On a family of finite moving-average trend filters for the ends of series*

Alistair G. Gray
Statistics New Zealand
Wellington, New Zealand

Peter J. Thomson†
Statistics Research Associates Ltd
Wellington, New Zealand

9 March 2000

Abstract

A family of finite end filters is constructed using a minimum revisions criterion and based on a local dynamic model operating within the span of a given finite central filter. These end filters are equivalent to evaluating the central filter with unavailable future observations replaced by constrained optimal linear predictions. Two prediction methods are considered; best linear unbiased prediction and best linear biased prediction where the bias is time invariant. The properties of these end filters are determined. In particular, they are compared to X-11 end filters and to the case where the central filter is evaluated with unavailable future observations predicted by global ARIMA models as in X-11-ARIMA or X-12-ARIMA.

Keywords: Local dynamic model; minimum revisions; best linear unbiased prediction; best linear biased prediction.

---

*To appear in the *Journal of Forecasting*.

†Formerly at the Institute of Statistics and Operations Research, Victoria University of Wellington, New Zealand.

1 Introduction

Many seasonal adjustment procedures decompose time series into trend, seasonal, irregular and other components using non-seasonal finite moving-average trend filters. This paper is concerned with the extension of the finite central moving-average trend filter used in the body of the series to the ends where there are missing observations.

For any given finite central moving-average trend filter, a family of end filters is constructed using a minimum revisions criterion and a local dynamic model operating within the span of the central filter. These end filters are equivalent to evaluating the central filter with unavailable future observations replaced by constrained optimal linear predictors. Two prediction methods are considered. Best Linear Unbiased Prediction (BLUP) and Best Linear Biased Prediction, where the bias is time invariant (BLIP). The BLIP end filters are shown to be a generalisation of those developed by Musgrave (1964) for the central X-11 Henderson filters and include the BLUP end filters as a special case.

The theoretical properties of BLUP's and BLIP's are examined. In particular, it is established that the BLIP end filters generally have smaller mean squared revisions than the BLUP end filters. However, unlike the BLUP filters, the BLIP filters are no longer independent of the parameters in the local dynamic model and so, in practice, it is possible that a mis-specification of these parameters will lead to BLIP end filters with greater mean squared revisions than BLUP end filters. The effects of such mis-specification are discussed. Comparisons are also made between these end filters and the Musgrave end filters used by X-11, and the end filters obtained when the central filter is evaluated with unavailable future observations predicted by global ARIMA models. The latter parallels the ARIMA forecast extension method used in X-11-ARIMA (Dagum (1980)) and X-12-ARIMA (Findley et al. (1998)). Finally these filters are evaluated on some New Zealand time series.

2 Local dynamic model

The conventional paradigm for trend filter design is to consider a moving window of $n = 2r + 1$ observations within which an estimate of the trend is to be calculated for the central time point.

Within the finite window we choose to model the observations as

$$y_t = g_t + \epsilon_t \tag{1}$$

where the trend g_t is given by

$$g_t = \sum_{j=0}^{p} \beta_j t^j + \xi_t, \tag{2}$$

and ϵ_t is a white noise process with variance σ^2. The zero mean stochastic process ξ_t is assumed to be correlated, but uncorrelated with ϵ_t, and ξ_t, ϵ_t are assumed to be not both zero. In particular we consider the situation where the β_j and σ^2 are parameters local to the window, but p, n and the model for ξ_t/σ involve global parameters which are constant across windows. Thus, although the parameters involved with the mean and variance of y_t vary across windows, the autocorrelation structure of y_t will be a function of time invariant parameters in addition to time itself.

Loosely speaking, the finite polynomial is intended to capture deterministic low order polynomial trend whereas ξ_t is intended to capture smooth deviations from the polynomial trend. Note that it is the incorporation of ξ_t which distinguishes this local model from the standard situation where it is zero. Among the anticipated benefits of including ξ_t are lower values of p and improved performance at the ends of series, in particular the recent end.

Because the window is not likely to be large the model will need to involve as few parameters as possible on the one hand, while allowing for a sufficiently flexible family of forms for g_t on the other. *With these points in mind we choose to model ξ_t as a (possibly integrated) random walk with initial value zero.* In particular, if Δ denotes the backwards difference operator satisfying $\Delta X_t = X_t - X_{t-1}$, we have in mind the situation where $\Delta^{p+1} g_t = \Delta^{p+1} \xi_t$ is a zero mean stationary process within the window. In keeping with this rationale, we shall always assume that the levels of integration of the random walk components that make up ξ_t do not exceed $p+1$.

This seems an appropriate and parsimonious model which should account for smooth deviations from the deterministic polynomial trend component. It also provides a dynamic trend model for g_t which is essentially of the same form as that used in the global structural models that have been successfully applied to economic and official data. See Akaike (1980), Schlicht (1981),Hillmer and Tiao (1982), Gersch and Kitagawa (1983), Bell and Hillmer (1984), Harvey (1989), Maravall (1993), Young et al. (1999) for example. For these global models there is no need for special end filters since they are automatically provided using the global model and, in most cases, recursive algorithms such as the Kalman filter and smoother. The final trend filters implied by these global procedures are infinite central moving-average filters. By contrast, here we consider local models and the situation where the central moving-average trend filters are finite as is the case in X11-ARIMA, X-12-ARIMA, SABL (Cleveland et al. (1978)) and STL (Cleveland et al. (1990)).

In the local linear case $p = 1$ a simple dynamic model for g_t is given by

$$y_t = g_t + \epsilon_t = \beta_0 + \beta_1 t + \xi_t + \epsilon_t \tag{3}$$

where β_0, β_1 are constants and ξ_t is a simple random walk satisfying

$$\xi_t = \xi_{t-1} + \eta_t$$

with $\xi_0 = 0$. Here $\Delta g_t = \beta_1 + \eta_t$ and ϵ_t, η_t are mutually uncorrelated white noise processes with variances σ^2 and $\sigma_\eta^2 = \lambda\sigma^2$ respectively. In the local quadratic case $p = 2$ a simple dynamic model for g_t is given by

$$y_t = g_t + \epsilon_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \xi_t + \epsilon_t \tag{4}$$

where β_0, β_1, β_2 are constants and ξ_t is a simple random walk satisfying the same conditions as in the local linear model. Preliminary analysis indicates that these models have properties that can be regarded as representative of other more general models of the type discussed above.

3 Trend filter design at the ends

Now consider a finite window of width $n = 2r+1$ points centred at time point t and within which the observations follow the local dynamic model given in Section 2. We consider the case where the trend g_t is to be estimated by a given finite central moving-average trend filter

$$\hat{g}_t = \sum_{s=-r}^{r} w_s y_{t+s}. \tag{5}$$

In keeping with standard practice, we assume that the w_s are constrained by the requirement that

$$E\{\hat{g}_t - g_t\} = 0 \tag{6}$$

so that $\hat{g}_t$ is an unbiased estimator of g_t. Note that this condition is equivalent to the requirement that the w_s satisfy

$$\sum_{s=-r}^{r} w_s = 1 \qquad \sum_{s=-r}^{r} s^j w_s = 0 \qquad (0 < j \leq p) \tag{7}$$

so that the moving-average filter passes polynomials of degree p.

At the ends of series the central moving-average filter (5) will involve unavailable future observations. How these missing observations should be treated is open to question.

A common and natural approach involves forecasting the missing values, either implicitly or explicitly, and then applying the desired central filter. The forecasting methods used range from simple extrapolation to model based methods, some based on the local trend model adopted, others based on global models for the series as a whole. The latter include the fitting of ARIMA models to produce forecasts (see Dagum (1980) in particular). The principle of using prediction at the ends of series seems a key one which goes back to DeForest (1877). See also the discussion in Cleveland (1983), Greville (1979) and Wallis (1983).

Yet another way to handle the missing values in the window is to employ additional criteria specific to the ends of the series. An important requirement, especially among official statisticians, is to keep seasonal adjustment revisions and therefore trend revisions to a minimum as more data comes to hand. Thus, at the ends of series, a natural criterion to consider is

$$R_q = E\{(\sum_{s=-r}^{r} w_s y_{t+s} - \tilde{g}_t)^2\} \tag{8}$$

where $\tilde{g}_t$ is a predictor of $\hat{g}_t = \sum_{s=-r}^{r} w_s y_{t+s}$ based on the data available. In general, given a history of observations $y_1, \ldots, y_T$, it is evident that R_q is minimised when

$$\tilde{g}_t = \sum_{s=-r}^{r} w_s \hat{y}_{t+s} \tag{9}$$

where $\hat{y}_{t+s} = E(y_{t+s} | y_1, \ldots, y_T)$ denotes the best predictor of y_{t+s} in the usual mean squared error sense. Thus there is a close relationship between the minimum revisions approach and that of forecasting the missing values in the window.

The minimum revisions strategy appears to have been originally proposed by Musgrave (1964) for the case where $\tilde{g}_t$ is restricted to be linear in the observations within the window. (See the discussion in Doherty (1991).) This approach has also been adopted by Lane (1972), Laniel (1986) and will also be adopted here. Geweke (1978) and Pierce (1980) established the result (9) for the case where $\tilde{g}_t$ is linear in past values of the time series (not just those within the window) and where the time series follows an appropriate global model. However the argument leading to (9) shows that in general $\hat{y}_{t+s}$, and hence $\tilde{g}_t$, need not necessarily be linear in the observations.

In this paper we adopt the minimum revisions strategy based on the moving window paradigm and the local dynamic trend model of Section 2. At the ends of the series we choose to predict $\hat{g}_t = \sum_{s=-r}^{r} w_s y_{t+s}$ by a linear predictor of the form

$$\tilde{g}_t = \sum_{s=-r}^{q} u_s y_{t+s} \tag{10}$$

where $q = T - t$ with $0 \leq q < r$, T denotes the time point of the last observation and the u_s are dependent on q. We shall consider two cases. The first imposes the condition that $\tilde{g}_t$ be an unbiased predictor of $\sum_{s=-r}^{r} w_s y_{t+s}$. The second weakens this requirement by considering biased predictors such as those developed by Musgrave (1964) for X-11 (see in particular Doherty (1991)).

3.1 Unbiased predictors

If $\tilde{g}_t$ is to be an unbiased predictor of $\sum_{s=-r}^{r} w_s y_{t+s}$ given by (5) and (7), then the u_s must satisfy

$$\sum_{s=-r}^{q} u_s = 1 \qquad \sum_{s=-r}^{q} s^j u_s = 0 \tag{11}$$

where $0 < j \leq p$. Thus the asymmetric moving-average filter implied by (10) passes polynomials of degree p. Moreover R_q is now given by

$$R_q = \mathbf{v}^T \mathbf{E}_1 \mathbf{v} \tag{12}$$

where $\mathbf{v}$ has typical element

$$v_s = \begin{cases} w_s - u_s & (-r \leq s \leq q) \\ w_s & (q < s \leq r) \end{cases} \tag{13}$$

and

$$\mathbf{E}_1 = \sigma^2 \mathbf{I} + \mathbf{\Omega}. \tag{14}$$

Here $\mathbf{I}$ denotes the identity matrix and the covariance matrix $\mathbf{\Omega}$ has typical element

$$\mathbf{\Omega}_{jk} = \mathrm{cov}(\xi_{t+j} - \xi_t, \xi_{t+k} - \xi_t) \tag{15}$$

for $-r \leq j, k \leq r$. Note, in particular, that $\mathbf{\Omega}$ does not depend on t, the absolute value of time indexing the origin of the window. This natural and important invariance property is a consequence of (7), (11) and the assumption that the levels of integration of the random walk components that make up ξ_t do not exceed $p + 1$.

For each q, appropriate values of the u_s can now be determined by minimising R_q subject to (11). As we show below, this results in an end filter that satisfies a particular form of (9) involving optimal prediction.

First consider predicting

$$Y = \sum_{s=-r}^{r} \delta_s y_{t+s}$$

from $y_{t-r}, \ldots y_{t+q}$ by a linear predictor of the form

$$\hat{Y} = \sum_{s=-r}^{q} u_s y_{t+s}$$

where $q < r$ and the δ_s are arbitrary known values. Then, in terms of the local dynamic model that applies in the window, $\hat{Y}$ is the best linear unbiased predictor (BLUP) of Y if the u_s are chosen so that $E(Y - \hat{Y}) = 0$ (unbiased prediction error) and the mean squared error criterion $E\{(Y - \hat{Y})^2\}$ is a minimum.

Additional notation is needed to present the results. For the model specified by (1) and (2) define the $n \times (p+1)$ dimensional matrix $\mathbf{C}$ and the $p+1$ dimensional vector $\mathbf{c}$ by

$$\mathbf{C} = \begin{pmatrix} 1 & -r & \dots & (-r)^p \\ 1 & -r+1 & \dots & (-r+1)^p \\ . & . & \dots & . \\ . & . & \dots & . \\ 1 & r-1 & \dots & (r-1)^p \\ 1 & r & \dots & r^p \end{pmatrix} \tag{16}$$

and $\mathbf{c} = (1, 0, \dots, 0)^T$. As well, define the $n \times (q+r+1)$ matrix $\mathbf{L_1}$ and the $n \times (r-q)$ matrix $\mathbf{L_2}$ by the relation

$$\mathbf{I} = [\mathbf{L_1}, \mathbf{L_2}]. \tag{17}$$

We now establish the following result.

Theorem 1 *Let y_t follow the local dynamic model specified by (1) and (2). Given $\delta = (\delta_{-r}, \dots, \delta_r)^T$ and observations $y_{t-r}, \dots, y_{t+q}$ for $0 \le q < r$, the BLUP of $\sum_{s=-r}^{r} \delta_s y_{t+s}$ is $\sum_{s=-r}^{q} u_s y_{t+s}$ where $\mathbf{u} = (u_{-r}, \dots, u_q)^T$ is given by*

$$\mathbf{u} = \mathbf{L}_1^T(\mathbf{I} - \mathbf{G}\mathbf{L}_2(\mathbf{L}_2^T\mathbf{G}\mathbf{L}_2)^{-1}\mathbf{L}_2^T)\delta.$$

Here

$$\mathbf{G} = \mathbf{E}_1^{-1} - \mathbf{E}_1^{-1}\mathbf{C}(\mathbf{C}^T\mathbf{E}_1^{-1}\mathbf{C})^{-1}\mathbf{C}^T\mathbf{E}_1^{-1},$$

$\mathbf{E}_1$ is given by (14) and $\mathbf{C}$ is given by (16).

In particular the BLUP of $\sum_{s=-r}^{r} \delta_s y_{t+s}$ is given by $\sum_{s=-r}^{r} \delta_s \hat{y}_{t+s}$ where $\hat{y}_{t+s}$ is the BLUP of y_{t+s} for $q < s \le r$ and y_{t+s} otherwise.

Proof For $\hat{Y} = \sum_{s=-r}^{q} u_s y_{t+s}$ to be an unbiased predictor of $Y = \sum_{s=-r}^{r} \delta_s y_{t+s}$ the u_s must satisfy

$$\sum_{-r}^{q} s^j u_s = \sum_{-r}^{r} s^j \delta_s \qquad (0 \le j \le p). \tag{18}$$

Given (18)

$$Y - \hat{Y} = \sum_{s=-r}^{r} v_s(\xi_{t+s} - \xi_t + \epsilon_{t+s})$$

where

$$v_s = \begin{cases} \delta_s - u_s & (-r \le s \le q) \\ \delta_s & (q < s \le r) \end{cases} \tag{19}$$

and

$$\sum_{-r}^{r} s^j v_s = 0 \qquad (0 \le j \le p). \tag{20}$$

In terms of the v_s these conditions can be written as

$$\mathbf{L}_2^T\mathbf{v} = \mathbf{L}_2^T\delta, \qquad \mathbf{C}^T\mathbf{v} = \mathbf{0} \tag{21}$$

where $\mathbf{v} = (v_{-r}, \ldots, v_r)^T$ and $\mathbf{C}$ is given by (16). We now need to minimise

$$E\{(Y - \hat{Y})^2\} = \mathbf{v}^T(\sigma^2\mathbf{I} + \mathbf{\Omega})\mathbf{v} = \mathbf{v}^T\mathbf{E}_1\mathbf{v} \tag{22}$$

subject to (21).

Using Lagrange multipliers, optimising with respect to the v_s, and incorporating the constraints (21) yields the equations

$$\begin{aligned}
\mathbf{v} &= \mathbf{E}_1^{-1}\mathbf{C}\mu + \mathbf{E}_1^{-1}\mathbf{L}_2\nu \\
\mathbf{L}_2^T\mathbf{E}_1^{-1}\mathbf{C}\mu + \mathbf{L}_2^T\mathbf{E}_1^{-1}\mathbf{L}_2\nu &= \mathbf{L}_2^T\delta \\
\mathbf{C}^T\mathbf{E}_1^{-1}\mathbf{C}\mu + \mathbf{C}^T\mathbf{E}_1^{-1}\mathbf{L}_2\nu &= \mathbf{0}
\end{aligned}$$

where μ, ν are the vectors of Lagrange multipliers. These together with (19) yield

$$\mathbf{u} = \mathbf{L}_1^T(\mathbf{I} - \mathbf{G}\mathbf{L}_2(\mathbf{L}_2^T\mathbf{G}\mathbf{L}_2)^{-1}\mathbf{L}_2^T)\delta$$

as stated.

Now the BLUP of y_{t+s} is obtained by setting $\delta = \mathbf{0}$ with the exception of the s-th element which is set to unity. Then the BLUP of y_{t+s} is given by

$$\hat{y}_{t+s} = \begin{cases} y_{t+s} & (-r \leq s \leq q) \\ \mathbf{h}_s^T\mathbf{y} & (q < s \leq r) \end{cases}$$

where $\mathbf{h}_s$ is the s-th column of $-\mathbf{L}_1^T\mathbf{G}\mathbf{L}_2(\mathbf{L}_2^T\mathbf{G}\mathbf{L}_2)^{-1}\mathbf{L}_2^T$ and $\mathbf{y} = (y_{t-r}, \ldots, y_{t+q})^T$. Thus, for arbitrary choice of δ,

$$\begin{aligned}
\sum_{s=-r}^{q} u_s y_{t+s} &= \delta^T(\mathbf{I} - \mathbf{G}\mathbf{L}_2(\mathbf{L}_2^T\mathbf{G}\mathbf{L}_2)^{-1}\mathbf{L}_2^T)^T\mathbf{L}_1\mathbf{y} \\
&= \sum_{s=-r}^{r} \delta_s \hat{y}_{t+s}
\end{aligned}$$

as required.

□

Replacing the arbitrary δ_s by the weights w_s of the central moving-average trend filter yields the following result.

Corollary 1 *Let y_t follow the local dynamic model specified by (1), (2) and let $\mathbf{w}$ denote the vector of weights w_s for the central filter used in the body of the series with the w_s satisfying (7). Furthermore let $\tilde{g}_t = \sum_{s=-r}^{q} u_s y_{t+s}$ be a linear unbiased predictor of $\sum_{s=-r}^{r} w_s y_{t+s}$ with the u_s satisfying (11). Then, for $0 \leq q < r$, the values of u_s that minimise R_q subject to (11) are given by Theorem 1 with $\delta = \mathbf{w}$ and*

$$\sum_{s=-r}^{q} u_s y_{t+s} = \sum_{s=-r}^{r} w_s \hat{y}_{t+s}.$$

Here $\hat{y}_{t+s}$ is the BLUP of y_{t+s} for $q < s \leq r$ and y_{t+s} otherwise.

Proof When $\delta = \mathbf{w}$ the conditions (11) and (18) are equivalent since the w_s satisfy (7). From Theorem 1, the BLUP predictor of $\sum_{s=-r}^{r} w_s y_{t+s}$ is given by $\sum_{s=-r}^{r} w_s \hat{y}_{t+s}$ and this minimises $E\{(\sum_{s=-r}^{r} w_s y_{t+s} - \tilde{g}_t)^2\} = R_q$ as required.

□

Because of their dependence on BLUP predictors we shall henceforth refer to the end filters specified by Corollary 1 as *BLUP end filters*. The properties of these end filters are investigated in Section 4.

3.2 Biased predictors

Again consider the situation where $\tilde{g}_t$ is a linear predictor of the form (10), but now no longer required to be unbiased. Instead we require the bias to be time invariant in the sense that it does not depend on the absolute time t indexing the origin of the window, whatever the parameters of the local dynamic model adopted. Unlike the end filters based on BLUP predictors, the end filters we construct will no longer be independent of these parameters. The requirement that the bias be time invariant does, however, lead to relatively straightforward procedures for estimating the parametric quantities involved. By virtue of the local dynamic models adopted, the end filters derived generalise and extend the current X-11 end filters which were developed by Musgrave (1964) and placed in a prediction context by Doherty (1991).

Following the development in Section 3.1 we first consider predicting $Y = \sum_{s=-r}^{r} \delta_s y_{t+s}$ using a linear predictor of the form $\hat{Y} = \sum_{s=-r}^{q} u_s y_{t+s}$ where $q < r$ is given and the δ_s are arbitrary known values. In general, the mean squared error criterion $E\{(Y-\hat{Y})^2\}$ is given by

$$E\{(Y-\hat{Y})^2\} = (\sum_{s=-r}^{r} v_s E y_{t+s})^2 + \mathrm{Var}\{\sum_{s=-r}^{r} v_s y_{t+s}\} \tag{23}$$

where $\mathbf{v}$ is given by (19) with typical element $v_s = \delta_s - u_s$ when $-r \le s \le q$ and δ_s otherwise. The bias term in (23) can be written as

$$\sum_{s=-r}^{r} v_s E y_{t+s} = \sum_{s=-r}^{r} v_s \sum_{j=0}^{p} \beta_j (t+s)^j = \sum_{k=0}^{p} (\sum_{j=0}^{p-k} \beta_{j+k} \, {}^{j+k}C_k \sum_{s=-r}^{r} v_s s^j) t^k \tag{24}$$

and this will be invariant to the location of the window's time origin t when $p = 0$ or if and only if

$$\sum_{s=-r}^{r} s^j v_s = 0 \qquad (0 \le j < p). \tag{25}$$

when $p > 0$. Note that (25) is equivalent to

$$\sum_{s=-r}^{q} s^j u_s = \sum_{s=-r}^{r} s^j \delta_s \qquad (0 \le j < p). \tag{26}$$

If $p > 0$ and the u_s satisfy (26) then

$$\mathrm{Var}\{\sum_{s=-r}^{r} v_s y_{t+s}\} = E\{(\sum_{s=-r}^{r} v_s(\xi_{t+s} - \xi_t + \epsilon_{t+s}))^2\}$$

and (23) becomes

$$E\{(Y-\hat{Y})^2\} = \beta_p^2 (\sum_{s=-r}^{r} s^p v_s)^2 + \mathbf{v}^T \mathbf{E}_1 \mathbf{v} \tag{27}$$

with $\mathbf{v}$ given by (19) and $\mathbf{E}_1 = \sigma^2 \mathbf{I} + \mathbf{\Omega}$ as in (14). As before, it is desirable for the mean squared error $E\{(Y-\hat{Y})^2\}$ as well as the bias to be time invariant.

However, as indicated in the discussion following Theorem 3 in Gray and Thomson (1996), this will only be the case if (20) holds and the levels of integration of the random walk components that make up ξ_t do not exceed $p+1$. Here the v_s satisfy (25) rather than (20). Thus, to ensure that the mean squared error for the biased predictors is time invariant, we need to impose the stronger condition that the levels of integration of the random walk components that make up ξ_t do not exceed p. In the case where $p = 0$ this necessarily leads to the requirement that $\xi_t = 0$ and so (27) continues to hold for $p = 0$ with $\mathbf{E}_1 = \sigma^2\mathbf{I}$.

These observations lead us to consider a *restricted local dynamic model* for the window centred at t where the levels of integration of the random walk components that make up ξ_t do not exceed p. For such a model we define $\hat{Y}$ to be the best linear time invariant predictor (BLIP) of Y if the u_s are chosen to satisfy (26) and the mean squared error criterion $E\{(Y - \hat{Y})^2\}$ is a minimum. Thus the expected prediction error $E(Y - \hat{Y})$ and the mean squared error $E\{(Y - \hat{Y})^2\}$ do not depend on t whatever the parameters of the local dynamic model concerned.

Theorem 2 *Let y_t follow the restricted local dynamic model specified above and let $Y = \sum_{s=-r}^{r} \delta_s y_{t+s}$ where $\delta = (\delta_{-r}, \ldots, \delta_r)^T$ is known. Given observations $y_{t-r}, \ldots, y_{t+q}$ and q satisfying $0 \le q < r$, the BLIP of Y is $\hat{Y} = \sum_{s=-r}^{q} u_s y_{t+s}$ where $\mathbf{u} = (u_{-r}, \ldots, u_q)^T$ is given by*

$$\mathbf{u} = \mathbf{L}_1^T(\mathbf{I} - \tilde{\mathbf{G}}\mathbf{L}_2(\mathbf{L}_2^T\tilde{\mathbf{G}}\mathbf{L}_2)^{-1}\mathbf{L}_2^T)\delta$$

with

$$\tilde{\mathbf{G}} = \tilde{\mathbf{E}}_1^{-1} - \tilde{\mathbf{E}}_1^{-1}\mathbf{C}_{p-1}(\mathbf{C}_{p-1}^T\tilde{\mathbf{E}}_1^{-1}\mathbf{C}_{p-1})^{-1}\mathbf{C}_{p-1}^T\tilde{\mathbf{E}}_1^{-1}.$$

Here

$$\tilde{\mathbf{E}}_1 = \begin{cases} \mathbf{E}_1 + \beta_p^2\mathbf{c}_p\mathbf{c}_p^T & (p > 0) \\ \sigma^2\mathbf{I} + \beta_0^2\mathbf{c}_0\mathbf{c}_0^T & (p = 0) \end{cases}.$$

and $\mathbf{E}_1$ is as given by (14). The $n \times p$ matrix $\mathbf{C}_{p-1}$ and the p dimensional vector $\mathbf{c}_p$ are defined implicitly by the partitioned matrix

$$\mathbf{C} = [\mathbf{C}_{p-1}, \mathbf{c}_p]$$

where $\mathbf{C}$ is given by (16) and $\mathbf{C}_{p-1}$ is null when $p = 0$.

In particular the BLIP of $\sum_{s=-r}^{r} \delta_s y_{t+s}$ is given by $\sum_{s=-r}^{r} \delta_s \hat{y}_{t+s}$ where $\hat{y}_{t+s}$ is the BLIP of y_{t+s} for $q < s \le r$ and y_{t+s} otherwise.

Proof The proof of this result proceeds in much the same way as Theorem 1. Consider first the case $p > 0$. Using Lagrange multipliers we optimise (27) subject to (26) and obtain

$$\begin{aligned} \mathbf{v} &= \tilde{\mathbf{E}}_1^{-1}\mathbf{C}_{p-1}\mu + \tilde{\mathbf{E}}_1^{-1}\mathbf{L}_2\nu \\ \mathbf{L}_2^T\tilde{\mathbf{E}}_1^{-1}\mathbf{C}_{p-1}\mu + \mathbf{L}_2^T\tilde{\mathbf{E}}_1^{-1}\mathbf{L}_2\nu &= \mathbf{L}_2^T\delta \\ \mathbf{C}_{p-1}^T\tilde{\mathbf{E}}_1^{-1}\mathbf{C}_{p-1}\mu + \mathbf{C}_{p-1}^T\tilde{\mathbf{E}}_1^{-1}\mathbf{L}_2\nu &= \mathbf{0} \end{aligned}$$

where μ, ν are the vectors of Lagrange multipliers. Solving these equations together with (19) yields the stated result. The case $p = 0$ follows similarly.

Setting $\delta = \mathbf{0}$ with the exception of the s-th element which is set to unity, we see that the BLIP of y_{t+s} is

$$\hat{y}_{t+s} = \begin{cases} y_{t+s} & (-r \le s \le q) \\ \tilde{\mathbf{h}}_s^T\mathbf{y} & (q < s \le r) \end{cases}$$

where $\tilde{\mathbf{h}}_s$ is the s-th column of $-\mathbf{L}_1^T\tilde{\mathbf{G}}\mathbf{L}_2(\mathbf{L}_2^T\tilde{\mathbf{G}}\mathbf{L}_2)^{-1}\mathbf{L}_2^T$ and $\mathbf{y} = (y_{t-r}, \ldots, y_{t+q})^T$. Thus, for arbitrary choice of δ,

$$\sum_{s=-r}^{q} u_s y_{t+s} = \delta^T(\mathbf{I} - \tilde{\mathbf{G}}\mathbf{L}_2(\mathbf{L}_2^T\tilde{\mathbf{G}}\mathbf{L}_2)^{-1}\mathbf{L}_2^T)^T\mathbf{L}_1\mathbf{y} = \sum_{s=-r}^{r} \delta_s\hat{y}_{t+s}$$

as required.

□

Note that the BLIP predictor given by Theorem 2 has the form of a shrinkage estimator since it is exactly the same as the BLUP predictor given by Theorem 1, but with $\mathbf{E}_1$ replaced by $\mathbf{E}_1 + \beta_p^2 \mathbf{c}_p \mathbf{c}_p^T$ and $\mathbf{C}$ replaced by $\mathbf{C}_{p-1}$. Indeed, when $\beta_p = 0$, the BLIP predictor becomes the BLUP predictor for the reduced model where g_t is replaced by

$$g_t = \begin{cases} \sum_{j=0}^{p-1} \beta_j t^j + \xi_t & (p > 0) \\ 0 & (p = 0) \end{cases},$$

but ξ_t and ϵ_t remain the same.

We now return to the minimisation of the revisions criterion R_q given by (8) where the w_s are the known central filter weights that apply in the body of the series and the w_s satisfy (7). Given a restricted local dynamic model, we choose to predict $\sum_{s=-r}^{r} w_s y_{t+s}$ by a linear predictor $\tilde{g}_t$ of the form (10) subject to the requirement that the bias of $\tilde{g}_t$ is time invariant. The latter quantity is given by (24) with $w_s = \delta_s$ $(-r \leq s \leq r)$. Thus, from (26) and (27), $\tilde{g}_t$ is a linear time invariant predictor of $\sum_{s=-r}^{r} w_s y_{t+s}$ if

$$\sum_{s=-r}^{q} u_s = 1, \qquad \sum_{s=-r}^{q} s^j u_s = 0 \qquad (0 < j < p) \tag{28}$$

when $p > 0$ and, for $p \geq 0$,

$$R_q = \beta_p^2 (\sum_{s=-r}^{q} s^p v_s)^2 + \mathbf{v}^T \mathbf{E}_1 \mathbf{v} \tag{29}$$

with

$$\beta_p^2 (\sum_{s=-r}^{q} s^p v_s)^2 = \begin{cases} \beta_p^2 (\sum_{s=-r}^{q} s^p u_s)^2 & (p > 0) \\ \beta_0^2 (1 - \sum_{s=-r}^{q} u_s)^2 & (p = 0) \end{cases}.$$

Here $\mathbf{v}$ is given by (13) and R_q is also time invariant. This leads to the following result which is a direct application of Theorem 2.

Corollary 2 *Let y_t follow the restricted local dynamic model specified in Theorem 2 and let $\mathbf{w}$ denote the vector of weights w_s for the central filter used in the body of the series with the w_s satisfying (7). Furthermore, let $\tilde{g}_t = \sum_{s=-r}^{q} u_s y_{t+s}$ be a linear predictor of $\sum_{s=-r}^{r} w_s y_{t+s}$ with time invariant bias so that the u_s satisfy (28). Then, for $0 \leq q < r$, the values of u_s that minimise R_q subject to (28) are given by Theorem 2 with $\delta = \mathbf{w}$ and*

$$\sum_{s=-r}^{q} u_s y_{t+s} = \sum_{s=-r}^{r} w_s \hat{y}_{t+s}.$$

Here $\hat{y}_{t+s}$ is the BLIP of y_{t+s} for $q < s \leq r$ and y_{t+s} otherwise.

Because of their dependence on BLIP predictors we shall henceforth refer to the end filters specified by Corollary 2 as *BLIP end filters*. Their properties are investigated in Section 4.

Now $\mathbf{E}_1/\sigma^2$ does not depend on σ^2 and $\tilde{\mathbf{E}}_1$ need only be known up to a constant of proportionality. Thus, unlike the BLUP end filters specified by Corollary 1, the BLIP end filters specified by Corollary 2 can only be made operational when β_p^2/σ^2 is known. Since this will rarely, if ever, be the case, estimates of β_p^2/σ^2 of one form or another need to be determined from the data. Such estimates will, necessarily, differ from their true values and it is therefore important to determine the effects of mis-specification of β_p^2/σ^2. This issue is addressed below in Theorem 3 and also Section 4 where the properties of these end filters and the estimation of β_p^2/σ^2 are discussed.

Note that Corollary 2 yields the X-11 end filters derived by Musgrave (1964) and Doherty (1991) when $\xi_t = 0$, $p = 1$, $\beta_p^2/\sigma^2 = 4/(\pi(3.5)^2)$ and $\mathbf{w}$ contains the Henderson central filter weights. (See, for example Henderson (1924) and the alternative derivations given in, e.g. Kenny and Durbin (1982), Gray and Thomson (1996).) Thus Corollary 2 provides a generalisation and extension of the current X-11 end filters.

The following result considers an alternative form of the BLIP end filters given by Corollary 2 which explicitly builds on the corresponding BLUP end filters given by Corollary 1. In addition the result provides a means of exploring the effects of mis-specification of β_p^2/σ^2.

Theorem 3 *Let y_t follow the restricted local dynamic model specified in Theorem 2 and let $\mathbf{w}$ denote the vector of weights w_s for the central filter used in the body of the series with the w_s satisfying (7). Given observations $y_{t-r}, \ldots, y_{t+q}$ and q satisfying $0 \leq q < r$, let $\tilde{g}_t = \sum_{s=-r}^{q} u_s(\phi) y_{t+s}$ be a linear predictor of $\sum_{s=-r}^{r} w_s y_{t+s}$ where $\mathbf{u}(\phi) = (u_{-r}(\phi), \ldots, u_q(\phi))^T$ is defined for every scalar ϕ by*

$$\mathbf{u}(\phi) = \mathbf{L}_1^T(\mathbf{I} - \mathbf{G}\mathbf{L}_2(\mathbf{L}_2^T\mathbf{G}\mathbf{L}_2)^{-1}\mathbf{L}_2^T)(\mathbf{w} + \phi\gamma)$$

and

$$\gamma = \mathbf{E}_1^{-1}\mathbf{C}(\mathbf{C}^T\mathbf{E}_1^{-1}\mathbf{C})^{-1}\mathbf{d}.$$

Here $\mathbf{L}_1$, $\mathbf{L}_2$, $\mathbf{G}$, $\mathbf{E}_1$ and $\mathbf{C}$ are as given in (17), Theorem 1, (14) and (16) respectively, and the $p+1$ dimensional vector $\mathbf{d}$ is zero save for the last element which is unity. Then

(a) $\tilde{g}_t(\phi)$ is a linear time invariant predictor of $\sum_{s=-r}^{r} w_s y_{t+s}$ with squared bias $\beta_p^2\phi^2$ and the $u_s(\phi)$ satisfy (28) when $p > 0$;

(b) the revisions criterion R_q is given in this case by

$$R_q(\phi) = \mathbf{w}^T\mathbf{H}\mathbf{w} + 2\phi\gamma^T\mathbf{H}\mathbf{w} + \phi^2(\beta_p^2 + \mathbf{d}^T(\mathbf{C}^T\mathbf{E}_1^{-1}\mathbf{C})^{-1}\mathbf{d} + \gamma^T\mathbf{H}\gamma)$$

where

$$\mathbf{H} = \mathbf{L}_2(\mathbf{L}_2^T\mathbf{G}\mathbf{L}_2)^{-1}\mathbf{L}_2^T;$$

(c) the optimal end filters of Corollary 1 and Corollary 2 are given by $\mathbf{u}(0)$ and $\mathbf{u}(\phi_0)$ respectively where

$$\phi_0 = -\frac{\gamma^T\mathbf{H}\mathbf{w}}{\beta_p^2 + \mathbf{d}^T(\mathbf{C}^T\mathbf{E}_1^{-1}\mathbf{C})^{-1}\mathbf{d} + \gamma^T\mathbf{H}\gamma}$$

minimises $R_q(\phi)$.

Proof First observe from Theorem 1 that $\tilde{g}_t(\phi)$ is the BLUP of $\sum_{s=-r}^{r}(w_s + \phi\gamma_s)y_{t+s}$ and $\mathbf{C}^T\gamma = \mathbf{d}$ so that

$$\sum_{s=-r}^{q} s^p\gamma_s = 1, \qquad \sum_{s=-r}^{q} s^j\gamma_s = 0 \qquad\qquad (0 \leq j < p).$$

Thus, from (18),

$$\sum_{s=-r}^{q} s^j u_s(\phi) = \sum_{s=-r}^{r} s^j (w_s + \phi\gamma_s) = \begin{cases} 1 & (j=0) \\ 0 & (0<j<p) \\ \phi & (j=p) \end{cases}$$

when $p > 0$ and $\sum_{s=-r}^{q} u_s(\phi) = 1 + \phi$ when $p = 0$. Moreover

$$E\{\sum_{s=-r}^{r} w_s y_{t+s} - \tilde{g}_t(\phi)\} = -\phi \sum_{s=-r}^{r} \gamma_s E(y_{t+s}) = -\phi\beta_p$$

since $\mathbf{C}^T\gamma = \mathbf{d}$. This establishes (a).

Now, from (29) and (13),

$$R_q(\phi) = \beta_p^2\phi^2 + \mathbf{v}^T\mathbf{E}_1\mathbf{v}$$

where

$$\mathbf{v} = \mathbf{w} - \mathbf{L}_1\mathbf{u}(\phi).$$

Noting that $\mathbf{L}_2^T(\mathbf{I} - \mathbf{GH}) = \mathbf{0}$ we can write $\mathbf{v}$ as

$$\begin{aligned} \mathbf{v} &= \mathbf{w} - (\mathbf{I} - \mathbf{GH})(\mathbf{w} + \phi\gamma) \\ &= \mathbf{GHw} - \phi(\mathbf{I} - \mathbf{GH})\gamma \end{aligned} \tag{30}$$

and $R_q(\phi)$ becomes

$$\begin{aligned} R_q(\phi) &= \mathbf{w}^T\mathbf{HGE}_1\mathbf{GHw} - 2\phi\gamma^T(\mathbf{I} - \mathbf{HG})\mathbf{E}_1\mathbf{GHw} \\ &\quad + \phi^2(\beta_p^2 + \gamma^T(\mathbf{I} - \mathbf{HG})\mathbf{E}_1(\mathbf{I} - \mathbf{GH})\gamma^T). \end{aligned} \tag{31}$$

Since $\mathbf{GE}_1\mathbf{G} = \mathbf{G}$, $\mathbf{HGH} = \mathbf{H}$ and $\gamma^T\mathbf{E}_1\mathbf{G} = \mathbf{0}$, the above now reduces to the expression for $R_q(\phi)$ given by (b).

From the form of $\mathbf{u}(\phi)$ it follows directly that $\mathbf{u}(0)$ is the optimal end filter given by Corollary 1. All that remains to prove is that $\mathbf{u}(\phi_0)$ is the optimal end filter given by Corollary 2. Consider again minimising (29) with respect to the u_s where now $\mathbf{v}$ is defined by (13) and the u_s satisfy (28) if $p > 0$. This is equivalent to optimising

$$\tilde{R}_q = \beta_p^2\phi^2 + \mathbf{v}^T\mathbf{E}_1\mathbf{v} - 2\mu^T(\mathbf{C}^T\mathbf{v} + \phi\mathbf{d}) - 2\nu^T\mathbf{L}_2^T(\mathbf{v} - \mathbf{w})$$

with respect to $\mathbf{v}$, ϕ and the Lagrange multipliers μ, ν. Here ϕ is a function of $\mathbf{v}$ defined by

$$\phi = -\sum_{s=-r}^{r} s^p v_s = \begin{cases} \sum_{s=-r}^{q} s^p u_s & (p>0) \\ \sum_{s=-r}^{q} u_s - 1 & (p=0) \end{cases} . \tag{32}$$

where the latter equality follows from (7) and (13). Optimising $\tilde{R}_q$ first with respect to $\mathbf{v}$, μ and ν we obtain the equations

$$\begin{aligned} \mathbf{v} &= \mathbf{E}_1^{-1}\mathbf{C}\mu + \mathbf{E}_1^{-1}\mathbf{L}_2\nu \\ \mathbf{L}_2^T\mathbf{E}_1^{-1}\mathbf{C}\mu + \mathbf{L}_2^T\mathbf{E}_1^{-1}\mathbf{L}_2\nu &= \mathbf{L}_2^T\mathbf{w} \\ \mathbf{C}^T\mathbf{E}_1^{-1}\mathbf{C}\mu + \mathbf{C}^T\mathbf{E}_1^{-1}\mathbf{L}_2\nu &= -\phi\mathbf{d} \end{aligned}$$

and these, together with (13), yield $\mathbf{u} = \mathbf{u}(\phi)$. Substituting this solution back into $\tilde{R}_q$ gives $R_q(\phi)$ which must now be optimised with respect to ϕ. Since $R_q(\phi)$ is a quadratic in ϕ and takes its minimum value when $\phi = \phi_0$ result (c) follows.

□

An immediate consequence of Theorem 3 is that end filters based on BLIP predictors will generally have smaller mean squared revisions than those based on the corresponding BLUP predictors since $R_q(\phi_0) \leq R_q(0)$. Moreover

$$\begin{aligned} R_q(\phi_0) &= \mathbf{w}^T\mathbf{H}\mathbf{w} - \frac{(\gamma^T\mathbf{H}\mathbf{w})^2}{\beta_p^2 + \mathbf{d}^T(\mathbf{C}^T\mathbf{E}_1^{-1}\mathbf{C})^{-1}\mathbf{d} + \gamma^T\mathbf{H}\gamma} \\ &= R_q(0) - |\phi_0||\gamma^T\mathbf{H}\mathbf{w}| \end{aligned}$$

so that $R_q(\phi_0)$ is a monotonically increasing function of β_p^2/σ^2 and a linearly decreasing function in $|\phi_0|$. In particular

$$\lim_{\beta_p^2/\sigma^2 \longrightarrow \infty} R_q(\phi_0) = R_q(0), \qquad \lim_{\beta_p^2/\sigma^2 \longrightarrow \infty} \phi_0 = 0$$

so that BLIP end filters converge to their corresponding BLUP end filters as β_p^2/σ^2 increases. The best possible mean squared revisions are achieved when $\beta_p^2/\sigma^2 = 0$. Then $R_q(\phi_0)$ is least and, as noted following Theorem 2, the BLIP end filters become BLUP end filters for the local dynamic model with order $p-1$, but the same stochastic structure.

From the representation (30) and the properties of G and H, observe in passing that

$$\text{cov}\{\sum_{s=-r}^{r} w_s y_{t+s} - \tilde{g}_t(0), \sum_{s=-r}^{r} w_s y_{t+s} - \tilde{g}_t(\phi_0)\} = \mathbf{w}^T\mathbf{H}\mathbf{w} + \phi_0\gamma^T\mathbf{H}\mathbf{w} = R_q(\phi_0)$$

so that the normalised quantity

$$\frac{R_q(\phi_0)}{R_q(0)} = 1 - |\phi_0|\frac{|\gamma^T\mathbf{H}\mathbf{w}|}{\mathbf{w}^T\mathbf{H}\mathbf{w}}$$

represents the regression coefficient of the BLIP revisions $\sum_{s=-r}^{r} w_s y_{t+s} - \tilde{g}_t(\phi_0)$ on the BLUP revisions $\sum_{s=-r}^{r} w_s y_{t+s} - \tilde{g}_t(0)$. Moreover, as noted in the proof to Theorem 3, $\tilde{g}_t(\phi)$ is the BLUP of $\sum_{s=-r}^{r} w_s y_{t+s} + \phi \sum_{s=-r}^{r} \gamma_s y_{t+s}$. Here $\sum_{s=-r}^{r} \gamma_s y_{t+s}$ is the best linear unbiased estimator (BLUE) of β_p given $y_{t-r}, \ldots, y_{t+r}$.

Both the BLIP and BLUP end filters are dependent on the global parameters specified by p, n and the model for ξ_t/σ. Although these global parameters are sufficient to determine the BLUP end filters, the BLIP end filters further require knowledge of β_p^2/σ^2. The latter is a function of local parameters whose values will not normally be known in practice and which will need to be estimated from the data. In this case it is possible that a mis-specified value of β_p^2/σ^2 could result in a BLIP end filter which has greater mean squared revisions than its corresponding BLUP end filter.

Now $\tilde{g}_t(\phi_0)$ depends only on β_p^2/σ^2 through ϕ_0 which is a one-to-one function of β_p^2/σ^2. Thus Theorem 3 enables us to consider the effects of mis-specification of β_p^2/σ^2. Let $\hat{\phi}_0$ denote ϕ_0 evaluated at $\hat{\beta}_p^2/\hat{\sigma}^2$, some estimated or target value of β_p^2/σ^2. Then, since $R_q(\phi)$ is a quadratic in ϕ, it is evident that $R_q(\hat{\phi}_0) \leq R_q(0)$ if and only if $\hat{\phi}_0$ lies between 0 and $2\phi_0$. The BLIP end filters will therefore have better mean squared revisions than their corresponding BLUP end filters when

$$\frac{\beta_p^2}{\sigma^2} < 2\frac{\hat{\beta}_p^2}{\hat{\sigma}^2} + (\mathbf{d}^T(\mathbf{C}^T\mathbf{E}_1^{-1}\mathbf{C})^{-1}\mathbf{d} + \gamma^T\mathbf{H}\gamma)/\sigma^2. \tag{33}$$

In particular, since $R_q(\phi_0 + \phi) = R_q(\phi_0 - \phi)$, it is sufficient to select $\hat{\beta}_p^2/\hat{\sigma}^2$ so that $\hat{\phi}_0$ is between 0 and ϕ_0 or, equivalently, $\hat{\beta}_p^2/\hat{\sigma}^2 \geq \beta_p^2/\sigma^2$. In practice this choice should lead to values for $\hat{\beta}_p^2/\hat{\sigma}^2$ that, if anything, over-estimate β_p^2/σ^2 thus controlling the mis-specification error by shrinking the BLIP end filter towards its BLUP counterpart. Note that if $\hat{\beta}_p^2/\hat{\sigma}^2 \geq \beta_p^2/\sigma^2$ then

$$\tilde{g}_t(\hat{\phi}_0) = (1 - \frac{\hat{\phi}_0}{\phi_0})\tilde{g}_t(0) + \frac{\hat{\phi}_0}{\phi_0}\tilde{g}_t(\phi_0) \qquad (0 \leq \frac{\hat{\phi}_0}{\phi_0} \leq 1)$$

so that $\tilde{g}_t(\hat{\phi}_0)$ is a convex combination of the two optimal end filters.

The properties of these end filters are investigated in Section 4.

4 Properties of the Filters

This section considers the properties of the end filters specified by Corollary 1 and Corollary 2 which are designed to minimise the expected mean squared revisions between the output of these filters and that of the central filters on which they are based. These end filters deal with a transition problem that ultimately goes away as the current time points are subsumed into the body of the series. The minimum revisions criterion therefore provides a measure of the total cost of this transition.

We restrict attention to the important case where the window is of length 13 with $r = 6$ and the central filter used in the body of the series is the central Henderson filter. This case is the default used in the X-11 or X-12-ARIMA seasonal adjustment programs. Furthermore, we focus on two particular local dynamic models likely to be used in practice. These are the *local linear model* ($p = 1$) given by (3) and the *local quadratic model* ($p = 2$) given by (4).

Three end filters are considered: the BLUP end filter (or BLIP end filter with $\hat{\beta}_p^2/\hat{\sigma}^2 = \infty$), the BLIP end filter with $\hat{\beta}_p^2/\hat{\sigma}^2 = 4/(\pi(3.5)^2)$ (this gives the X-11 end filter in the case where $p = 1$ and $\lambda = 0$), and the BLIP end filter with $\hat{\beta}_p^2/\hat{\sigma}^2 = 0$. As noted in the discussion following Theorem 3, the smallest value of $R_q(\phi)$ is achieved when $\phi = \phi_0$ and $\beta_p^2/\sigma^2 = \hat{\beta}_p^2/\hat{\sigma}^2 = 0$. Also, if $\hat{\beta}_p^2/\hat{\sigma}^2$ is chosen so that (33) is satisfied or $\hat{\beta}_p^2/\hat{\sigma}^2 \geq \beta_p^2/\sigma^2$, then the largest value of $R_q(\hat{\phi}_0)$ is $R_q(0)$ which is the minimum revisions value for the BLUP end filter. These values serve as useful bounds for $R_q(\hat{\phi}_0)$.

To provide a range of possible values of β_p^2/σ^2 we briefly consider the so-called $\bar{I}/\bar{C}$ ratio used by X-11 to specify the length of the trend moving-average adopted. Here $\bar{I}$ and $\bar{C}$ are the respective averages of the absolute values of the month to month changes in the (estimated) irregular ϵ_t and trend g_t. Thus the $\bar{I}/\bar{C}$ ratio measures the importance of month to month changes in ϵ_t relative to those in the trend g_t. X-11 recommends that the central 9 point Henderson filter be used when $\bar{I}/\bar{C} < 1$, the central 13 point Henderson trend filter when $1 \leq \bar{I}/\bar{C} < 3.5$, and the central 23 point Henderson trend filter when $\bar{I}/\bar{C} \geq 3.5$. Following Musgrave (1964) and Doherty (1991) we consider the local linear model and

$$\frac{I}{C} = \frac{E|\Delta\epsilon_t|}{E|\Delta g_t|} \tag{34}$$

which can be thought of as the population parameter that $\bar{I}/\bar{C}$ is estimating. Under Gaussian assumptions

$$\frac{I}{C} = \frac{2}{\sqrt{\pi\tilde{\beta}_1^2}}(\Phi(\frac{|\tilde{\beta}_1|}{\sqrt{\lambda}}) - \Phi(-\frac{|\tilde{\beta}_1|}{\sqrt{\lambda}}) + \sqrt{\frac{2\lambda}{\pi\tilde{\beta}_1^2}}\exp(-\frac{\tilde{\beta}_1^2}{2\lambda}))^{-1}$$

where $\tilde{\beta}_1 = \beta_1/\sigma$ and $\Phi(x)$ is the standard normal cumulative distribution function. Simple evaluations show that the values of β_p^2/σ^2 for which $1 \leq I/C < 3.5$ lie in $[0, 4/\pi]$. This is the interval of values for β_p^2/σ^2 that we choose to adopt.

We now examine the mean squared revisions of the given end filters for β_p^2/σ^2 in $[0, 4/\pi]$. Note that it is sufficient to consider the extremes of this interval since, from Theorem 3, $R_q(\phi)$ is linearly increasing in β_p^2 over the interval for given ϕ. The mean squared revisions criterion $R_q(\hat{\phi}_0)/\sigma^2$ is plotted in Figure 1 as a function of q for a selection of models and for the particular end filters considered.

As expected, the mean squared revisions are greatest when q is least with $q = 0$ yielding the greatest revisions followed by $q = 1$. The mean squared revisions for the other values of q are negligible by comparison. Although not apparent on this scale, $R_q(\hat{\phi}_0)/\sigma^2$ is not necessarily a monotonic function of q. For example the local quadratic model yields mean squared revisions when $q = 2$ that are typically smaller than those for $q = 3$. This is a consequence of the shape of the central Henderson filter adopted.

If $\hat{\beta}_p^2/\hat{\sigma}^2$ has been selected appropriately, $R_q(\hat{\phi}_0)/\sigma^2$ will be bounded above by the mean squared revisions for the BLUP end filter and below by the mean squared revisions for the BLIP end filter with $\beta_p^2/\sigma^2 = \hat{\beta}_p^2/\hat{\sigma}^2 = 0$. These bounds are plotted in Figure 1 and show that there are gains to be had using BLIP end filters, although these are likely to be modest in the case of the local linear model. As evident from the form of $R_q(\hat{\phi}_0)/\sigma^2$ given by Theorem 3, the gains are greatest when β_p^2/σ^2 is least.

Note that the mean squared revisions generally increase as λ and p increase. This is of marginal utility in practice, since the local dynamic model chosen is determined from the particular time series concerned. However it is possible that a local linear model with large λ might describe a time series as well as a local quadratic model with small λ. From Figure 1 it would appear that, in such cases, the BLIP end filter for the quadratic model may have greater capacity to achieve lower revisions.

Consider now the BLIP end filters based on the X-11 value $\hat{\beta}_p^2/\hat{\sigma}^2 = 4/(\pi(3.5)^2)$. For the local linear model, these end filters appear to offer only marginal gains over the BLUP end filters when β_p^2/σ^2 satisfies (33). In the local quadratic case, they are clearly too conservative and a lower value of $\hat{\beta}_p^2/\hat{\sigma}^2$ might more profitably be considered. For both models, the upper limit of $\beta_p^2/\sigma^2 = 4/\pi$ for these particular end filters lead to end filters with unacceptably high revisions. This may explain, in part, the reason why forecast extension using ARIMA models has largely superseded the use of the X-11 end filters in practice. The former yield (global) BLUP end filters with properties that one might expect are close to the (local) BLUP end filters considered here. On the other hand, for the local linear model, the $\bar{I}/\bar{C}$ guidelines imply that the X-11 end filters are inflexibly applied whenever β_p^2/σ^2 satisfies $\beta_p^2/\sigma^2 \leq 4/\pi$ rather than (33).

It is of interest to consider the ranges of values of $(|\hat{\beta}_p|/\hat{\sigma}, |\beta_p|/\sigma)$ that satisfy (33) and to ascertain the general shape of $R_q(\hat{\phi}_0)/\sigma^2$ as a function of $(|\hat{\beta}_p|/\hat{\sigma}, |\beta_p|/\sigma)$ or, equivalently, $(\hat{\beta}_p^2/\hat{\sigma}^2, \beta_p^2/\sigma^2)$. Plots of the lines

$$\frac{|\beta_p|}{\sigma} = \sqrt{2\frac{\hat{\beta}_p^2}{\hat{\sigma}^2} + (\mathbf{d}^T(\mathbf{C}^T\mathbf{E}_1^{-1}\mathbf{C})^{-1}\mathbf{d} + \gamma^T\mathbf{H}\gamma)/\sigma^2}$$

are given in Figure 2 for $q = 3$, $q = 0$ and the local dynamic models (3) and (4) based on the central Henderson filter and selected values of λ. From (33) the BLIP end filters have better mean squared revisions when $(|\hat{\beta}_p|/\hat{\sigma}, |\beta_p|/\sigma)$ lies below the lines. For large

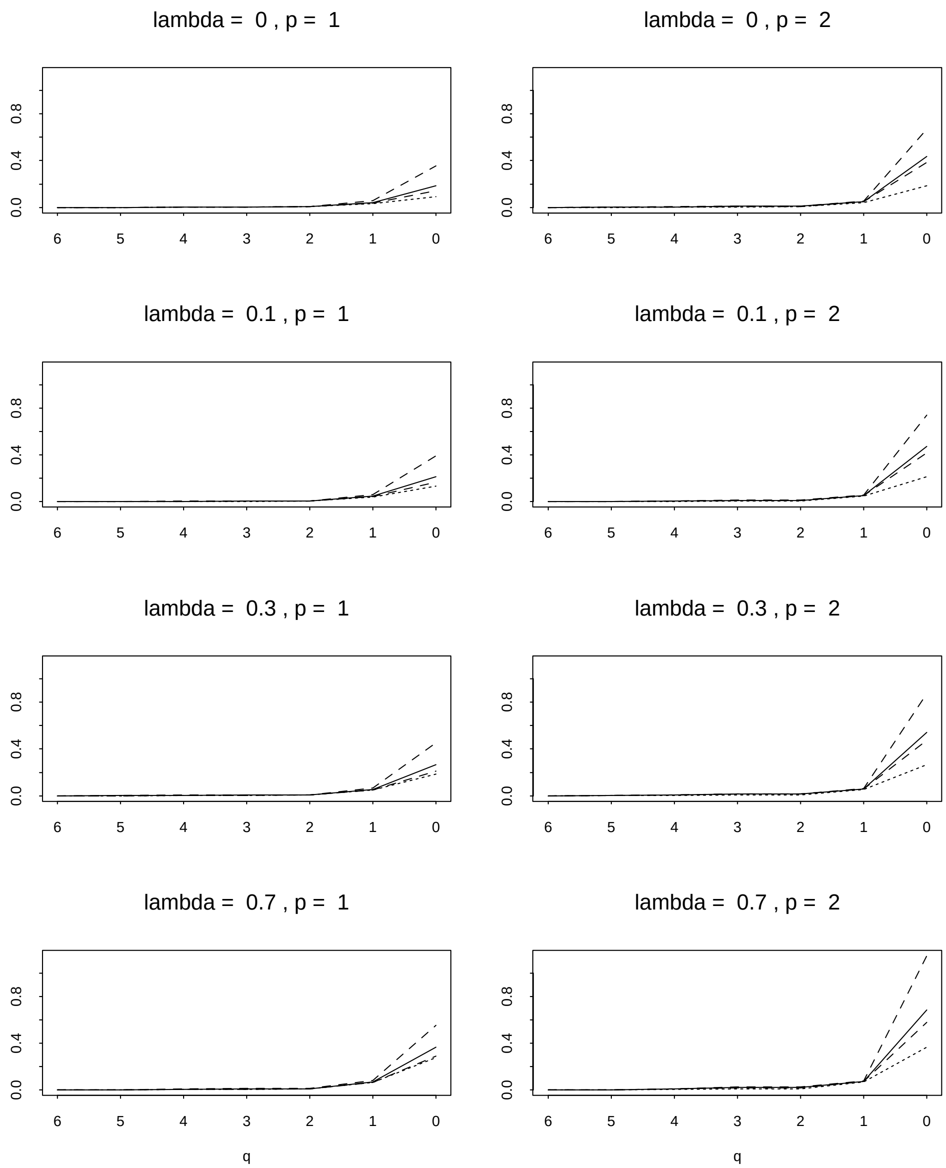

Figure 1: Plots of the mean squared revisions criterion R_q/σ^2 for the BLIP end filters based on the central 13 point Henderson filter and the local linear and quadratic models specified by λ and p. The solid lines correspond to the BLUP end filters or the BLIP end filters with $\hat{\beta}_p^2/\hat{\sigma}^2 = \infty$. The dashed lines correspond to $\hat{\beta}_p^2/\hat{\sigma}^2 = 4/(\pi(3.5)^2)$ (this value gives the X-11 end filters when $p = 1$, $\lambda = 0$) in the two cases where $\beta_p^2/\sigma^2 = 0$ (lower limit) and $\beta_p^2/\sigma^2 = 4/\pi$ (upper limit). The dotted lines correspond to $\hat{\beta}_p^2/\hat{\sigma}^2 = 0$ in the case where $\beta_p^2/\sigma^2 = 0$. This is the BLUP end filter for the reduced model of order $p-1$ and represents the best mean squared revisions possible.

values of $|\hat{\beta}_p|/\hat{\sigma}$ the BLIP end filters approach the BLUP end filters and the boundary becomes $|\beta_p|/\sigma = \sqrt{2}|\hat{\beta}_p|/\hat{\sigma}$. However, for smaller values of $|\hat{\beta}_p|/\hat{\sigma}$ a somewhat greater range of possibilities is evident, especially for the larger values of λ. There is less latitude when $p = 2$ as might be expected. Given a range of (time-varying) values of $|\beta_p|/\sigma$ one would wish to select $|\hat{\beta}_p|/\hat{\sigma}$ as low as possible in order to maximise the gains in terms of expected revisions.

Now consider Figure 3 which plots $R_q(\hat{\phi}_0)/R_q(0)$, the mean squared revisions of the BLIP end filters normalised by the mean squared revisions of the BLUP end filters, as a function of $(|\hat{\beta}_p|/\hat{\sigma}, |\beta_p|/\sigma)$. Here only $q = 0$ is shown and the BLIP end filters are based on the local dynamic models (3) and (4), the central Henderson filter and selected values of λ, $|\beta_p|/\sigma$. Given $|\hat{\beta}_p|/\hat{\sigma}$, it follows from Theorem 3 that the normalised revisions $R_q(\hat{\phi}_0)/R_q(0)$ increase quadratically with increasing $|\beta_p|/\sigma$. From the plots and Theorem 3 a three dimensional picture of $R_q(\hat{\phi}_0)/R_q(0)$ can now be visualised. Looking out from the origin along the line $|\beta_p|/\sigma = |\hat{\beta}_p|/\hat{\sigma}$, $R_q(\hat{\phi}_0)/R_q(0)$ has the appearance of a rising valley with a sharply increasing left hand side (representing the unacceptably high mean squared revisions incurred when $(|\hat{\beta}_p|/\hat{\sigma}, |\beta_p|/\sigma)$ does not satisfy (33)) and a right hand side that levels out at the BLUP limit $R_q(0)/R_q(0) = 1$. Note that the optimum mean squared revisions $R_q(\phi_0)/R_q(0)$ occur when $|\beta_p|/\sigma = |\hat{\beta}_p|/\hat{\sigma}$ and, as noted in the discussion following Theorem 3, $R_q(\phi_0)$ is a monotonically increasing function of $|\beta_p|/\sigma$. Clearly the greatest gains are to be had when β_p^2/σ^2 and hence $\hat{\beta}_p^2/\hat{\sigma}^2$ are small.

If the BLUP end filters yield results that are comparable to ARIMA forecast extension, then it would appear that judiciously selected BLIP end filters may offer modest performance gains in terms of improved revisions. However the price of this improvement is a better understanding of the time varying values of β_p^2/σ^2.

5 Practical Study

The outcomes of a preliminary study on selected official time series where the end filters given by Corollary 1 and Corollary 2 were compared with X-11 end filters and ARIMA forecast extension were reported in Gray and Thomson (1996). In this section we shall discuss salient features from that study and illustrate how the theoretical properties mentioned in the previous section are realized in practice.

5.1 Description of the time series

All the time series we examined were seasonal and some of them had large outliers. Since the trend filters under study are for non-seasonal time series and have not yet been modified to handle outliers, the first task in this analysis was to remove the seasonal component and large outliers from the time series so that we could examine the relative performance of these filters. For simplicity we used the modified seasonally adjusted series from the X-11 seasonal adjustment method as the data for this study. We chose a window of the series which we believed would give a good X-11 decomposition. Other than choosing the commonly used limits of 1.8 and 2.8 for treating outliers we used the default options of X-11 to produce the seasonally adjusted data.

The resulting non-seasonal times series presented different types of short and long term

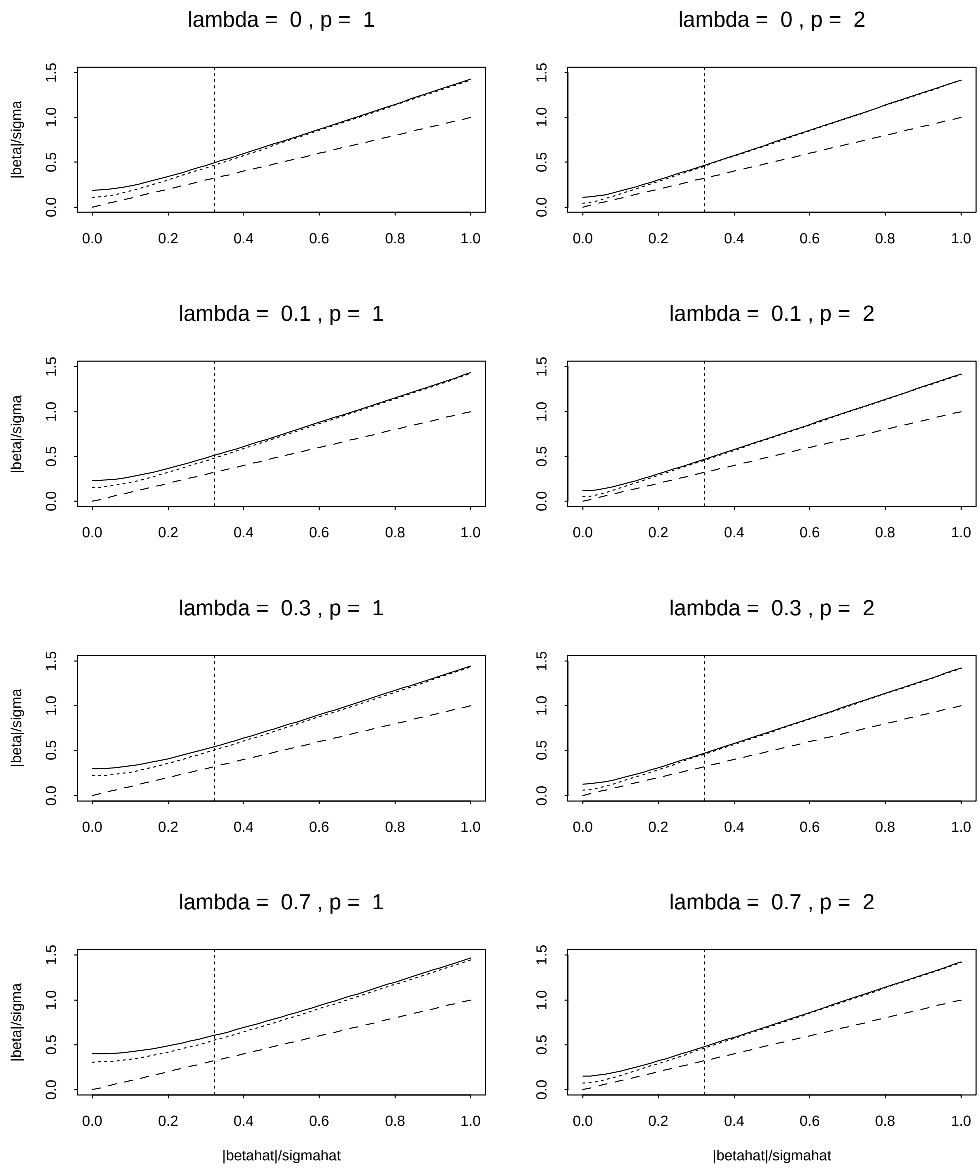

Figure 2: Plots of the lines in the $(|\hat{\beta}_p|/\hat{\sigma}, |\beta_p|/\sigma)$ plane below which BLIP end filters have better mean squared revisions than BLUP end filters. Here the BLIP end filters are based on the central 13 point Henderson filter, and the local linear and quadratic models are specified by p and λ. The solid lines correspond to $q = 0$, the dotted lines to $q = 3$ and the dashed line to the case $|\beta_p|/\sigma = |\hat{\beta}_p|/\hat{\sigma}$ when the optimum revisions are achieved for a given value of $|\beta_p|/\sigma$. The vertical dotted line is $|\hat{\beta}_p|/\hat{\sigma} = \sqrt{4/(\pi(3.5)^2)}$, a value derived from recommendations made by X-11 concerning I/C ratios.

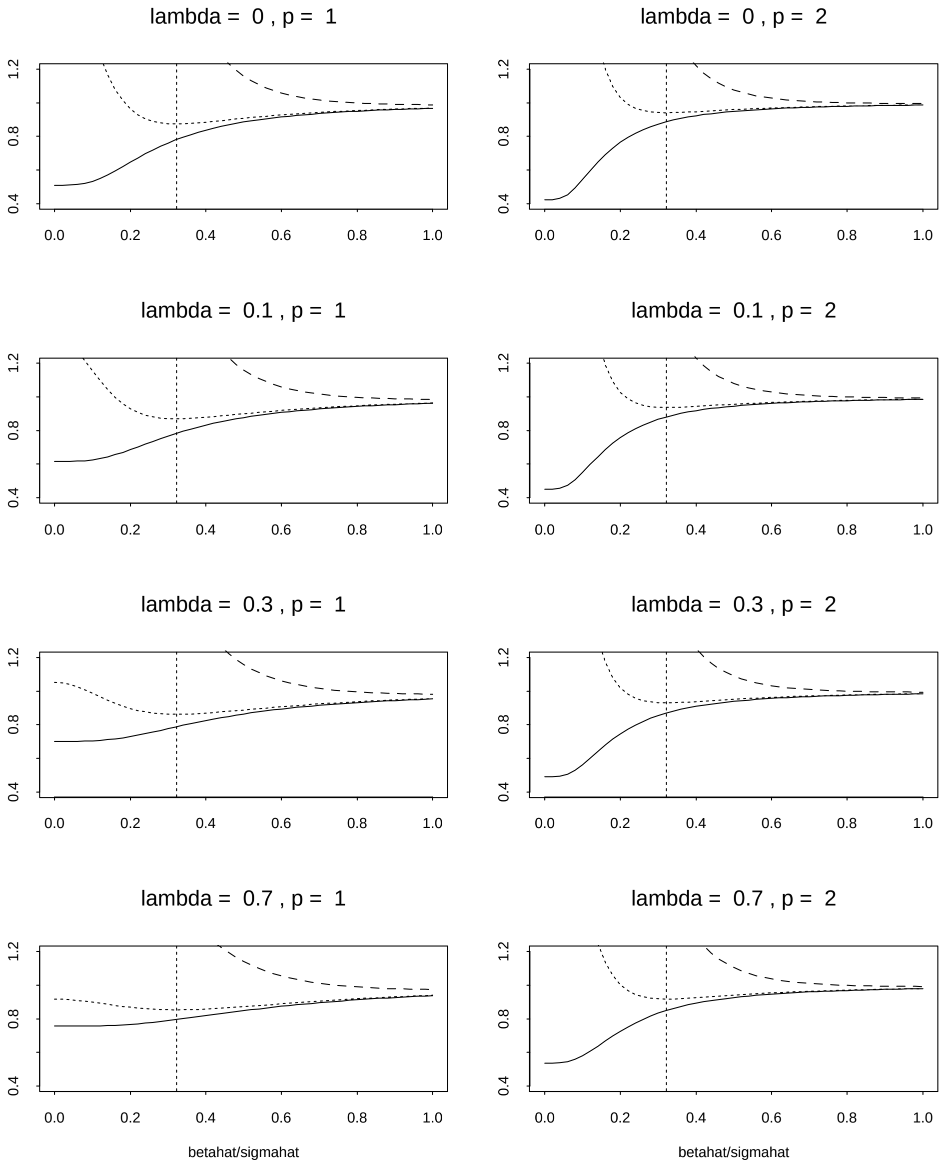

Figure 3: Plots of $R_q(\hat{\phi}_0)/R_q(0)$, the mean squared revisions of the BLIP end filters normalised by the mean squared revisions of the BLUP end filters, as a function of $(|\hat{\beta}_p|/\hat{\sigma}, |\beta_p|/\sigma)$. Here $q = 0$ and the BLIP end filters are based on the local linear ($p = 1$) and quadratic ($p = 2$) models, the central 13 point Henderson filter and selected values of λ, $|\beta_p|/\sigma$. For given $|\hat{\beta}_p|/\hat{\sigma}$ the normalised revisions $R_q(\hat{\phi}_0)/R_q(0)$ increase quadratically with increasing $|\beta_p|/\sigma$. The solid lines correspond to $\beta_p^2/\sigma^2 = 0$ (this yields the best mean squared revisions possible for any given value of $\hat{\beta}_p^2/\hat{\sigma}^2$), the dotted lines to $\beta_p^2/\sigma^2 = 4/(\pi(3.5)^2)$, the dashed lines to $\beta_p^2/\sigma^2 = 4/\pi$ and the vertical dotted line to $|\hat{\beta}_p|/\hat{\sigma} = \sqrt{4/(\pi(3.5)^2)}$. These values are derived from recommendations made by X-11 concerning I/C ratios.

trends and their local dynamic models included random walk components with a range of variances. In this section we shall discuss three monthly series representing this range of trend behaviour. These are as follows.

Building Permits New Zealand data from January 1985 to December 1992 on the number of Building Permits issued by Local Authorities for the construction of private houses and flats. This subset of the data was chosen in order to get a consistent series where, for example, there was no Government intervention in Housing Policy.

Merchandise Exports New Zealand data from January 1986 to December 1992 on the value of Merchandise Exports. This subset of the data was chosen in order to get a consistent series where the economy had adjusted to the impact of deregulation in the mid 1980s.

Permanent Migration New Zealand data on the Net Permanent and Long Term Migration from January 1982 to December 1992.

5.2 Fitting the Local Dynamic Model

Recall that the parameters of the local dynamic model are the window length $n = 2r + 1$, the order of the polynomial p, and a value of λ. To study the filters based on the local dynamic model all these parameters must be determined from the data, although to some extent the choice of n will be be influenced by the overall objectives of the smoothing.

Typically n and p are identified from a simple graphical analysis of the data although more formal methods could be used. Given n and p, there are a variety of methods for determining λ. These range from trial and error and simple variational arguments based on quantities like X-11's I/C ratio, through to likelihood analysis based on fitting the local dynamic model within non-overlapping windows, etc. Which approach turns out to be most useful is still very much an open question and the subject of further research. Here we adopted a simple and direct method of estimating λ that took advantage of the unbiased BLUP predictors.

Consider the family of BLUP end filters based on Corollary 1 for a given central filter and given values of n and p. Note that since these end filters are unbiased they depend only on q and λ. Applying the end filters to the data yields the revisions

$$r_t(q,\lambda) = \sum_{-r}^{r} w_s y_{t+s} - \sum_{-r}^{q} u_s(\lambda) y_{t+s}$$

where the w_s are the given central filter weights and the $u_s(\lambda)$ are the BLUP end filter weights given by Corollary 1. An appropriate cost function such as

$$P = \sum_{q=0}^{r-1} \alpha_q r_t^2(q,\lambda) \tag{35}$$

can now be constructed with the given positive weights α_q reflecting the relative costs of the respective revisions. Finally λ is determined by minimising P with respect to λ.

For our study $n = 13$, $p = 1$ and the central filter chosen was the central Henderson filter. Moreover we chose $\alpha_0 = 1$ and $\alpha_q = 0$ for $q \neq 0$ yielding a cost function which focused on the case of worst revisions. However, other central filters and other values for the α_q could have been chosen and would lead to different results. Again, more research is needed, but it is believed that the results presented here are indicative of what might be achieved more generally. With these caveats, this procedure was applied to Building Permits, Exports and Permanent Migration, yielding values for λ of 0.046, 0.52, and 3.8 respectively.

Finally we consider the estimation of the ratio $|\beta_p|/\sigma$ which involves the local parameters β_p and σ. Although local estimates of $|\beta_p|/\sigma$ should be explored in order to minimise revisions from BLIP end filters, this option was left for further study. Instead we adopted the more conservative strategy of estimating a common value for $|\beta_p|/\sigma$ for each of the series considered.

To determine a global value for $|\hat{\beta}_p|/\hat{\sigma}$ from the data we adopted a similar approach to that used for determining λ, but now focused on the BLIP end filters given by Corollary 2. Here we determined the revisions

$$\tilde{r}_t(q, |\hat{\beta}_p|/\hat{\sigma}) = \sum_{-r}^{r} w_s y_{t+s} - \sum_{-r}^{q} u_s(|\hat{\beta}_p|/\hat{\sigma}) y_{t+s}$$

where the w_s are the known central filter weights, the $u_s(|\hat{\beta}_p|/\hat{\sigma})$ are the BLIP end filter weights given by Corollary 2 and n, p are given. The value of λ is set at its previously estimated value. Now the cost function $\tilde{P}$ given by (35) with r_t replaced by $\tilde{r}_t$ can be evaluated and minimised with respect to $|\hat{\beta}_p|/\hat{\sigma}$.

As before we chose $\alpha_0 = 1$, $\alpha_q = 0$ for $q \neq 0$ and set $n = 13$, $p = 1$ with the central filter chosen to be the central Henderson filter. Minimising $\tilde{P}$ in this way for Building Permits, Exports and Permanent Migration, yielded global values for $|\hat{\beta}_p|/\hat{\sigma}$ of 0.22, 0.3, and 0.32 respectively. By contrast, the X-11 value for $|\hat{\beta}_p|/\hat{\sigma}$ is 0.322. As was noted before in relation to the estimation of λ, different central filters and different values of α_q could have been chosen and would give different results.

5.3 Comparative Performance

In this subsection we consider the comparative performance of the BLIP, BLUP, X-11 and ARIMA forecast extension filters on the data sets under study. Here, as before, the window length is $n = 13$ and only the local linear model ($p = 1$) is considered with λ and a global value for $|\hat{\beta}_p|/\hat{\sigma}$ determined by the methods described in Section 5.2. As mentioned in that section, a more appropriate procedure would have been to determine local values for $|\hat{\beta}_p|/\hat{\sigma}$ and use these in the BLIP end filters. However this was not done and remains a topic for further research. It might be expected that the use of local values for $|\hat{\beta}_p|/\hat{\sigma}$ would lead to BLIP end filters with improved performance over those using a global value for $|\hat{\beta}_p|/\hat{\sigma}$.

The following empirical comparison of the competing end filters was undertaken. First, for each local window we calculated the trend for the central point of that window using the central filter (5). Next, for each q and for each local window, we calculated the trend using the appropriate end filter. For the BLUP and BLIP cases, this meant predicting the trend as in (10), where the u_s are the weights of the end filter determined by Corollary 1 and 2 respectively. For the X-11 case the standard X-11 end filters were applied. Recall that these are given by Corollary 2 using the central Henderson filter, $\hat{\beta}_p^2/\hat{\sigma}^2 = 4/(\pi(3.5)^2)$ and the local linear model with $\lambda = 0$. For the ARIMA forecast extension case this meant predicting the trend as in (9) with y_{t+s} now predicted using a global ARIMA model fitted to all the data available up to and within the local window. Finally, for each of these methods, the differences between the central filter trend estimates and the end filter trend estimates were calculated to give the revisions which occur as predictions of missing data are replaced by their actual values. Boxplots of the absolute value of the revisions are given in Figure 4.

Further comments about the ARIMA forecast extension case are in order. We note that the resulting end filter is essentially that proposed in Dagum (1996) and, by contrast to the BLIP and BLUP end filters which use only observations within the local window, the ARIMA forecast extension end filter involves the available observations within the window and, in principle, all previous observations as well. To obtain appropriate ARIMA model parameters we fitted global ARIMA models to all the data. This was done for reasons of simplicity and to provide stable estimates. However, it puts our procedure at a disadvantage since the ARIMA global models are taking advantage of future values, which is not possible in practice. Using the AIC criterion and other standard diagnostics we selected ARIMA (0,1,1) models for Building Permits and Permanent Migration, and an ARIMA (2,1,0) model for Exports. Other global predictive models could have been chosen instead of ARIMA models including structural unobserved component models such as those given in Harvey (1989). However, this has not been investigated, in part because these particular models are a subset of the more general ARIMA model class considered here.

In accord with the discussion of the theoretical properties of these filters, Figure 4 shows that the greatest revisions and spread of revisions occurred when $q = 0$ followed by $q = 1$. It is not clear in making comparisons whether one should focus on the median or the lower quartile of the distribution of the revisions, or the upper quartile or the interquartile range, or even the outliers. Depending on the cost function associated with revisions, a case can be made for each of these. Here we focus on the median because that provides a robust estimate of the square root of the mean squared revisions (up to a scale factor). The mean squared revisions criterion was used to evaluate the theoretical performance of the filters in Section 4. To further assist comparison, Table 1 provides the values of the medians for each of the series for the various values of q.

As expected from Theorem 3, the BLIP end filters where $|\hat{\beta}_p|/\hat{\sigma}$ is determined from the data generally have smaller revisions than BLUP end filters and smaller revisions than the Musgrave end filters used in X-11. (These are BLIP end filters based on the central Henderson filter, a local linear model with $\lambda = 0$ and $\hat{\beta}_p^2/\hat{\sigma}^2$ set to $4/(\pi(3.5)^2)$.) For the Exports and Permanent Migration series, although the choice of $|\hat{\beta}_p|/\hat{\sigma}$ is similar to the X-11 value, the value of λ used in the X-11 end filters is inappropriate and so here, in the main, the X-11 end filters have larger revisions than the BLUP end filters. For all cases when $q = 0$ and for Building permits and Permanent Migration when $q = 1$, the BLIP end filter with $|\hat{\beta}_p|/\hat{\sigma}$ determined from the data performed at least as well as if not better than ARIMA forecast extension.

The results presented in this section are indicative and promising. A larger study is being planned to validate these results more generally. In particular the study will also check whether further improvement can be achieved by using local (adaptive) estimates of $|\hat{\beta}_p|/\hat{\sigma}$.

Acknowledgements

The second author gratefully acknowledges support provided by an ASA/NSF/Census Research Fellowship which he held at the US Bureau of the Census. In particular, both

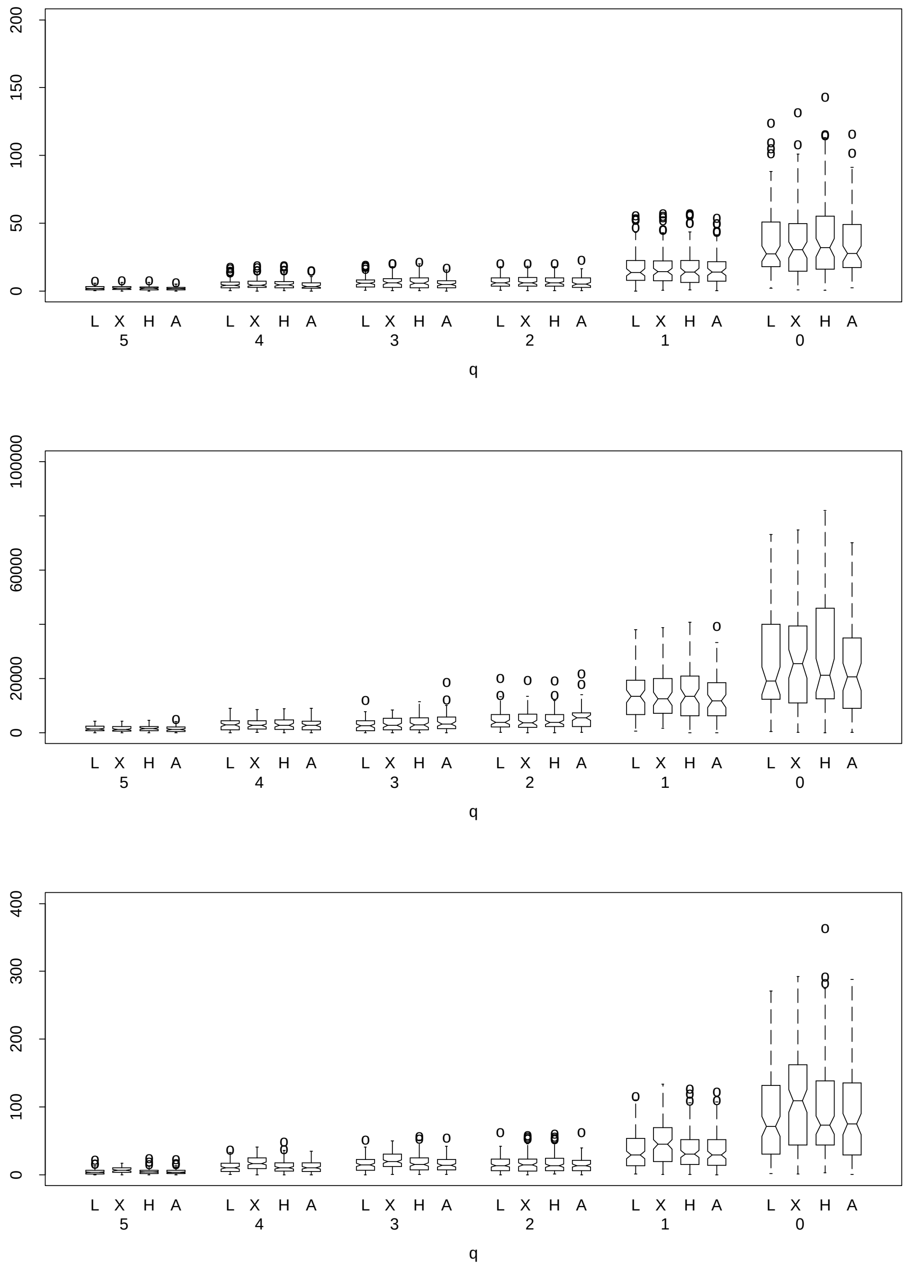

Figure 4: The top, middle and bottom graphs are for the Building Permits, Merchandise Exports, and Permanent Migration data respectively. Using the central 13 point Henderson filter in the body and the local linear model, comparisons are made between the various end filters and for the various values of q. The choice of λ is determined by searching for values which improve the revisions for the BLUP end filter for the $q = 0$ case. The choice of $\hat{\beta}_p^2/\hat{\sigma}^2$ is determined by searching for values which improve the revisions given λ. Here L refers to the BLIP end filter based on the local linear model and the choices of λ and $\hat{\beta}_p^2/\hat{\sigma}^2$ given in Section 5.2. Likewise, X refers to the standard X-11 end filter and H refers to the BLUP end filter based on the local linear model. Note that the X-11 end filters are BLIP end filters based on the central 13 point Henderson filter, a local linear model with $\lambda = 0$ and $\hat{\beta}_p^2/\hat{\sigma}^2 = 4/(\pi(3.5)^2)$. Finally A refers to the filter obtained by using the central filter with unavailable observations predicted by a global ARIMA model.

Table 1: Medians of absolute value of revisions for various end filters.

		Building Permits		Exports		Permanent Migration	
		m	d	m	d	m	d
$q=5$	L	1.8	-7.8	1503	21.0	3.7	-44.6
	X	1.9	-	1242	-	6.6	-
	H	1.9	-0.6	1562	25.7	4.3	-35.7
	A	1.6	-15.0	1247	0.4	3.7	-43.9
$q=4$	L	4.2	-0.6	2804	2.5	10.4	-37.2
	X	4.2	-	2737	-	16.5	-
	H	4.5	6.7	2852	4.2	10.4	-37.0
	A	3.3	-21.1	2827	3.3	10.4	-37.4
$q=3$	L	5.7	-4.3	2579	-7.0	14.4	-26.3
	X	6.0	-	2775	-	19.5	-
	H	5.8	-2.1	2889	4.1	15.0	-23.1
	A	4.7	-20.8	3299	18.9	13.9	-28.7
$q=2$	L	5.9	-0.5	3931	6.6	13.6	-8.2
	X	5.9	-	3688	-	14.8	-
	H	5.9	-0.7	3866	4.8	13.3	-10.6
	A	5.0	-16.1	5557	50.7	13.6	-8.5
$q=1$	L	13.8	-3.1	13459	8.0	29.2	-34.8
	X	14.2	-	12570	-	44.8	-
	H	13.9	-2.5	13534	7.7	30.7	-31.5
	A	14.1	-1.1	11777	-6.3	29.3	-34.5
$q=0$	L	27.5	-9.3	19187	-27.4	71.3	-34.6
	X	30.3	-	25472	-	108.9	-
	H	32.1	5.7	21201	-16.8	73.2	-32.8
	A	27.5	-9.2	20584	-19.2	74.9	-31.3

These abbreviations are used in the table above:

- m median of the absolute value of the revisions.
- d percentage differences (on unrounded values) between the median for the standard X-11 end filter and the medians for the other end filters.
- L the BLIP end filter based on the local linear model and where the choices of λ and $\hat{\beta}_p^2/\hat{\sigma}^2$ are given in Section 5.2.
- X the standard X-11 end filter, which is the BLIP end filter based on the central 13 point Henderson filter and a local linear model with $\lambda = 0$ and $\beta_p^2/\hat{\sigma}^2 = 4/(\pi(3.5)^2)$.
- H the BLUP end filter based on the local linear model.
- A the filter obtained by using the central 13 point Henderson filter with unavailable observations predicted by a global ARIMA model.

authors wish to record their gratitude to Dr David Findley of that organisation for the helpful advice and encouragement he has provided throughout the project. The data was kindly provided by Statistics New Zealand and the authors are grateful to the referees and the editor for suggesting a number of improvements.

References

Akaike, H. (1980). Seasonal adjustment by a Bayesian modeling. *Journal of Time Series Analysis*, 1(1):1–13.

Bell, W. R. and Hillmer, S. C. (1984). Issues involved with the seasonal adjustment of economic time series. *Journal of Business and Economic Statistics*, 2.

Cleveland, R. B., Cleveland, W. S., McRae, J. E., and Terpenning, I. J. (1990). STL: A seasonal-trend decomposition procedure based on loess. *Journal of Official Statistics*, 6:3–73.

Cleveland, W. S. (1983). Seasonal and calendar adjustment. In Brillinger, D. and Krishnaiah, P., editors, *Handbook of Statistics* **3**, pages 39–72. North Holland, Amsterdam.

Cleveland, W. S., Dunn, D. M., and Terpenning, I. J. (1978). SABL - A resistant seasonal adjustment procedure with graphical methods for interpretation and diagnosis. In Zellner, A., editor, *Seasonal Analysis of Economic Time Series*, pages 201–231. U.S. Department of Commerce, U.S. Bureau of the Census, Washington, DC.

Dagum, E. B. (1980). The X-11-ARIMA seasonal adjustment method. Research paper, Statistics Canada, Ottawa K1A 0T6.

Dagum, E. B. (1996). A new method to reduce unwanted ripples and revisions in trend-cycle estimates from X11ARIMA. *Survey Methodology*, 22(1):77–83.

DeForest, E. L. (1877). On adjustment formulas. *The Analyst*, 4:79–86, 107–113.

Doherty, M. (1991). Surrogate Henderson filters in X-11. Technical report, NZ Department of Statistics, Wellington, New Zealand.

Findley, D. F., Monsell, B. C., Bell, W. R., Otto, M. C., and Chen, B.-C. (1998). New capabilities and methods of the X-12-ARIMA seasonal-adjustment program. *Journal of Business and Economic Statistics*, 16(2):127–177.

Gersch, W. and Kitagawa, G. (1983). The prediction of time series with trends and seasonalities. *Journal of Business and Economic Statistics*, 1:253–264.

Geweke, J. (1978). The revision of seasonally adjusted time series. In *Proceedings of the Business and Economic Statistics Section, American Statistical Association*, pages 320–325.

Gray, A. G. and Thomson, P. (1996). Design of Moving-Average Trend Filters using Fidelity, Smoothness and Minimum Revisions Criteria. Research Report CENSUS/SRD/RR-96/1, Statistical Research Division, Bureau of the Census, Washington, D.C 20233-4200.

Greville, T. N. E. (1979). Moving-weighted-average smoothing extended to the extremities of the data. Technical Summary Report #2025, Mathematics Research Center, University of Wisconsin, Madison, Wisconsin.

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

Henderson, R. (1924). A new method of graduation. *Transactions of the Actuarial Society of America*, 25:29–40.

Hillmer, S. C. and Tiao, G. C. (1982). An ARIMA model based approach to seasonal adjustment. *Journal of the American Statistical Association*, 77:63–70.

Kenny, P. B. and Durbin, J. (1982). Local trend estimation and seasonal adjustment of economic and social time series. *Journal of the Royal Statistical Society, Series A*, 145:1–41.

Lane, R. O. D. (1972). Minimal revision trend estimates. Technical Report Research Exercise Note 8/72, Central Statistical Office, London.

Laniel, N. (1986). Design criteria for 13 term Henderson end-weights. Technical Report Working Paper TSRA-86-011, Statistics Canada, Ottawa K1A 0T6.

Maravall, A. (1993). Stochastic linear trends, models and estimators. *Journal of Econometrics*, 56:5–37.

Musgrave, J. C. (1964). A set of end weights to end all end weights. Working paper, Bureau of the Census, US Department of Commerce, Washington, D.C.

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

Schlicht, E. (1981). A seasonal adjustment principle and a seasonal adjustment method derived from this principle. *Journal of the American Statistical Association*, 76:374–378.

Wallis, K. F. (1983). Models for X-11 and X-11-FORECAST procedures for preliminary and revised seasonal adjustments. In Zellner, A., editor, *Applied Time Series Analysis of Economic Data*, pages 3–11, Washington, D. C. U. S. Bureau of the Census.

Young, P. C., Pedregal, D., and Tych, W. (1999). Dynamic harmonic regression (DHR). *J. Forecast.*, 18:369–394.