

Journal of Official Statistics, Vol. 30, No. 4, 2014, pp. 811-838, http://dx.doi.org/10.2478/JOS-2014-0049

Estimation of Mean Squared Error of X-11-ARIMA and Other Estimators of Time Series Components

Danny Pfeffermann¹ and Michail Sverchkov²

This article considers the familiar but very important problem of how to estimate the mean squared error (MSE) of seasonally adjusted and trend estimators produced by X-11-ARIMA or other decomposition methods. The MSE estimators are obtained by defining the unknown target components such as the trend and seasonal effects to be the hypothetical X-11 estimates of them that would be obtained if there were no sampling errors and the series were sufficiently long to allow the use of the symmetric filters embedded in the programme, which are time invariant. This definition of the component series conforms to the classical definition of the target parameters in design-based survey sampling theory, so that users should find it comfortable to adjust to this definition. The performance of the MSE estimators is assessed by a simulation study and by application to real series obtained from an establishment survey carried out by the Bureau of Labor Statistics in the U.S.A.

Key words: Bias correction; canonical decomposition; seasonal adjustment; state-space model; survey sampling; trend; X-13ARIMA-SEATS.

1. Introduction

In this article, we consider estimation of the mean squared error (MSE) of seasonally adjusted and trend estimators produced by X-11-ARIMA or other decomposition methods. In particular, we compare the MSE of estimators obtained by application of X-11-ARIMA with the MSE of estimators obtained by fitting state-space models that account for correlated sampling errors. We define the target seasonally adjusted and trend components to be the hypothetical X-11 estimates of those that would be obtained in the absence of sampling errors and if the time series under consideration was sufficiently long for application of the symmetric filters embedded in the original X-11 procedure, which are time invariant. This definition of the component series conforms to the classical definition of target finite population parameters in design-based survey sampling theory. In fact, in one variant of the proposed definition, the target components are shown to be linear combinations of finite

¹ Central Bureau of Statistics, Israel, Hebrew University of Jerusalem, Israel and University of Southampton, Southampton SO17 1BJ, UK. Email: msdanny@soton.ac.uk

² Bureau of Labor Statistics, 2 Massachusetts Avenue, NE, Suite 1950, Washington DC 20212, U.S.A. Email: Sverchkov.Michael@bls.gov

Acknowledgments: We are grateful to the Associate Editor and two reviewers for many excellent and constructive comments. We also thank Brian Monsell from the Census Bureau in the U.S.A. for writing a special module within X-13ARIMA-SEATS, which permits forecasting any number of signal components' values. The opinions expressed in this article are those of the authors and do not necessarily represent the policies of the Central Bureau of Statistics in Israel or the Bureau of Labor Statistics in the U.S.A.

population means or totals. The MSE of X-11-ARIMA and state-space model estimators are defined with respect to this definition.

We estimate the MSE by conditioning on the target components, thereby accounting for possible conditional bias in estimating them. The results are illustrated by use of simulated series and by application to real series obtained from an establishment survey carried out by the Bureau of Labor Statistics (BLS) in the U.S.A. The latter results also contrast the performance of our proposed MSE estimators with estimators proposed by Bell and Kramer (1999).

2. Target Components, Bias and MSE of X-11-ARIMA Estimators

2.1. Target Components

We begin with the usual notion that an economic time series, Y_t ; t = 1, 2, ... can be decomposed into a trend or trend-cycle component T_t , a seasonal component S_t , and an irregular component I_t ; $Y_t = T_t + S_t + I_t$. Here we consider the additive decomposition, but the results of this article can be generalized to the multiplicative decomposition $Y_t = T_t \times S_t \times I_t$ by applying the log transformation and employing similar considerations to those in Pfeffermann et al. (1995). In practice, it is often the case that the series Y_t is unobserved and the available series consists of sample estimates, y_t , obtained from repeated sample surveys. Consequently, the series y_t can be expressed as the sum of the true population value, Y_t , and a sampling error, ε_t . More generally, the observed series can be viewed as the sum of a signal, G_t , and an error, e_t ; $y_t = G_t + e_t$, where the signal, and hence the error, may be defined in two alternative ways:

- *GE*1. $G_t = T_t + S_t$, $e_t = I_t + \varepsilon_t$. In this case e_t is the combined error of the time series irregular term and the sampling error (Pfeffermann 1994);
- *GE2.* $G_t = T_t + S_t + I_t$, $e_t = \varepsilon_t$. In this case the irregular term is part of the signal, and e_t is the sampling error (Bell and Kramer 1999)

We assume without loss of generality that the series started at time $-\infty < t_{start} < 1$, but y_t is only observed for the time points t = 1, ..., N, such that

$$y_t = G_t + e_t, \quad t = \underbrace{t_{start}, \dots, 0}_{unobserved}, \underbrace{1, \dots, N}_{y_t \ observed}, \underbrace{N+1, \dots, \infty}_{unobserved}$$
 (1)

It is assumed also that under both definitions of the signal, e_t is independent of $\mathbf{G} = \{G_t, t = t_{start}, \ldots, \infty\}$ for all t, with $E(e_t) = 0$, $Var(e_t) < \infty$, although in practice the sampling error, and in particular the variance of the sampling error, sometimes depends on the magnitude of the signal.

The X-11-ARIMA program first forecasts and backcasts the time series under consideration based on an ARIMA model fitted to the observed series, and then applies a sequence of moving averages (linear filters) to the series augmented by the forecasts and backcasts. It follows that the X-11-ARIMA estimators of the trend and the seasonal components can be approximated as,

$$\hat{T}_t = \sum_{k=-(t-1)}^{N-t} w_{kt}^T y_{t+k}, \quad \hat{S}_t = \sum_{k=-(t-1)}^{N-t} w_{kt}^S y_{t+k},$$
(2)

where the coefficients $\{w_{kt}^T\}$, $\{w_{kt}^S\}$ are defined in general by the program options for the observed time interval $t = 1, \ldots, N$, by the ARIMA model used to forecast and backcast the series and by the number of backcasts and forecasts. However, at the central part of the series, the filters in (2) are time-invariant and symmetric; $w_{kt}^T = w_k^T$, $w_{-k}^T = w_k^T$ for $a_T < t \le N - a_T$; $w_{kt}^S = w_k^S$, $w_{-k}^S = w_k^S$ for $a_S < t \le N - a_S$, where a_T, a_S are defined by the X-11-ARIMA program options. The length of the symmetric filters is thus $2a_T + 1$ ($2a_S + 1$). For example, for the default X-11-ARIMA options, $a_S = 84$, $a_T = 90$, but a_S , for example, may be as low as 70 or as high as 149 when using other options. Note that in the central part of the series the X-11-ARIMA estimators are the same as the X-11 estimators with no ARIMA extrapolations, such that the symmetric filters only depend on the X-11 program options and not on the ARIMA extrapolations.

Remark 1. The use of X-11-ARIMA also involves 'non-linear' operations such as the identification and estimation of ARIMA models used for forecasting and backcasting the original series, and the identification and gradual replacement of extreme observations. We assume that the time series under consideration is already modified for extreme values, thus robustifying the variance estimates described in Subsection 2.3. As illustrated in Pfeffermann et al. (1995) and Pfeffermann et al. (2000), the effects of the identification and nonlinear estimation of ARIMA models are generally minor.

Definition 1. Assuming $t_{start} < \min(-a_T, -a_S)$ and following Bell and Kramer (1999), we define the trend component at time t to be $T_t^{X11} = \sum_{k=-a_T}^{a_T} w_k^T G_{t+k}$. Analogously, the seasonal component is defined as $S_t^{X11} = \sum_{k=-a_S}^{a_S} w_k^S G_{t+k}$. The target components T_t^{X11} and S_t^{X11} are thus the hypothetical components that would be obtained by application of the X-11 symmetric filters to the signal **G** at time point t, $t = 1, \ldots, N$. It follows therefore that the observed series may be decomposed as the sum of the 'X-11-trend', T_t^{X11} , the 'X-11-seasonal component', S_t^{X11} , and the 'X-11 error', $e_t^{X11} = y_t - T_t^{X11} - S_t^{X11}$:

$$y_t = T_t^{X11} + S_t^{X11} + e_t^{X11}.$$
(3)

Result 1. For $a_T < t \le N - a_T$, $T_t^{X11} = E(\hat{T}_t | \mathbf{G})$ and for $a_S < t \le N - a_S$, $S_t^{X11} = E(\hat{S}_t | \mathbf{G})$, where \hat{T}_t, \hat{S}_t are the X-11-ARIMA estimators defined in (2) and the expectation is taken over the distribution of the errors $\{e_t, t = 1, \ldots, N\}$, with the signal \mathbf{G} held fixed. It follows therefore from our definition that in the central part of the series, the X-11-ARIMA estimators \hat{T}_t, \hat{S}_t of the trend and the seasonal component are unbiased. (As noted before, we assume that the observed series is already modified for extreme values. The identification and estimation of ARIMA models are irrelevant at the center of the series.)

Remark 2. For X-11 filters $a_T > a_S$ because the final trend filter is applied after the final seasonal and seasonally adjusted components are computed. Thus, max $(a_T, a_S) = a_T$.

Remark 3 We define the trend and seasonal components to be the (hypothetical) outputs that would be obtained when applying the symmetric filters to the signal, since the filters at the non-central parts of the series are asymmetric and depend on the time points with data.

In particular, the filters applied for a time point $t > N - a_T$ change every time that a new observation is added to the series until $t \le N - a_T$, when the symmetric filter is applied. As mentioned before, the decomposition (3) has been used by Bell and Kramer (1999) with the error defined by the sampling error, such that the irregular term is part of the signal; $G_t = T_t + S_t + I_t$ (Definition GE2). See Subsection 2.5 for details of their approach. Note that with this definition, the target values are just linear combinations of the unadjusted population values of the series, which in most cases are finite population means or totals, in line with classical survey sampling theory.

2.2. Conditional Bias and MSE of X-11-ARIMA Estimators

The conditional bias, variance and MSE of the X-11-ARIMA estimator of the trend with respect to the decomposition (3), given the signal, are as follows:

$$Bias(\hat{T}_t | \mathbf{G}) = E[(\hat{T}_t - T_t^{X11}) | \mathbf{G}] = \sum_{k=-(t-1)}^{N-t} w_{kt}^T G_{t+k} - \sum_{k=-a_T}^{a_T} w_k^T G_{t+k}.$$
 (4)

$$Var[\hat{T}_{t}|\mathbf{G}] = E\left\{ \left[\sum_{k=-(t-1)}^{N-t} w_{kt}^{T} y_{t+k} - E\left(\sum_{k=-(t-1)}^{N-t} w_{kt}^{T} y_{t+k} \middle| \mathbf{G} \right) \right]^{2} \middle| \mathbf{G} \right\}$$

$$= E\left\{ \left[\sum_{k=-(t-1)}^{N-t} w_{kt}^{T} (y_{t+k} - G_{t+k}) \right]^{2} \middle| \mathbf{G} \right\} = E\left(\sum_{k=-(t-1)}^{N-t} w_{kt}^{T} e_{t+k} \right)^{2}$$

$$MSE(\hat{T}_{t}|\mathbf{G}) = E\left[\left(\hat{T}_{t} - T_{t}^{X11} \right)^{2} \middle| \mathbf{G} \right] = Var(\hat{T}_{t}|\mathbf{G}) + Bias^{2}(\hat{T}_{t}|\mathbf{G}).$$
(5)
(6)

Similar expressions hold for the seasonal and seasonally adjusted estimators.

Expressions (4)–(6) are general and apply to any linear estimator with arbitrary coefficients $\{w_{kt}^T\}$, as defined by the X-11-ARIMA options, the ARIMA model used for extrapolations and the number of forecasts and backcasts. In fact, as will be shown in Section 3, the Expressions (4)–(6) hold equally for other linear filters, not necessarily embedded in the X-11-ARIMA program. In the following sections we discuss ways of estimating the MSE in (6).

2.3. Variance Estimation

Under Definition GE2 of the signal and error in Subsection 2.1, $e_t = \varepsilon_t$ is the sampling error, and by (5),

$$Var(\hat{T}_t|\mathbf{G}) = E\left(\sum_{k=-(t-1)}^{N-t} w_{kt}^T \varepsilon_{t+k}\right)^2 = \sum_k \sum_l w_{kt}^T w_{lt}^T Cov(\varepsilon_{t+k}, \varepsilon_{t+l}).$$

Similar expressions apply when estimating the seasonal or the seasonally adjusted component. We assume the availability of estimates of the variances and covariances of the sampling errors, which enables estimation of the variance $Var(\hat{T}_t|\mathbf{G})$ and the variance of any other component estimator.

Next, consider the estimation of the variance under Definition GE1 of the signal and error, by which $e_t = I_t + \varepsilon_t$. By (5), the variance of the X-11-ARIMA trend estimator is in this case a linear combination of the covariances $v_{tm} = Cov(e_t, e_m)$, t, m = 1, ..., N. Following Pfeffermann (1994) and Pfeffermann and Scott (1997), let $R_t = y_t - \hat{S}_t - \hat{T}_t =$ $\sum_{k=-(t-1)}^{N-t} w_{kt}^R y_{t+k}$ define the linear approximation of the X-11-ARIMA residual term at time *t*, where $w_{0t}^R = 1 - w_{0t}^S - w_{0t}^T$ and $w_{kt}^R = -w_{kt}^S - w_{kt}^T$ for $k \neq 0$. Then,

$$Var(R_{t}|\mathbf{G}) = E\left\{ \left[\sum_{k=-(t-1)}^{N-t} w_{kt}^{R}(y_{t+k} - E(y_{t+k}|\mathbf{G})) \right]^{2} |\mathbf{G}\right\} = Var\left(\sum_{k=-(t-1)}^{N-t} w_{kt}^{R}e_{t+k} \right),$$

$$Cov(R_{t}, R_{m}|\mathbf{G}) = Cov\left[\sum_{k=-(t-1)}^{N-t} w_{kt}^{R}e_{t+k}, \sum_{l=-(m-1)}^{N-m} w_{lm}^{R}e_{m+l} \right] = \sum_{k} \sum_{l} w_{kt}^{R}w_{lm}^{R}Cov(e_{t+k}, e_{m+l}).$$

(7)

The residuals R_t are not stationary because of the use of asymmetric filters towards the two ends of the series. However, let $U(m) = \frac{1}{N-m} \sum_{t=1}^{N-m} Cov(R_t R_{t-m}), m=0,...,N-1$, and suppose that the errors $e_t = I_t + \varepsilon_t$ are stationary (see Remark 4 below). Then, by (7), the vector **U** of the means U(m) and the vector **V** of the covariances $V_k = Cov(e_t, e_{t+k}) = Cov(I_t + \varepsilon_t, I_{t+k} + \varepsilon_{t+k}), k = 0,...,N-1$ are related by the system of linear equations,

$$\mathbf{U} = D\mathbf{V},\tag{8}$$

where the matrix D is defined by the known weights $\{w_{kt}^R\}$. Since the X-11-ARIMA residuals are known for every $t=1,\ldots,N$, one may estimate U(m) by $\tilde{U}(m) = \frac{1}{N-m} \sum_{t=1}^{N-m} R_t R_{t-m}$. Substituting $\tilde{U}(m)$ for U(m) in (8) enables estimation of V by solving the resulting equations; see Pfeffermann (1994) and Pfeffermann and Scott (1997). Note that the use of (8) does not require the availability of estimates of the variances and covariances of the sampling errors. However, the estimators obtained in this way can be very unstable since the number of unknown variances and covariances generally equals the number of equations. A possible solution to this problem is to assume that the covariances V_k are negligible beyond some lag C and set them to zero, and then solve the reduced set of equations for V_0, \ldots, V_C . This is a mild ergodicity condition assumed for the series e_i . Note that with this assumption it is no longer necessary to consider the estimates $\tilde{U}(m)$ for large m. Additionally, when estimates for the autocovariances of the sampling errors are available, they can be substituted into the vector \mathbf{V} and taken as known, in which case one only needs to estimate the unknown variance and covariances of the time series irregular terms, I_i . This reduces the number of unknown covariances and hence the number of equations very drastically. Note that all these procedures are basically 'model free'. See Chen et al. (2003) for a different approach to estimating U and V. Bell and Kramer (1999) consider model-based estimation of the variance and covariances of the sampling errors.

Remark 4. The linear equations in (8) can easily be extended to the case of heteroscedastic sampling errors for which $V_{tk} = Cov(e_t, e_{t+k}) = L_{tk}V_k$ with known

coefficients L_{tk} . Another potential modification consists of utilizing all (or most of) the equations in (8), and estimating V_0, \ldots, V_C by a discounted least-squares procedure.

2.4. Bias and MSE Estimation

Estimation of the conditional bias of the estimator \hat{T}_t (or any other linear estimator) given the signal, and hence the conditional MSE is more involved. We propose to estimate the bias by estimating the signal and then substituting the estimate in the right hand side of the bias expression (4). A possible way of estimating the signal is by application of the programme X-13ARIMA-SEATS (X-13A-S Reference Manual, Version 0.1 2013). This program is now in common use in many statistical bureaus around the world (replacing X-12-ARIMA). The programme enables to extract the models holding for the trend and the seasonal effects from the ARIMA model fitted to the observed series, and use these models in order to estimate the signal within the observation period, and forecast and backcast the signal for a_T time points with no observations. Denote by \hat{G}_t the estimated signal for time t, including before or after times $1, \ldots, N$. The bias is estimated as,

$$Bi\hat{a}s[\hat{T}_{t}|\mathbf{G}] = \hat{E}\left[\left(\hat{T}_{t} - T_{t}^{X11}\right)|\mathbf{G}\right] = \sum_{k=-(t-1)}^{N-t} w_{kt}^{T}\hat{G}_{t+k} - \sum_{k=-a_{T}}^{a_{T}} w_{k}^{T}\hat{G}_{t+k}, \qquad (9)$$

$$t = 1, \dots, N.$$

Use a similar expression for estimating the bias of the seasonally adjusted estimator.

The SEATS models are obtained by application of canonical signal extraction and under correct model specification, the estimators have Minimum MSE (MMSE) (Hilmer and Tiao, 1982).

Remark 5. Wecker (1979) noted that the MMSE signal estimator has a different spectrum from the true signal and proposed another estimator which preserves the spectrum of the true signal. Application of this proposal requires external information and the loss in MSE compared to the use of the MMSE estimator can be large. We do not consider this estimator in the present article.

A simpler way of estimating the signal, which can be implemented by application of the X-11-ARIMA programme (or within X-13ARIMA-SEATS), consists of the following two steps:

- (a) Use the ARIMA model fitted by X-11-ARIMA to the original series to augment the series with a_T forecasts and backcasts;
- (b) Estimate the signal of the augmented series as,

$$\hat{G}_t = \sum_{k=-(t^1 a_T - 1)}^{N + a_T - t} w_{kt}^{G, aug} y_{t+k}^{aug}, \quad t = -a_T + 1, \dots, N + a_T.$$
(10)

where $y_t^{aug} = y_t$ if y_t is observed, y_t^{aug} is the forecasted (backcasted) value if y_t is not observed and $w_{kt}^{G,aug} = w_{kt}^{T,aug} + w_{kt}^{S,aug}$, with $w_{kt}^{T,aug}, w_{kt}^{S,aug}$ defining the X-11 weights for

the longer (augmented) series. Substituting the estimated signal (10) into (4) yields the trend bias estimator, similarly to (9). Similar expressions hold for the seasonal and seasonally adjusted estimators.

Remark 6. The difference between the two methods of estimating and predicting the signal described above lies in the linear filters applied to the original series. The first method uses the optimal filters for extracting the trend and seasonal component under the ARIMA model fitted to the series. The second method uses the ARIMA model for backcasting and forecasting the original series (for given model coefficients, the backcasts and forecasts are linear combinations of the original series), and then uses the original X-11 filters for estimating the trend and seasonals of the augmented series. The resulting filters generally differ from the optimal filters.

Having estimated the conditional variance and bias, a conservative estimator of the conditional MSE defined by (6) is obtained by adding the variance estimator to the square of the bias, i.e.,

$$M\hat{S}E(\hat{T}_t|\mathbf{G}) = V\hat{a}r(\hat{T}_t|\mathbf{G}) + Bi\hat{a}s^2(\hat{T}_t|\mathbf{G}).$$
(11)

The estimator in (11) is conservative since $E[Bi\hat{a}s^2(\hat{T}_t|\mathbf{G})|\mathbf{G}] = \{E[Bi\hat{a}s(\hat{T}_t|\mathbf{G})]|\mathbf{G}\}^2 + Var[Bi\hat{a}s(\hat{T}_t|\mathbf{G})|\mathbf{G}] > \{E[Bi\hat{a}s(\hat{T}_t|\mathbf{G})]|\mathbf{G}\}^2$. The overestimation of the MSE can be corrected by subtracting an estimate of $Var[Bi\hat{a}s(\hat{T}_t|\mathbf{G})|\mathbf{G}]$. Note that $Bi\hat{a}s(\hat{T}_t|\mathbf{G})$ is a linear combination of the signal estimates, \hat{G}_t , which in turn are linear combinations of the observed series, y_t . Thus, $Bi\hat{a}s(\hat{T}_t|\mathbf{G})$ is a linear combination of the signal estimated similarly to the estimation of $Var[\hat{T}_t|\mathbf{G}]$ discussed in Subsection 2.3. The weights defining $Bi\hat{a}s(\hat{T}_t|\mathbf{G})$ can be obtained similarly to Burck and Sverchkov (2001) and Findley and Martin (2006) (See Section 3).

Remark 7. The procedure proposed for estimating the bias and MSE of the X-11-ARIMA estimators raises two valid questions:

i) The predictors of the signal many years ahead, required for estimating the MSE of the estimators of the component series may be severely biased for time points far away from the last time point N with an observation, because of possible changes in the behavior of the signal over time. So how can one rely on these predictors? To answer this question, note first that even if the signal predictors are biased (given the true signal), the trend bias estimator in (9) may still be unbiased or only have a small bias. For example, if $E[(\hat{G}_t - G_t)|\mathbf{G}] = constant$ for all t, the bias estimator (9) is unbiased for the true bias since $\sum_{k=-(t-1)}^{N-t} w_{kt}^T = \sum_{k=-a_T}^{a_T} w_k^T = 1$. The same holds when estimating the bias of the seasonally adjusted estimators (SAE). While this may not be a realistic scenario, what is more important is that the weights of the symmetric filters, used to predict the trend and the SAE decay to zero very fast when moving away from the time point of interest, so that even large biases of the predictors of the signal for time points far away from the last time with an observation may have little effect on the bias of the estimator of the bias of the trend or the SAE. Figure 1 shows the central weights of the trend filters used in our simulation study described in Section 4. The plot of the Basic Structural Model (BSM) filter weights looks like



Fig. 1. Central weights applied to the signal for predicting the trend under X-11 and under the basic structural model (BSM, Section 3)

a trend filter for a nonseasonal series because there seems to be no seasonal pattern to the weights, which is counterintuitive. We have no explanation to this behavior of the weights, but all our checks show that they are correct.

ii) If we believe that we have good predictors of the signal and hence good estimators of our target trend, why not use these estimates in the first place instead of using the X-11-ARIMA estimates of the trend? The answer to this question is simple. Our aim in this article is not to propose new trend or seasonally adjusted estimators. In fact, the model-based predictors of the trend and seasonal component that we use to estimate the bias are produced by one of the modules of X-13ARIMA-SEATS, following the pioneering work of Gómez and Maravall (1996). Rather, our aim is to develop a method of estimating the conditional MSE of linear estimators such as the X-11-ARIMA estimators, which are in common use. We may refer to our method of bias estimation as 'model-based'.

Remark 8. When the signal is estimated by the MMSE estimator under the models extracted for the trend and the seasonal component, the estimator of the signal coincides with the conditional expectation of the signal, given the observed series. In this case the bias estimator (9) is the conditional expectation of the bias over all possible realizations of the signal given the observed series.

2.5. Comparison With the Method of Bell and Kramer

As noted before, Bell and Kramer (1999) use a similar definition of the target components. The authors estimate these components by augmenting the observed series

with a_T MMSE forecasts and backcasts under an appropriate ARIMA model, such that the symmetric X-11 filters can be applied to the augmented series at every time point *t* with an observation. The trend estimator, for example, can be written then as $\hat{T}_t^{BK} = \sum_{k=-a_T}^{a_T} w_k^T y_{t+k}^*$, where $y_{t+k}^* = y_{t+k}$ if y_{t+k} is observed $(1 \le t + k \le N)$, and y_{t+k}^* is the forecasted or backcasted value otherwise. The authors focus on $Var(\hat{T}_t^{BK} - T_t^{X11})$ under the GE2 definition of the signal by which the irregular term is part of the signal, so that the variance is taken over the distributions of the sampling errors and the forecast and backcast prediction errors. Notice that since $E(y_{t+k}^* - y_{t+k}) = 0$ under the model (unconditional on G),

$$E(\hat{T}^{BK} - T_t^{X11}) = E\left[\sum_{k=-a_T}^{a_T} w_k^T y_{t+k}^* - \sum_{k=-a_T}^{a_T} w_k^T G_{t+k}\right]$$
$$= E\left[\sum_{k=-a_T}^{a_T} w_k^T y_{t+k}^* - \sum_{k=-a_T}^{a_T} w_k^T y_{t+k}\right] = 0, \quad (12)$$

such that the estimators of the trend are likewise unbiased unconditionally. However, when conditioning on the signal, in general $E\left[\left(\hat{T}_{t}^{BK} - T_{t}^{X11}\right) \middle| \mathbf{G}\right] \neq 0$. As is evident from (4), a bias may also exist even unconditionally when forecasting and backcasting less than a_{T} observations, depending on the distribution of the signal.

Our approach differs from Bell and Kramer (1999) in three main aspects.

I- Our definition of the MSE and its estimation is not restricted to the case of full forecasts and backcasts, and it can be applied for any linear estimator of the form $\tilde{H}_t = \sum_{k=-(t-1)}^{N-t} h_{kt} y_{t+k}$. In particular, it applies to the case where the seasonally adjusted and trend components are estimated by use of X-11-ARIMA with only one or two years of forecasts and backcasts, the common case in practice, or even without ARIMA extrapolations. It also applies when estimating the components by signal extraction under appropriate ARIMA models as in X-13ARIMA-SEATS, or by fitting a state-space model to the series as in Sections 3 and 4.

II- We attempt to estimate the conditional MSE given the signal, even though the signal is not observed. We believe that many users of seasonally adjusted and trend estimators would feel most comfortable with the notion that the corresponding target components are fixed values, which conforms to classical sampling theory under which the target parameters are functions of the population values, which are viewed as fixed, nonstochastic quantities. In fact, under definition *GE2* of the signal, the target component values are just linear combinations of the unadjusted population values defining the series, which in most cases are finite population means or totals. On the other hand, as already stated in Remark 8, our bias estimators may also be viewed as estimating the unconditional bias over all possible realizations of the signal under an appropriate model, given the observed series.

III- Our approach is applicable to the case where the signal consists of only the trend and the seasonal effect, and the time series irregular is part of the error (definition GE1 of the signal). We mention also that in its present state, the application of Bell and Kramer's procedure is not straightforward and requires many intermediate steps. See Bell and Kramer (1999) and Scott et al. (2012) for details.

3. Estimation of MSE of Model-Based and Other Estimators of X-11 Components

Consider any other set of component estimators of the form

$$\tilde{T}_t = \sum_{k=-(t-1)}^{N-t} h_{kt}^T y_{t+k}, \quad \tilde{S}_t = \sum_{k=-(t-1)}^{N-t} h_{kt}^S y_{t+k}.$$
(13)

Then, similarly to the X-11-ARIMA estimators in Section 2, we can calculate the conditional bias and MSE with respect to the target X-11 components defined in Definition 1, yielding the same expressions as in (4)–(6) but with the weights $w_{kt}^T(w_{kt}^S)$ replaced by the weights $h_{kt}^T(h_{kt}^S)$. Note that unlike the X-11 estimators, the estimators defined by (13) are potentially biased when conditioning on the signal, even at the center of the series.

In the present article, besides X-11-ARIMA estimators, we also consider estimators obtained by fitting a simple state-space model (see Subsection 4.1). The state-space model estimates of the seasonal component and the trend for a given time t are again linear combinations of all the observed values. We calculated the weights defining the corresponding filters by using the impulse response method (Findley and Martin 2006). According to this method, the weight of an observation at time τ when applying the filter at time t is computed by applying the model fitted to the observed series to a series composed of 1 at time τ and 0 elsewhere, and then observing the filter value for time t, $t = 1, \ldots, N$. Calculation of the weights for all time points of a series of length N therefore requires running the model N times, each time with a vector observation defined by a different column of the identity matrix I_N . As in Subsection 2.4, in this case the bias is estimated by estimating the augmented signal $\mathbf{G}^{aug} = (G_{-a_T+1}, \ldots, G_0, \ldots, G_N, \ldots, G_{N+a_T})$ under an appropriate model. The bias and MSE estimators are obtained similarly to Eqs. (9)–(11).

In a recent article, Tiller (2012) suggested another approach to trend estimation which consists of applying time series model-based signal extraction to estimate and remove the sampling errors from the original series, and then applying the X-11-ARIMA trend filter to the adjusted series. Under definition *GE2* of the signal, the use of this approach reduces to applying the trend filter to the estimated signal under the model. Note that since the estimated signal is a linear filter and the X-11-ARIMA trend filter is linear as well, the trend estimators obtained under this approach are again linear combinations of all the observed values and we may apply our proposed approach to estimate the bias and MSE of the trend estimators obtained this way.

Remark 9. We have not considered Tiller's (2012) approach in the simulation study described below, but Pfeffermann et al. (1998) applied this approach to Labour Force series in Australia and found that the resulting trend estimators were very similar to the trend estimators obtained directly under the model.

4. Simulation Study

In this section we apply the estimators considered in Sections 2 and 3 to simulated series, generated from a state-space model fitted by the Bureau of Labour Statistics (BLS) in the

U.S.A. to the series *Employment to Population Ratio in the District of Columbia*, abbreviated hereafter by EP-DC. The EP series is obtained from the Current Population Survey (the US Labour Force Survey) and it estimates the percentage of employed persons out of the total population aged 15+. This is one of the key economic series in the U.S.A., produced monthly by the BLS for each of the 50 States and DC. The BLS uses a similar model for the production of the major employment and unemployment estimates in all the states of the U.S.A.; see Tiller (1992) for details. In order to assess the performance of the various estimators, we generated a large number of series from the EP-DC model. The model depends on 18 estimated hyperparameters but for the present experiment we consider the hyperparameter estimates as true known parameters.

4.1. Model Fitted to EP-DC Series

The EP-DC series is very erratic: The residual component (calculated by X-11-ARIMA) explains 55% of the month to month changes and 32% of the yearly changes. A large portion of the residual component is sampling errors. The series is plotted in Figure 2, along with the trend estimated under the EP-DC model defined below, and the trend estimated by X-11-ARIMA with twelve months forecasts when fitting the familiar airline model ARIMA(0,1,1)(0,1,1), selected by the program. The two trends behave similarly, but the trend estimated under the EP-DC model is smoother, a phenomenon observed in many other series. The X-11-ARIMA trend is below the EP-DC trend for most of the time points, but the average values of the two trends are very close: Av.(trend EP-DC) = 63.11, Av.(trend X11 ARIMA) = 63.01, Av.(original series) = 63.00.



Fig. 2. Employment to Population Ratio in DC (in percentages), Jan2001-Dec2010. Original series and trends estimated by X-11-ARIMA with 12 forecasts, and by the EP-DC model

Let y_t define, the direct sample estimate for time t and Y_t the corresponding true population ratio such that $\varepsilon_t = y_t - Y_t$ is the sampling error. The state-space model fitted to the series y_t combines a model for Y_t with a model for ε_t . The model postulated for Y_t is the basic structural model (BSM, Harvey 1989)

$$Y_{t} = T_{t} + S_{t} + I_{t}, \quad I_{t} \sim N(0, \sigma_{I}^{2}); \quad T_{t} = T_{t-1} + R_{t-1}, \quad R_{t} = R_{t-1} + \eta_{Rt},$$

$$\eta_{Rt} \sim N(0, \sigma_{R}^{2})$$

$$S_{j,t} = \cos \omega_{j} S_{j,t-1} + \sin \omega_{j} S_{j,t-1}^{*} + \eta_{j,t}, \quad \eta_{j,t} \sim N(0, \sigma_{S}^{2})$$

$$S_{j,t}^{*} = -\sin \omega_{j} S_{j,t-1} + \cos \omega_{j} S_{j,t-1}^{*} + \eta_{j,t}^{*}, \quad \eta_{j,t}^{*} \sim N(0, \sigma_{S}^{2}),$$

$$S_{t} = \sum_{j=1}^{6} S_{j,t}; \quad \omega_{j} = 2\pi j/12, \quad j = 1...6.$$
(14)

The error terms I_t , η_{Rt} , $\eta_{j,t}$, $\eta_{j,t}^*$ are mutually independent normal disturbances. In this model, T_t is the trend level, R_t is the slope and S_t is the seasonal effect. The trend model approximates a local linear trend, whereas the model for the seasonal effects uses the traditional decomposition of the seasonal component into eleven cyclical components corresponding to the six seasonal frequencies. The innovations $\eta_{j,t}$, $\eta_{j,t}^*$ allow the seasonal effects to evolve over time.

The model fitted for the sampling errors is AR(15); see Pfeffermann and Tiller (2005) for the considerations leading to the choice of this model.

The separate models holding for the population ratios and the sampling errors are cast into a single state-space model. In what follows we refer to the combined model holding for the observed series $y_t = Y_t + \varepsilon_t$ as the extended BSM (EBSM). Note that the state vector consists of the trend, slope, seasonal effects and sampling errors. The variances and covariances of the sampling errors are estimated from the survey micro-data using a replication approach with a large number of replications. The *AR*(15) model coefficients are then estimated by solving the corresponding Yule-Walker equations and they are set to their estimated values when estimating the population model variances by maximum likelihood. The variances and AR coefficients used for the present simulation experiment are the same as in Pfeffermann and Tiller (2005). See that paper for further details on the way we generated series under the model and for the values of the model variances and AR coefficients.

4.2. Simulation Plan

We generated three sets of 1,000 monthly series of length 300; { $y_{t,b}$, t = 1, ..., 300, b = 1, ..., 1,000}. The first set was generated by simulating for every month t a trend value, T_t , a seasonal effect, S_t , and an irregular term, I_t , from the model (14), and a sampling error, ε_t , from the AR(15) model, and then adding the separate components; $y_{t,b} = T_{t,b} + S_{t,b} + I_{t,b} + \varepsilon_{t,b}$, b = 1, ..., 1,000. The second set was obtained from the first set by halving the sampling errors, that is, $y_{t,b} = T_{t,b} + S_{t,b} + I_{t,b} + \varepsilon_{t,b}/2$. The third set was obtained from the first set by doubling the sampling errors, i.e., $y_{t,b} = T_{t,b} + S_{t,b} + I_{t,b} + 2\varepsilon_{t,b}$. Considering the three data sets allows an assessment of the effect of the magnitude of the sampling errors on the performance of the estimators. For the present study we employ the definition *GE*2 of the signal by which $G_{t,b} = T_{t,b} + S_{t,b} + I_{t,b}$. We computed the default X-11 estimator of the trend and seasonal component for each simulated signal of length 300, so as to obtain the target X-11 components defined by (3) for the central 180 months. (For the default X-11 estimator $a_S = 84$, $a_T = 90$, but augmenting the series with only 60 forecasts and backcasts yields almost identical target components.) We defined the target seasonally adjusted component as the difference between the original series without sampling error and the target seasonal component, that is,

$$T_t^{X11} = \sum_{k=-a_T}^{a_T} w_k^T G_{t+k}; \quad S_t^{X11} = \sum_{k=-a_S}^{a_S} w_k^S G_{t+k}; \quad SA_t^{X11} = (y_t - S_t^{X11} - \varepsilon_t).$$
(15)

Finally, we removed the first and last 60 monthly observations from the simulated series and applied X-11-ARIMA with twelve and 60 forecasts to the reduced series of length 180, using the default X-11 filters but setting the ARIMA model as the airline model ARIMA(0,1,1),(0,1,1) (Remark 10 below). Thus, the X-11-ARIMA estimators are

$$\hat{T}_{t,b}^{X11} = \sum_{k=-(t-1)}^{N-t} w_{kt}^T y_{t+k,b}, \quad S\hat{A}_{t,b}^{X11} = y_{t,b} - \sum_{k=-(t-1)}^{N-t} w_{kt}^S y_{t+k,b}, \tag{16}$$

where the weights $\{w_{kt}^T\}$, $\{w_{kt}^S\}$ are defined by the ARIMA model, the program default options and the number of forecasts (12 or 60). We also computed the EBSM estimators, obtained by replacing the weights $\{w_{kt}^T\}$, $\{w_{kt}^S\}$ in (16) by the weights $\{h_{kt}^T\}$, $\{h_{kt}^S\}$ (Section 3).

Remark 10. In the simulation study we did not select new ARIMA models or re-estimated the model coefficients for every simulated series. We used the airline model for all the simulated series and estimated the model parameters once for each set of series, based on a randomly selected series from the set. Selecting a model and re-estimating the model coefficients for each simulated series would require new computation of the filter weights for every series and every model, which is not feasible in a simulation study with 3,000 series. See Section 3 for the method used for estimating the filter weights. Notwithstanding, X-13ARIMA-SEATS selected the airline model as the preferred model for most of the series in all three sets. Moreover, for series of length 180 (quite typical for monthly economic series), the estimation of the model coefficients is generally very stable and is not expected to affect the results. (The sampling error variances and hence the AR(15) model coefficients are taken as known. The model variances are estimated by MLE, which are known to be consistent.) We also reiterate the statement made in Remark 3 above that our purpose in this article is to propose a method of estimating the conditional bias and RMSE of linear estimators of the proposed target components, and not to search for the most appropriate model and estimators. In practice, one would let the program select the model and estimate the unknown coefficients, and then compute the required filter weights for the particular choice of model and estimates.

4.3. Computations

Because of space limitations, in subsequent subsections we restrict ourselves to the estimation of the target trend. Estimation of the MSE of seasonally adjusted estimators is considered in Section 5. We computed the following statistics:

4.3.1. Conditional Variance of X-11-ARIMA and EBSM Estimators

$$V_t^{T,X11} = \sum_k \sum_l w_{kt}^T w_{lt}^T Cov(\varepsilon_{t+k}, \varepsilon_{t+l}), \quad V_t^{T,EBSM} = \sum_k \sum_l h_{kt}^T h_{lt}^T Cov(\varepsilon_{t+k}, \varepsilon_{t+l}).$$
(17)

The variances and covariances of the sampling errors are taken as known.

4.3.2. Conditional Bias and Root MSE of X-11-ARIMA and EBSM Estimators

The conditional bias and root MSE (RMSE) for a given signal when estimating the target trend in Eq. 3 are:

$$B_{t,b}^{T,X11} = \sum_{k=-(t-1)}^{N-t} w_{kt}^{T} G_{t+k,b} - \sum_{k=-a_{T}}^{a_{T}} w_{k}^{T} G_{t+k,b},$$

$$RMSE(\hat{T}_{t,b}^{X11}) = [V_{t}^{T,X11} + (B_{t,b}^{T,X11})^{2}]^{1/2}.$$
(18)

The bias and RMSE of the EBSM estimators are obtained in similar manner.

4.3.3. Estimation of Squared Bias and MSE

Denote by $\hat{B}_{t,b}^{T,X11}$ the estimate of the bias, obtained by predicting the unknown signal using the models extracted for the trend and seasonal effects by X-13ARIMA-SEATS (Eq. 9), or by predicting the signal using the X-11-ARIMA forecasts and backcasts (Eq. 10).

The RMSE is estimated as,

$$M\hat{S}E_{t,b}^{T,X11} = V_t^{T,X11} + (\hat{B}_{t,b}^{T,X11})^2 - V(\hat{B}_{t,b}^{T,X11}),$$

$$M\hat{S}E_{t,b}^{T,EBSM} = V_t^{T,EBSM} + (\hat{B}_{t,b}^{T,EBSM})^2 - V(\hat{B}_{t,b}^{T,EBSM}).$$
(19)

where $V(\hat{B}_{t,b}^{T,X11})$, $V(\hat{B}_{t,b}^{T,EBSM})$ are the variances of the bias estimates, computed similarly to the variances of the estimators in Eq. 17. As explained in Subsection 2.4, subtracting the variance of the bias estimator is necessary for unbiased MSE estimation.

4.3.4. Error of Estimators of Squared Bias and RMSE

$$EBS_{t,b}^{B,X11} = (B_{t,b}^{T,X11})^2 - [(\hat{B}_{t,b}^{T,X11})^2 - V(\hat{B}_{t,b}^{T,X11})],$$

$$EM_{t,b}^{RMSE,X11} = RMSE(\hat{T}_{t,b}^{X11}) - \sqrt{M\hat{S}E_{t,b}^{T,X11}}.$$
(20)

Similar expressions apply when using the EBSM estimators.

4.4. Results

The results are summarized in Tables 1-3 and Figures 3-8. The tables show average results obtained for the three sets of series, for each of the last six months of the reduced series (time

set of 1,000 series, last six months of series								
	Month	Jul	Aug	Sep	Oct	Nov	Dec	
X-11 ARIMA	ABS	0.020	0.021	0.021	0.035	0.113	0.332	
60 forecasts	AEBS(9)	0.002	0.002	0.004	0.017	0.084	0.273	
	SDEBS(9)	(0.042)	(0.043)	(0.044)	(0.062)	(0.177)	(0.508)	
	AEBS(10)	-0.009	-0.012	0.009	0.000	0.012	0.022	
	SDEBS(10)	(0.326)	(0.333)	(0.333)	(0.357)	(0.508)	(0.831)	
	ARMSE	1.268	1.267	1.265	1.285	1.347	1.457	
	AEM(9)	0.001	0.001	0.001	0.006	0.030	0.089	
	SDEM(9)	(0.016)	(0.017)	(0.017)	(0.024)	(0.062)	(0.154)	
	AEM(10)	-0.026	-0.026	-0.026	-0.023	-0.008	0.019	
	SDEM(10)	(0.046)	(0.048)	(0.049)	(0.055)	(0.090)	(0.187)	
X-11 ARIMA	ABS	0.024	0.025	0.026	0.041	0.120	0.343	
12 forecasts	AEBS(9)	0.018	0.020	0.021	0.035	0.108	0.036	
	SDEBS(9)	(0.035)	(0.034)	(0.036)	(0.060)	(0.182)	(0.514)	
	AEBS(10)	-0.003	-0.001	0.004	0.012	0.019	0.022	
	SDEBS(10)	(0.254)	(0.247)	(0.241)	(0.276)	(0.445)	(0.774)	
	ARMSE	1.268	1.274	1.274	1.291	1.352	1.463	
	AEM(9)	0.007	0.008	0.008	0.013	0.039	0.100	
	SDEM(9)	(0.014)	(0.013)	(0.014)	(0.023)	(0.063)	(0.154)	
	AEM(10)	-0.008	-0.006	-0.004	-0.001	0.014	0.041	
	SDEM(10)	(0.027)	(0.025)	(0.023)	(0.032)	(0.075)	(0.176)	
EBSM	ABS	0.240	0.221	0.209	0.221	0.287	0.441	
	AEBS(9)	0.010	0.022	0.018	0.001	-0.006	0.044	
	SDEBS(9)	(0.396)	(0.355)	(0.328)	(0.358)	(0.525)	(0.845)	
	AEBS(10)	-0.004	0.038	0.078	0.101	0.097	0.061	
	SDEBS(10)	(0.570)	(0.510)	(0.493)	(0.541)	(0.660)	(0.852)	
	ARMSE	1.124	1.149	1.185	1.230	1.286	1.367	
	AEM(9)	0.005	0.009	0.007	0.001	-0.001	0.016	
	SDEM(9)	(0.155)	(0.137)	(0.125)	(0.132)	(0.179)	(0.265)	
	AEM(10)	-0.067	-0.047	-0.028	-0.024	-0.036	-0.041	
	SDEM(10)	(0.207)	(0.182)	(0.172)	(0.189)	(0.235)	(0.312)	

Table 1. Means of true squared bias and RMSE, simulation means of error when estimating the squared bias and RMSE, and SD of simulation means of error as obtained by application of X-11-ARIMA and by EBSM. First set of 1,000 series, last six months of series

points 175–180). As stated before, we restrict to estimation of the target trend and we show the results of estimating the squared bias and the RMSE (Eq. 20). We use the following abbreviations: ABS is the simulation mean (over 1,000 series) of true $bias^2$ (average of square of Eq. 18), ARMSE is the simulation mean of the true RMSE (Eq. 18); Note that the variance of any given estimator is fixed for all simulated series in a given set, but the signal, and hence the bias, changes from one simulated series to another. AEBS(9) is the simulation mean of the error in estimating $bias^2$ (Eq. 20) when the signal is estimated as in (9), AEBS(10) is the simulation mean of the error in estimating $bias^2$ when the signal is estimated by (10); SDEB(9) and SDEBS(10) are the standard deviations (SD) of the simulation means AEBS(9) and AEBS(10) respectively. AEM(9) is the simulation mean of the error of the RMSE estimates (Eq. 20) when the signal is estimated as in (9), AEM(10) is the simulation mean of the error of the RMSE estimates (Eq. 20) when the signal is estimated as in (9), AEM(10) is the simulation mean of the error of the RMSE (Eq. 20) when the signal is estimated as in (9), AEM(10) is the simulation mean of the error of the RMSE estimates (Eq. 20) when the signal is estimated as in (9), AEM(10) is the simulation mean of the error of the RMSE estimates (Eq. 20) when the signal is estimated by (10); SDEM(9) and SDEM(10) are the signal is estimated as in (9), AEM(10) is the simulation mean of the error of the RMSE estimates when the signal is estimated by (10); SDEM(9) and SDEM(10) are the standard deviations of the means AEM(9) and AEM(10) respectively.

Figures 3-8 show the means of the true and estimated squared bias and RMSE, as obtained for the last four years of the series for each of the three sets of series by

	Month	Jul	Aug	Sep	Oct	Nov	Dec
X-11 ARIMA	ABS	0.020	0.021	0.021	0.035	0.113	0.332
60 forecasts	AEBS(9)	0.010	0.011	0.012	0.026	0.098	0.298
	SDEBS(9)	(0.031)	(0.032)	(0.034)	(0.054)	(0.172)	(0.501)
	AEBS(10)	-0.007	-0.009	-0.006	0.002	0.011	0.017
	SDEBS(10)	(0.238)	(0.248)	(0.246)	(0.268)	(0.427)	(0.739)
	ARMSE	0.644	0.644	0.644	0.661	0.725	0.851
	AEM(9)	0.008	0.008	0.009	0.019	0.064	0.161
	SDEM(9)	(0.023)	(0.024)	(0.025)	(0.038)	(0.101)	(0.227)
	AEM(10)	-0.014	-0.015	-0.014	-0.005	0.033	0.104
	SDEM(10)	(0.042)	(0.046)	(0.048)	(0.057)	(0.113)	(0.241)
X-11 ARIMA	ABS	0.024	0.025	0.026	0.041	0.120	0.343
12 forecasts	AEBS(9)	0.022	0.023	0.024	0.039	0.115	0.323
	SDEBS(9)	(0.034)	(0.033)	(0.035)	(0.059)	(0.181)	(0.512)
	AEBS(10)	- 0.003	- 0.001	0.003	0.011	0.016	0.017
	SDEBS(10)	(0.185)	(0.185)	(0.180)	(0.216)	(0.385)	(0.693)
	ARMSE	0.647	0.651	0.651	0.667	0.731	0.859
	AEM(9)	0.017	0.017	0.018	0.029	0.075	0.175
	SDEM(9)	(0.025)	(0.025)	(0.026)	(0.041)	(0.105)	(0.229)
	AEM(10)	0.007	0.009	0.010	0.019	0.056	0.128
	SDEM(10)	(0.030)	(0.029)	(0.029)	(0.044)	(0.109)	(0.239)
EBSM	ABS	0.240	0.221	0.209	0.221	0.287	0.441
	AEBS(9)	0.160	0.140	0.120	0.120	0.167	0.291
	SDEBS(9)	(0.279)	(0.264)	(0.260)	(0.260)	(0.375)	(0.648)
	AEBS(10)	-0.002	0.038	0.076	0.096	0.090	0.052
	SDEBS(10)	(0.340)	(0.330)	(0.355)	(0.403)	(0.494)	(0.665)
	ARMSE	0.690	0.694	0.717	0.749	0.794	0.871
	AEM(9)	0.109	0.092	0.078	0.076	0.099	0.152
	SDEM(9)	(0.163)	(0.153)	(0.147)	(0.147)	(0.195)	(0.297
	AEM(10)	0.016	0.021	0.0290	0.033	0.035	0.057
	SDEM(10)	(0.172)	(0.166)	(0.179)	(0.203)	(0.248)	(0.332)

Table 2. Means of true squared bias and RMSE, simulation means of error when estimating the squared bias and RMSE, and SD of simulation means of error as obtained by application of X-11-ARIMA and by EBSM. Second set of 1,000 series, last six months of series

application of X-11-ARIMA and the EBSM. The X-11-ARIMA estimators refer to the case of twelve months forecasts (generally similar results to the case of 60 months forecasts). The signal in all the figures is estimated by use of X-11-ARIMA (Eq. 10), which produces somewhat less biased estimators of the RMSE than the use of Eq. 9, although occasionally with larger SD (see summary below).

The main conclusions from the simulation study can be summarized as follows:

- 1. The simulation mean of the errors over all realizations of the signal and the sampling errors of the 1,000 series in each set, when estimating the true bias of the estimators are all very close to zero for all the estimators and all three data sets (not shown in the tables). Thus, our proposed estimators of the bias of the estimators of the target trend are unbiased *unconditionally*, although occasionally with large standard errors.
- 2. The true ARMSEs of the estimators increase by a magnitude of around two when increasing the SD of the sampling errors by two. Thus, the ARMSEs in Table 1 are

Month

ABS

AEBS(9) SDEBS(9) AEBS(10) SDEBS(10) DMOD

X-11 ARIMA

60 forecasts

bias and RM neans of erro s of series	SE, simulatio r as obtained	n means of ei by applicatio	rror when esti n of X-11-ARI	mating the sq MA and by El	uared bias 3SM. Third
Jul	Aug	Sep	Oct	Nov	Dec
0.020	0.021	0.021	0.035	0.113	0.332
-0.042	-0.039	-0.034	-0.025	0.019	0.157
(0.121)	(0.118)	(0.101)	(0.126)	(0.248)	(0.595)
-0.011	-0.016	-0.015	-0.005	0.012	0.030
(0.552)	(0.549)	(0.552)	(0.586)	(0.747)	(1.122)
2.530	2.530	2.520	2.550	2.651	2.802
-0.008	-0.008	-0.007	-0.005	0.004	0.027
(0.023)	(0.022)	(0.020)	(0.024)	(0.046)	(0.102)
0.051	0.040	0.050	0.052	0.050	0.002

Table 3. Means of true squared and RMSE, and SD of simulation r set of 1,000 series, last six month

	ARMSE	2.550	2.330	2.320	2.550	2.031	2.002
	AEM(9)	-0.008	-0.008	-0.007	-0.005	0.004	0.027
	SDEM(9)	(0.023)	(0.022)	(0.020)	(0.024)	(0.046)	(0.102)
	AEM(10)	-0.051	-0.049	-0.050	-0.052	-0.058	-0.083
	SDEM(10)	(0.074)	(0.075)	(0.075)	(0.083)	(0.115)	(0.206)
X-11 ARIMA	ABS	0.024	0.025	0.026	0.041	0.120	0.343
12 forecasts	AEBS(9)	-0.009	-0.004	-0.002	0.007	0.060	0.209
	SDEBS(9)	(0.064)	(0.058)	(0.058)	(0.086)	(0.209)	(0.555)
	AEBS(10)	-0.003	-0.001	0.005	0.014	0.023	0.031
	SDEBS(10)	(0.429)	(0.440)	(0.399)	(0.441)	(0.633)	(1.028)
	ARMSE	2.525	2.536	2.532	2.557	2.654	2.807
	AEM(9)	-0.002	-0.001	-0.000	0.001	0.011	0.036
	SDEM(9)	(0.012)	(0.011)	(0.011)	(0.017)	(0.039)	(0.095)
	AEM(10)	-0.027	-0.023	-0.021	-0.023	-0.023	-0.054
	SDEM(10)	(0.044)	(0.038)	(0.036)	(0.044)	(0.079)	(0.173)
EBSM	ABS	0.240	0.221	0.209	0.221	0.287	0.441
	AEBS(9)	-0.173	-0.141	-0.082	-0.043	-0.081	-0.173
	SDEBS(9)	(1.452)	(1.200)	(0.938)	(0.976)	(1.473)	(2.268)
	AEBS(10)	0.006	0.049	0.089	0.107	0.094	0.044
	SDEBS(10)	(1.186)	(0.935)	(0.847)	(0.879)	(1.044)	(1.300)
	ARMSE	2.090	2.150	2.221	2.301	2.391	2.499
	AEM(9)	-0.023	-0.021	-0.011	-0.002	-0.004	-0.010
	SDEM(9)	(0.291)	(0.238)	(0.190)	(0.193)	(0.270)	(0.386)
	AEM(10)	-0.205	-0.156	-0.116	-0.107	-0.137	-0.178
	SDEM(10)	(0.321)	(0.267)	(0.223)	(0.224)	(0.277)	(0.363)

around twice the ARMSEs in Table 2, and the ARMSEs in Table 3 are around twice the ARMSEs in Table 1. The increase in the ARMSEs is somewhat lower for the EBSM estimators. This outcome is explained by the fact that under the present simulation setup, the major component of the RMSE is the variance of the estimator, which of course depends on the variances and covariances of the sampling errors. (The autocorrelations of the sampling errors are the same for all the three sets.)

3. Interestingly enough, the SD of the mean error when estimating the true RMSE does not show a similarly stable pattern. For example, the SDEM(9) values in Table 3 with the largest variance of the sampling errors are, in the case of X-11-ARIMA with 60 forecasts, smaller for the last two months than the corresponding values in the other two tables. This seemingly odd outcome is explained by the fact that the SD of the true RMSE (not shown) actually decreases as the variance of the sampling errors increases. The latter property follows from the fact that for a given estimator and data set, the variance of the estimator is constant and under general conditions $SD{[bias^2 + var(est.)]^{1/2}}$ decreases as var(est.) increases (can be shown by second-order linearization). Note that unlike the variance, which is fixed in a given



Fig. 3. Means of true and estimated squared bias by application of X-11-ARIMA with twelve months forecasts and EBSM. First set of 1,000 series, last 48 months of data

set, the true $bias^2$ changes from one simulation to another, depending on the random realization of the signal.

The aforementioned phenomenon with the SD of the mean error when estimating the true RMSE does not repeat itself when estimating the true squared bias. Thus, SDEBS(9) and SDEBS(10) are smaller for all estimators and all the months in Table 2 than in Table 1, and smaller in Table 1 than in Table 3.

- 4. The estimators of the true squared bias when estimating the signal by forecasting the series using X-11-ARIMA (Eq. 10) are generally less biased unconditionally (over all realizations of the signal in a given set) than when estimating the signal by the model identified by X13ARIMA-SEATS (Eq. 9), particularly in the last two months (November, December). This outcome may look odd but note that the models used to generate the series for our simulation study (Eq. 14 for the population values and AR(15) for the sampling errors) do not combine to the airline model fitted to the data, so that the model extracted for the trend levels by X13ARIMA-SEATS is not the correct model. On the other hand, the SD of the mean errors when estimating the signal by application of Eq. 9 than when estimating the signal by application of Eq. 9 than when estimating the signal by application of Eq. 9 than when estimating the signal by application is in Table 3 when estimating the trend using the EBSM.
- 5. The conclusions referring to the estimation of the true squared bias generally also apply to the estimation of the true RMSE, particularly with regard to the SD of the mean of the estimation errors. (Compare the rows SDEM(9) and SDEM(10).) In general, we find that our proposed estimators of the RMSE when using the X-11-ARIMA method are



Fig. 4. Means of true and estimated RMSE by application of X-11-ARIMA with twelve months forecasts and EBSM. First set of 1,000 series, last 48 months of data

unbiased in our simulation study when averaging over all possible realizations of the signal, except perhaps for the last two time points, although even there, the bias is never significant using the ordinary *t*-statistic (the ratio of AEM to SDEM is always smaller than 1).

6. Finally, by comparing the results obtained for the three estimators we notice that the ARMSEs are very similar when using the ARIMA estimators with 60 forecasts or with only twelve forecasts. What we find very interesting is that the EBSM produces estimators with lower ARMSEs (lower variances), except in the case of the small sampling errors. As noted before, we used the EBSM for generating the simulated series, but this only partly explains this outcome because the target trend defined by Eq. 3 is not the trend generated under the model and we predicted the signal by use of the airline ARIMA model fitted to the data and not under the EBSM.

Figures 3–8 show the means of the true and estimated squared bias and RMSE as obtained for the last four years of the series for each of the three sets of series, by application of X-11-ARIMA with twelve forecasts (using Eq. 10 for estimating the signal) and the EBSM. The main conclusion from these figures is that the use of X11-ARIMA yields unbiased estimators of the squared bias and the RMSE, except when estimating the RMSE for the last two months in the second set with the small sampling errors. The EBSM estimators seem to be biased, especially in the case of the third set with the large sampling errors, but as can be seen in the tables, the biases are highly insignificant. As already mentioned, the true RMSEs of the EBSM estimators are lower than the true RMSEs of the X11-ARIMA estimators, except for the second set of data with the small sampling errors.

Fig. 5. Means of true and estimated squared bias by application of X-11-ARIMA with twelve months forecasts and EBSM. Second set of 1,000 series, last 48 months of data

Remark 11. We emphasize again that although our proposed estimators condition on a given signal, the results in the tables and figures are unconditional by averaging over the 1,000 realizations of the true signal and the estimators.

5. Application to Current Employment Statistics Series

5.1. Series Considered

In this section we study the performance of the proposed method when applied to real series. We consider four leading monthly employment series, each spreading over 17 years, from February 1990 to January 2007. The series are produced by the BLS based on the Current Employment Statistics (CES) survey, which covers more than 300,000 establishments. The target of interest is the monthly change in employment. The variance and autocovariances of the sampling errors of the unadjusted estimators are estimated each month using the balanced repeated replication (BRR) method, with a modification proposed by Robert Fay, using a factor of 0.5 to reflect the sampling design (see http://www.bls.gov/web/empsit/cestn2.htm#4). The CES survey has the advantage of having time-lagged true population figures from the Unemployment Insurance Program (UIP). Quarterly business tax forms collected by the UIP include monthly employment data which are assembled first at the state level and then at the national level. The true population value for March of each year becomes available by the following January and then all the estimates from March of the previous year up to the current January are

Fig. 6. Means of true and estimated RMSE by application of X-11-ARIMA with twelve months forecasts and EBSM. Second set of 1,000 series, last 48 months of data

benchmarked for the difference between the population and estimated levels in March of last year. No benchmarking is carried out in the months of February to December of a current year. The employment estimate for a current month t is computed as a "link-relative" estimator,

$$\hat{E}_t = E_0 \times r_1 \times r_2 \times \ldots \times r_t, \tag{21}$$

where E_0 is the latest available population value and subsequent subscripts denote months after the benchmark month. The links r_j are ratios between employment estimates in adjacent months,

$$r_{j} = \frac{\sum_{i \in M_{j}} d_{ij} x_{ij}}{\sum_{i \in M_{j}} d_{i,j-1} x_{i,j-1}},$$
(22)

where x_{ij} represents the number of employees in establishment *i* at month *j*, d_{ij} is the survey weight and M_j represents the set of establishments for which the number of employees is reported for both months *j* and *j* – 1.

In this study we focus on the estimation of the MSEs of seasonally adjusted estimators (SAE) produced by X-11-ARIMA with 24 months of forecasts, which is the common routine at the BLS for these series. We focus on SAE since it allows us to compare our proposed MSE estimators with estimators produced by application of Bell and Kramer's (1999) approach, reported in Scott et al. (2012). Traditionally, the CES employment series

Fig. 7. Means of true and estimated squared bias by application of X-11-ARIMA with twelve months forecasts and EBSM. Third set of 1,000 series, last 48 months of data

have been seasonally adjusted multiplicatively. This suggests considering monthly changes in the log scale,

$$y_t = \log(\hat{E}_t) - \log(\hat{E}_{t-1}) = \log\left(\frac{\hat{E}_t}{\hat{E}_{t-1}}\right).$$
 (23)

Under the multiplicative decomposition, $y_t = (\log (E_t) - \log (E_{t-1})) + \log (\varepsilon_t / \varepsilon_{t-1})$, decomposing the estimated monthly change as the sum of the population value, $Y_t = \log (E_t) - \log (E_{t-1})$ and a sampling error component, $\tilde{\varepsilon}_t = \log (\varepsilon_t / \varepsilon_{t-1})$.

Remark 12. For the present illustrative study, the input series are the ratios of the benchmarked estimators. Previous studies show that the ratios of the benchmarked estimators are very close to the ratios of the unbenchmarked estimators, and in what follows, we refer to the observed series as the ratios r_t . BRR estimates for the variances and covariances of the sampling errors $\tilde{\varepsilon}_t$ of the log ratios, $\log(r_t)$, have been produced and are used for the computations of the various estimators. As stated above, the benchmarking changes the current estimates, and hence the variances and covariances, very little.

Following the methodology of the previous sections we fit ARIMA models to $\log(\hat{E}_t)$ with one regular difference, such that the observed input series has the general form, $y_t = (1 - B) \log(\hat{E}_t) = \log(r_t)$. Furthermore, assuming that the ratios r_t fluctuate around 1 and using a Taylor expansion, $\log(r_t) \approx r_t - 1 = (\hat{E}_t - \hat{E}_{t-1})/\hat{E}_{t-1}$. Thus, the seasonally adjusted values of the series $y_t = \log(r_t)$ can be interpreted as estimating the seasonally adjusted values of the percentage change in employment, which is the focus of estimation.

Fig. 8. Means of true and estimated RMSE by application of X-11-ARIMA with twelve months forecasts and EBSM. Third set of 1,000 series, last 48 months of data

5.2. Results

We present the results obtained for the last five years of data when applying our proposed method of RMSE estimation and the method proposed by Bell and Kramer (1999, hereafter B-K), to the following four series: "Total Employment in Education and Health Services"; "Total Employment in Manufacturing, Durable Goods"; "Total Employment in Manufacturing, Nondurable Goods" and "Total Employment in Retail Trade". Using standard ARIMA model fitting and diagnostic techniques, we fit the model (1,0,1) (0,1,1) to the first three series and the model (1,0,0) (0,1,1) to the last series. (The input series is in all cases $y_t = \log(I_t^-)$). As mentioned above, we used two years of monthly forecasts when computing the X-11-ARIMA estimator $S\hat{A}_t$ of the seasonally adjusted value SA_t for time *t*. The MSE estimator under the proposed method is,

$$M\hat{S}E(S\hat{A}_t|\mathbf{G}) = V\hat{a}r(S\hat{A}_t|\mathbf{G}) + B\hat{a}s^2(S\hat{A}_t|\mathbf{G}) - V\hat{a}r[B\hat{a}s(S\hat{A}_t|\mathbf{G})|\mathbf{G}], \quad (24)$$

with the signal estimated by X-11 ARIMA (Eq. 10, the same as in the Figures 3–8). See Eq. 15 for the definition of the SA estimator, and Subsection 2.5 for discussion of the difference between the proposed MSE estimator and the B-K method. We also show the RMSE estimators obtained under the proposed method when the irregular term is part of the error (definition *GE*1 of the signal in Subsection 2.1). As noted before, no B-K estimators are available for this definition of the signal. In addition, we show the conditional standard deviations (SD) of the estimators $S\hat{A}_t$ given the signal *GE*2 (SQRT of Eq. 5 with respect to the SAE), and the SD of the original, unadjusted estimators. The last two SDs only account for the variance of the sampling errors. All the values in the figures are multiplied by 10,000.

Fig. 9. Results for Total Employment in Education and Health Services

Fig. 10. Results for Total Employment in Manufacturing, Durable Goods

Fig. 11. Results for Total Employment in Manufacturing, Nondurable Goods

It is hard to assess the performance of the various estimators because the true MSEs are unknown when analyzing real series, but the following points are worth mentioning.

- 1. The SD of the SAE given the signal *GE*2, by which the error consists only of the sampling error without the irregular term, is always smaller than the SD of the unadjusted estimator. This result is explained by the fact that the SAEs are weighted averages of the unadjusted estimators with weights that sum to 1.
- 2. The RMSE estimates given the signal *GE*1 are always higher than the RMSE estimates given the signal *GE*2, which is obvious since under definition *GE*1 the signal consists only of the trend and seasonal effect and the irregular term is part of the error.
- 3. The RMSE estimates given the signal *GE*2 are literally the same as the B-K SD estimates in the center of the series (until around time point 168). However, except for Figure 9 (Total Employment in Education and Health Services), for the last three years of data the B-K SD estimates are higher than the conditional RMSE estimates given the signal *GE*2. As discussed in Subsection 2.5, the two estimators differ in the definition of the estimators of the SAE (B-K assume that the X-11 ARIMA SAE use seven years of forecasts whereas in our present application the SAE use only two years of forecasts), and in the definition of the target MSE (we condition on the actual signal, whereas the B-K variance is over all possible realizations of the signal under the ARIMA model fitted to the series, thus accounting for the forecast and backast prediction errors).
- 4. Except for Figure 9, The RMSE estimates given the signal *GE*2 are very close to the SD of the SAE given the signal *GE*2, and the bias corrections contribute only marginally. On the other hand, in Figure 9, the SDs are much smaller than the

Fig. 12. Results for Total Employment in Retail Trade

RMSEs in the last two years where the SAE use asymmetric weights, indicating a significant contribution of the bias corrections.

5. The RMSE estimates given the signal GE1 are slightly higher than the SD of the unadjusted series in two of the series, but are appreciably lower in the two other series.

6. Summary

In this article we propose a new method for the estimation of the MSE of X-11-ARIMA estimators or other linear estimators of the underlying components of a time series. Our approach has some important advantages over other approaches proposed in the literature. First, we follow Bell and Kramer (1999) by defining the target component values as the corresponding X-11 estimators that would be obtained if the series were free of sampling errors and long enough to permit the use of the symmetric filters embedded in the program. In other words, the target components are real entities defined as linear combinations of finite population means or totals over time, in close correspondence to the target values in classical finite population sampling. In particular, under definition GE2 of the signal, the target component values are just linear combinations of the unadjusted finite population values. Interestingly, while the programme X-11 for seasonal adjustment and its previous and subsequent versions have been in wide use for many decades, the target estimated values were never defined in a precise form. This is rather unusual in statistics, where an estimator is defined but not what is estimated. This problem does not exist when using model-dependent methods where the targets are defined by the model, such as in the BSM, the Tramo and Seats program (Gómez and Maravall 1996) and in one of the modules of X-13ARIMA-SEATS, but purely model dependent estimators are not in common use, at least not in national statistical offices.

A second notable advantage of our procedure is that for definition GE2 of the signal, the procedure is basically automatic and does not require new programs or external intervention beyond what is required for the production of the component estimators themselves. Thus, the X-13ARIMA-SEATS programme produces the models for the trend and seasonal components and hence for the signal. These models are then used to estimate the signal within the observation period and to predict the signal outside the observation period. The weights required to define the X-11-ARIMA estimators and the bias estimators (Eq. 9) can be obtained by repeated runs of X-11-ARIMA, as described in Section 3 and in Burck and Sverchkov (2001). In the case of definition GE1 of the signal, the application of our procedure additionally requires the estimation of the variances and covariances of the combined errors or at least the variance and covariances of the irregular terms (Subsection 2.3), for which an additional program has to be used.

A third important advantage of the procedure is its flexibility in terms of the target values and the estimators used. It is applicable to the case where the signal consists of only the trend and the seasonal effect and the time series irregular component is part of the error (definition GE1 of the signal and error), and to the case where the irregular component is part of the signal, as under the B-K approach. It is up to the user to decide which definition of the signal is more appropriate. In addition, the procedure is applicable to any linear estimator with known coefficients.

Finally, and most importantly, we have illustrated the good performance of the procedure in estimating the true unknown MSEs, as defined in this article.

Taking into account the clear interpretation of the target values and the estimated MSE and the other advantages listed above, we hope that our proposed procedure will be experimented with by other users and we shall be happy to receive questions arising from these experiments.

7. References

- Bell, W.R. and M. Kramer. 1999. "Toward Variances for X-11 Seasonal Adjustments." Survey Methodology 25: 13–29.
- Burck, L. and M. Sverchkov. 2001. "A General Method for Estimating the Variances of X-11-ARIMA Estimators." *Federal Committee on Statistical Methodology Research Conference* 3: 1–11. November 14–16, 2001, Washington DC, U.S.A. Available at: http://fcsm.sites. usa.gov/files/2014/05/2001FCSM_Burck.pdf (accessed October 2014).
- Chen, Z.G., P. Wong, M. Morry, and H. Fung. 2003. Variance Estimation for X-11 Seasonal Adjustment Procedure: Spectrum Approach and Comparison. Statistics Canada (Report BSMD-2003-001E).
- Findley, D.F. and D.E.K. Martin. 2006. "Frequency Domain Diagnostics of SEATS and X-11/12-ARIMA Seasonal Adjustment Filters for Short and Moderate-Length Time Series." *Journal of Official Statistics* 22: 1–34.
- Gómez, V. and A Maravall. 1996. Programs TRAMO and SEATS, Introduction for User (Beta Version). Banco de España: Banco de España Working Papers 9628.
- Harvey, A.C. 1989. *Forecasting Structural Time Series With the Kalman Filter*. Cambridge: Cambridge University Press.

- Hilmer, S.C. and G.S. Tiao. 1982. "An ARIMA-Model-Based Approach to Seasonal Adjustment." *Journal of the American Statistical Association* 77: 63–70. DOI: http://dx.doi.org/10.1080/01621459.1982.10477767
- Pfeffermann, D. 1994. "A General Method for Estimating the Variances of X-11 Seasonally Adjusted Estimators." *Journal of Time Series Analysis* 15: 85–116. DOI: http://dx.doi.org/10.1111/j.1467-9892.1994.tb00179.x
- Pfeffermann, D., M. Morry, and P. Wong. 1995. "Estimation of the Variances of X-11 ARIMA Seasonally Adjusted Estimators for a Multiplicative Decomposition and Heteroscedastic Variances." *International Journal of Forecasting* 11: 271–283. DOI: http://dx.doi.org/10.1016/0169-2070(94)00573-U
- Pfeffermann, D. and S. Scott. 1997. "Variance Measures for X-11 Seasonally Adjusted Estimators: Some Developments with Application to Labor Force Series." In Proceedings of the Section on Business & Economic Statistics: American Statistical Association, 211–216. August 10–14, 1997, Anaheim, California, U.S.A.
- Pfeffermann, D., M. Feder, and D. Signorelli. 1998. "Estimation of Auto-correlations of Survey Errors with Application to Trend Estimation in Small Areas." *Journal of Business and Economic Statistics* 16: 339–348.
- Pfeffermann, D., S. Scott, and R. Tiller. 2000. "Comparison of Variance Measures for Seasonally Adjusted and Trend Series." In *Proceedings of the 2nd International Conference* on *Establishment Surveys*, 755–764. June 17–21, 2000, Buffalo, New York, U.S.A.
- Pfeffermann, D. and R. Tiller. 2005. "Bootstrap Approximation to Prediction MSE for State-Space Models with Estimated Parameters." *Journal of Time Series Analysis* 26: 893–916. DOI: http://dx.doi.org/10.1111/j.1467-9892.2005.00448.x
- Scott, S., M. Sverchkov, and D. Pfeffermann. 2012. "Estimating Variance in X-11 Seasonal Adjustment." In *Economic Time Series: Modeling and Seasonality*, edited by William R. Bell, Scott H. Holan, and Tucker S. McElroy, 185–210. London: Chapman and Hall.
- Tiller, R.B. 1992. "Time Series Modeling of Sample Survey Data from the U.S. Current Population Survey." *Journal of Official Statistics* 8: 149–166.
- Tiller, R.B. 2012. "Frequency Domain Analysis of Seasonal Adjustment Filters Applied to Periodic Labor Force Survey Series." In *Economic Time Series: Modeling and Seasonality*, edited by William R. Bell, H. Holan Scott, and Tucker S. McElroy, 135–158. London: Chapman and Hall.
- Wecker, W.E. 1979. "A New Approach to Seasonal Adjustment." In Proceedings of the Section on Business and Economic Statistics: American Statistical Association, 322–323. August 13–16, 1979, Washington DC, U.S.A.
- *X-13A-S Reference Manual, Version 0.1 (Beta).* Time Series Staff, Statistical Research Division, Room 3000-4, U.S. Census Bureau, Washington, DC 20233-9100. Available at: https://www.google.com/?gws_rd=ssl#q=X-13-A-S+Reference+Manual%2c+Version+0.1+(Beta).

Received December 2012 Revised July 2014 Accepted July 2014