephard, N. (2000). *STAMP: structural time series analyser, modeller and predictor*, Timberlake Consultants Ltd, Harrow.

[27] Ozaki, T. (1997). Dynamic X11 model and nonlinear seasonal adjustment II: numerical examples and discussion (in Japanese with English abstract), *Proceedings of the Institute of Statistical Mathematics*, Vol. 45, No. 2, 287–300.

[28] Ozaki, T. and Thomson, P. J. (1992). A dynamical system approach to X-11 type seasonal adjustment, Research Memorandum No. 498, the Institute of Statistical Mathematics, Tokyo.

[29] Sato, S. (1997). Introduction to “Web-Decomp" — Seasonal Adjustment System on WWW — (in Japanese with English abstracts), *Proceedings of the Institute of Statistical Mathematics*, Vol. 45, No. 2, 233–243. See also `http://ssnt.ism.ac.jp/inets2/title.html`.

[30] Shephard, N. and Pitt, M. (1997). Likelihood analysis of non-Gaussian measurement time series, *Biometrika*, Vol. 84, 653–667.

[31] Wallis, K. F. (1982). Seasonal adjustment and revision of current data: linear filters for the X-11 method, *Journal of the Royal Statistical Society, Series A*, Vol. 145, 74–85.

[32] West, M. and Harrison, P. J. (1986). Monitoring and adaptation in Bayesian forecasting models, *J. Amer. Stat. Assoc.*, Vol. 81, 741–750.

[33] West, M., Harrison, P. J. and Migon, H. S. (1985). Dynamic generalized linear models and Bayesian forecasting (with discussion), *J. Amer. Stat. Assoc.*, Vol. 80, 73–97.