Empirical Comparison of Two Methods for Non-Gaussian Seasonal Adjustment

Andrew G. Bruce
and
Simon R. Jurke

May 27, 1992

Institute of Statistics and Operations Research
Victoria University
P.O. Box 600
Wellington
New Zealand

This work was started while Dr. Bruce was a visiting research statistician at the U.S. Bureau of the Census. It was completed through the support of a Joint Statistical Agreement between the U.S. Bureau of the Census and Victoria University of Wellington.

The authors are indebted to David Findley (U.S. Bureau of the Census), who inspired and supported this project. William Bell, Brian Monsell, and Mark Otto (U.S. Bureau of the Census) also gave invaluable support, providing the data, X-12-ARIMA software, REGCMPNT software, and the choice of ARIMA models for X-12-ARIMA. In addition, Genshiro Kitagawa provided software upon which the GAUSUM-STM program is based. Finally, Peter Thomson (Victoria University) and Alistair Gray (New Zealand Department of Statistics) contributed many insightful comments and guidance.

Andrew Bruce is currently a research scientist at Statistical Sciences Inc., 1700 Westlake Ave. N, Suite 500, Seattle, WA 98109, USA.

Abstract

This study compares two new seasonal adjustment methods designed to handle outliers and structural changes: X-12-ARIMA and GAUSUM-STM. X-12-ARIMA is a successor to the X-11-ARIMA seasonal adjustment method, and is being developed at the U.S. Bureau of the Census (Findley et al. (1988)). GAUSUM-STM is a non-Gaussian method using time series structural models, and was developed for this study based on methodology proposed by Kitagawa (1990).

The procedures are compared using 29 macroeconomic time series from the U.S. Bureau of the Census. These series have both outliers and structural changes, providing a good testbed for comparing non-Gaussian methods. For these series, the X- 12-ARIMA decomposition consistently leads to smoother seasonal factors which are as or more "flexible" than the GAUSUM-STM seasonal component. On the other hand, with some significant exceptions, GAUSUM-STM generally handles outliers and level shifts better than X-12-ARIMA. The differences between GAUSUM-STM and X-12-ARIMA in handling outliers and structural changes are swamped by the fundamental differences in the nature of the seasonal decompositions.

Recognizing that seasonal adjustment is a subjective enterprise, we feel the X-12-ARIMA procedure yields more appealing seasonal adjustments for most of the series examined. However, GAUSUM-STM potentially offers some important advantages. This study gives guidance on what problems need to be tackled to improve STM-based seasonal adjustments.

1 Introduction

Seasonal adjustment at most statistical agencies is currently done using a procedure based on the "X-11-ARIMA" method for seasonal adjustment. While X-11-ARIMA has proven to be reliable and effective, it is primarily an *ad hoc* method. Researchers have explored *model* based alternatives to X-11-ARIMA. Gaussian "time series structural models" (STM), introduced by Gersch and Kitagawa (1983) and Harvey and Todd (1983), are a class of models currently enjoying a surge of interest. Time series structural models are based on using simple intuitive component models for the trend, seasonal and irregular. In a comparison of X-11 and STM with relevant Dutch macroeconomic time series, den Butter and Mourik (1990) conclude that STM was a competitive method.

Outliers and "structural" changes (e.g., level shifts or ramps) cause problems with both X-11-ARIMA and the Gaussian STM methods. While X-11-ARIMA provides some protection against outliers, it is not fully robust and cannot handle level shifts or other structural changes. The X-12-ARIMA procedure, a successor to X-11-ARIMA, is being developed at the U.S. Bureau of the Census to handle additive outliers and level shifts (Findley et al. (1988), Monsell (1990)). Time series structural models can be adapted to non-Gaussian situations by assuming that the innovations of the component models are non-Gaussian (Kitagawa (1990)). For this study, we have developed "GAUSUM-STM", which extends the STM based seasonal adjustment to handle outliers and structural changes. GAUSUM-STM is derived from a computer program by Kitagawa (1991).

X-12-ARIMA and GAUSUM-STM differ in very significant ways. X-12-ARIMA is a nonparametric method while GAUSUM-STM is model based. The X-12-ARIMA seasonal filters are manually selected on the basis of diagnostic plots. By contrast, the seasonal decomposition of GAUSUM-STM is automatically obtained by maximizing the likelihood. Finally, the two procedures adopt very different methods for handling outliers and level shifts.

The X-12-ARIMA procedure is compared with GAUSUM-STM using 29 macroeconomic time series from the U.S. Bureau of the Census. These series have both outliers and structural changes, providing a good testbed for comparing non-Gaussian methods. For these series, the X-12-ARIMA decomposition consistently leads to smoother seasonal factors which are as or more "flexible" than the GAUSUM-STM seasonal component. On the other hand, with some significant exceptions, GAUSUM-STM generally handles outliers and level shifts better than X-12-ARIMA. The differences between GAUSUM-STM and X-12-ARIMA in handling outliers and structural changes are swamped by the fundamental differences in the nature of the seasonal decompositions.

Recognizing the seasonal adjustment is a subjective enterprise, we feel the X-12-ARIMA procedure yields more appealing seasonal adjustments for most of the series examined. GAUSUM-STM, or some similar procedure, is not yet a serious competitor. Perhaps the main contribution of this study is to give guidance on what problems need to be tackled to improve STM-based seasonal adjustments.

Section 2 discusses the outlier handling scheme for X-12-ARIMA. The GAUSUM-STM procedure is described in section 3. Section 4 describes the data and the associated filters and models. The plots and diagnostics used to assess and compare the seasonal adjustments are discussed in section 5. The heart of the paper lies in section 6, which summarizes the conclusions of the empirical comparison. Some additional issues regarding the GAUSUM-STM method are explored in section 7. Finally, conclusions and directions for future research are discussed in section 8.

2 X-12-ARIMA

The X-11 method for seasonal adjustment was developed at the U.S. Bureau of the Census by Shiskin et al. (1967). X-11-ARIMA is an extension of the X-11 method, developed at Statistics Canada by Dagum (1980). X-11-ARIMA eliminates the asymmetric filters of X-11 by using ARIMA models to forecast beyond the ends of the series. Both X-11 and X-11-ARIMA are nonparametric procedures, with a design based on practical considerations.

Numerous empirical studies have examined the X-11-ARIMA method: see, for example, Dagum (1978), Dagum and Morry (1984), den Butter et al. (1985), and Jain (1989). X-12-ARIMA offers several new features, including a new "language oriented" interface and the "sliding spans" diagnostics (Findley et al. (1990)). The primary new feature of interest in this study is the procedure for automatic detection of additive outliers and level shifts. This procedure is discussed in more detail below.

X-12-ARIMA outlier/level shift identification procedure

To avoid problems caused by additive outliers (AO's) and level shifts (LS's), X-12-ARIMA does a prior adjustment. AO's and LS's are identified using hypothesis tests based on the appropriate parametric intervention and ARIMA model. The series is adjusted using the estimated interventions.

The idea of doing hypothesis tests to identify the type of outlier was first introduced by Fox (1972). Suppose X_t is a time series which behaves according to the multiplicative Gaussian ARIMA $(p,d,q)\times(P,D,Q)_S$ model. Let Y_t be the observed series, which is related to X_t by

$$Y_t = X_t + \sum_{j=1}^{k} \zeta_j Z_t^{(j)}$$

Y_t might contain outliers, level shifts, etc., which are modeled by the 0-1 processes

$Z_t^{(j)}$ and the parameters ζ_j. To model an "additive outlier" (AO) at time T, we set

$$Z_s^{(j)} = \begin{cases} 1 & s = T \\ 0 & s \neq T \end{cases}$$

A "level shift" (LS) at time T is given by

$$Z_s^{(j)} = \begin{cases} 1 & s \geq T \\ 0 & s < T \end{cases}$$

A hypothesis test for the presence of an AO (or LS) at time T takes the form

$$\begin{aligned} \mathrm{H}_0: &\quad \zeta_j = 0 \\ \mathrm{H}_1: &\quad \zeta_j \neq 0 \end{aligned}$$

A large test statistic is indicative of an AO (or LS). These ideas generalize to other types of interventions, such as innovations outliers, ramps, or variance changes.

X-12-ARIMA incorporates tests for AO's and LS's in an iterative method for estimating parameters in a multiplicative seasonal ARIMA model. Suppose we have an initial estimate of the ARIMA parameters $\hat{\alpha}_0$. The algorithm proceeds as follows:

Step 0: $j \leftarrow 0$.

Forward Addition

Step 1: Given $\hat{\alpha}_j$, compute the t-statistics $\hat{\tau}_t^{\mathrm{AO}}$ and $\hat{\tau}_t^{\mathrm{LS}}$ corresponding to the hypothesis tests for AO's and LS's at times $t = 1,2,\ldots,N$.

Step 2: If

$$\max_t \left\{ \left|\hat{\tau}_t^{\mathrm{AO}}\right|, \left|\hat{\tau}_t^{\mathrm{LS}}\right| \right\} < C,$$

then go to step 5. Otherwise, flag the observation which is the most significant AO or LS according to the t-statistics.

Step 3: Subtract the least squares estimate $\hat{\zeta}_j$ of the flagged intervention from the series Y_t. Re-estimate the parameters $\hat{\alpha}_{j+1}$ with the adjusted data.

Step 4: $j \leftarrow j + 1$. Go to step 1.

Backward Elimination

Step 5: Let Ω be the set of indices corresponding to the identified AO's and LS's. Re-estimate the t-statistics for all identified AO's and LS's.

Step 6: If

$$\min_{t \in \Omega} \{|\hat{\tau}_t|\} > C,$$

then we are done. Otherwise, drop the least significant estimated intervention from the index set Ω and go to step 5.

For step 1, an efficient algorithm is available for ARIMA models, reducing the computational burden of computing the test statistic for all observations simultaneously. In step 2, the cutoff C is used to determine if there are any more significant AO's or LS's remaining in the series. In this study, a cutoff of $C = 3.1$ is used.

Note that only the most significant AO or LS is identified on each iteration. This "one-at-a-time" approach is computationally slower than identifying all significant AO's and LS's on each pass. However, for several of the series examined in this study, the multiple identification procedure is unstable and leads to poor decompositions. Hence, the multiple identification option is not recommended for general use.

Other types of interventions could be incorporated into the procedure. However, for economic time series, the most important and natural situations to attempt to model in this manner seem to be additive outliers and level shifts. This iterative identification procedure was first developed by Chang and Tiao (1983). Hillmer et al. (1983) applied this iterative estimation method in the context of ARIMA model based seasonal adjustment. See Bell (1986), Chang et al. (1988), and Tsay (1988) for further development of the method.

3 GAUSUM-STM

Many model based approaches to seasonal adjustment have been proposed. Advantages of model based seasonal adjustment are articulated by Bell and Hillmer (1984). Models provide an interpretable decomposition whose characteristics adapt to the nature of each series. One approach towards model based seasonal adjustment is based on fitting ARIMA models: see Box et al. (1978), Burman (1980), Hillmer and Tiao (1982), Maravall (1985), and Maravall and Pierce (1987). The ARIMA model is decomposed into trend, seasonal and irregular components, maximizing the variance of the irregular. This is often call the canonical decomposition.

In this study, we work with an approach based on time series structural models (STM). Gersch and Kitagawa (1983) and Harvey (1984) have explored the use of structural models for seasonal adjustment (see also Kitagawa and Gersch (1984), Harvey (1989), and den Butter and Mourik (1990)). There are several advantages of STM-based seasonal adjustment. Structural models are constructed by using simple component models for the trend, seasonal, and irregular. Hence, Harvey (1989) argues that structural models are more interpretable than ARIMA models. Harvey and Valls Pereira (1989) claim that structural models yield superior seasonal decompositions and adjustments than the canonical decomposition of ARIMA models. The fitting process is simpler for structural models, with only one or two basic model forms needed for a broad range of series (of the three models considered in this study, one model was consistently superior). Finally, structural models easily and naturally incorporate simple structural changes, such as level shifts and ramps. This is discussed further below.

3.1 Gaussian Time Series Structural Models

A time series structural model is based on forming models directly for each of the components in the decomposition

$$Y_t = T_t + S_t + I_t. \tag{1}$$

Y_t is some suitable transformation of the original observed series (in this study, Y_t is the log-transformed data). T_t, S_t, and I_t are the trend, seasonal and irregular. The irregular is usually considered to be Gaussian white noise with zero mean and variance σ_I^2. This is denoted by $I_t \sim \text{GWN}(0, \sigma_I^2)$.

A typical model for the trend is given by

$$T_t = T_{t-1} + b_{t-1} + \eta_t \tag{2}$$

where $\eta_t \sim \text{GWN}(0, \sigma_\eta^2)$. The term b_t acts as a "slope", and is permitted to evolve according to a random walk

$$b_t = b_{t-1} + \xi_t$$

where $\xi_t \sim \text{GWN}(0, \sigma_\xi^2)$.

Following Harvey (1984), we consider two different models for the seasonal component. Let s be the seasonal period (for monthly data $s = 12$). The first seasonal model, which makes up part of Harvey's "Basic Structural Model" (BSM), is defined by

$$S_t = -\sum_{j=1}^{s-1} S_{t-j} + \omega_t \tag{3}$$

where $\omega_t \sim \text{GWN}(0, \sigma_\omega^2)$. An alternative and more flexible seasonal model is given by

$$S_t = \sum_{j=1}^{[s/2]} \gamma_{j,t} \tag{4}$$

with

$$\begin{aligned} \gamma_{j,t} &= \gamma_{j,t-1} cos\lambda_j + \gamma_{j,t-1}^* sin\lambda_j + \omega_{j,t} \\ \gamma_{j,t}^* &= -\gamma_{j,t-1} sin\lambda_j + \gamma_{j,t-1}^* cos\lambda_j + \omega_{j,t}^* \end{aligned}$$

where $\lambda_j = 2\pi j/s$. The innovations $\omega_{j,t}$ and $\omega_{j,t}^*$ are uncorrelated with $\omega_{j,t} \sim \text{GWN}(0, \sigma_{\omega,j}^2)$ and $\omega_{j,t}^* \sim \text{GWN}(0, \sigma_{\omega,j}^2)$ for $j = 1, \ldots, 6$.

Modeling Calendar Effects

Structural models are easily extended to handled such things as calendar and holiday effects. Instead of (1), we might use the model

$$Y_t = T_t + C_t + S_t + I_t. \tag{5}$$

where C_t represents the calendar effect. C_t is estimated as a fixed effect by construction of the appropriate regression variables: see Bell and Hillmer (1983). Dynamic models for trading days also fit nicely within the structural model framework: see Monsell (1983) and Dagum et al. (1988).

3.2 Robustness through Gaussian Mixtures Models

One of the strengths of the structural model is the simplicity and interpretability of the component models. This is illustrated when we consider non-Gaussian extensions to the trend model. It can readily be seen that outliers in I_t, η_t, and ξ_t translate directly into additive outliers, local level shifts, and ramps respectively. Hence, these types of events can be accommodated for in the model by assuming that I_t, η_t, and ξ_t are generated from an appropriate outlier producing distribution.

Gaussian mixture distributions are one way to model outliers. For example, to generate additive outliers, we assume that

$$I_t \sim \begin{cases} \mathrm{N}(0, \sigma_{\mathrm{I}}^2) & \text{with probability } 1 - \epsilon_{\mathrm{I}} \\ \mathrm{N}(0, \tilde{\sigma}_{\mathrm{I}}^2) & \text{with probability } \epsilon_{\mathrm{I}} \end{cases} \tag{6}$$

where $\tilde{\sigma}_{\mathrm{I}}^2 \gg \sigma_{\mathrm{I}}^2$. The factor ϵ_{I} represents the "prior" probability of an additive outlier.

To avoid too many mixture terms with small probabilities, it can be assumed that only one type of structural change can occur at a given time point. In other words, we shall assume that either a level shift or a ramp can occur, but not both. Hence, the joint distribution of η_t and ξ_t is given by

$$\begin{pmatrix} \eta_t \\ \xi_t \end{pmatrix} \sim \begin{cases} \mathrm{N}\left(\mathbf{0}, \begin{pmatrix} \sigma_\eta^2 & 0 \\ 0 & \sigma_\xi^2 \end{pmatrix}\right) & \text{with probability } 1 - \epsilon_\eta - \epsilon_\xi \\ \mathrm{N}\left(\mathbf{0}, \begin{pmatrix} \tilde{\sigma}_\eta^2 & 0 \\ 0 & \sigma_\xi^2 \end{pmatrix}\right) & \text{with probability } \epsilon_\eta \\ \mathrm{N}\left(\mathbf{0}, \begin{pmatrix} \sigma_\eta^2 & 0 \\ 0 & \tilde{\sigma}_\xi^2 \end{pmatrix}\right) & \text{with probability } \epsilon_\xi \end{cases} \tag{7}$$

where $\tilde{\sigma}_\eta^2 \gg \sigma_\eta^2$ and $\tilde{\sigma}_\xi^2 \gg \sigma_\xi^2$. The factors ϵ_η and ϵ_ξ represent the prior probability of a level shift and ramp.

In the statistics literature, the model defined by (6) and (7) was introduced by Harrison and Stevens (1976), who called it the "Multiprocess Model" (see also Harrison and Stevens (1971)). This model has been successfully used in a Bayesian setting for several applications: see, for example, Smith and West (1983) and Gordon and Smith (1990). Kitagawa (1990) uses a similar model for robust seasonal adjustment, except that σ_η^2 and $\tilde{\sigma}_\eta^2$ are constrained to be zero (so ramps but not level shifts are modeled).

It is interesting to note that from the perspective of generating outlier models, the seasonal models given by (3) and (4) may not be so good. In any case, seasonal breaks are not considered in this study.

3.3 Technical Issues

Evaluation of the Likelihood

In structural time series models, the likelihood function is often decomposed in the form

$$L(Y_1, Y_2, \ldots, Y_N) = p(Y_1)p(Y_2|Y_1)\ldots p(y_N|Y_1, Y_2, \ldots, Y_{N-1}) \tag{8}$$

In the purely Gaussian case, (8) is readily computed by casting the model in state space form and applying the Kalman filter (see, for example, Harvey (1989)). For the Gaussian mixture model of (6) and (7), exact computation of (8) involves an algorithm with complexity of 6^N! This is because the one-step ahead predictive distribution for Y_t is a Gaussian mixture of 6^t components.

Different approaches have been adopted to circumvent this difficulty. Alspach and Sorenson (1972) develop a "Gaussian sums" method in which low probability components of the mixture are pruned. Harrison and Stevens (1976) invoke a collapsing procedure in which a number of terms in the mixture at each time are replaced with a Gaussian distribution. The replacement is done through moment matching, and minimizes the Kulback-Leibler distance. Kitagawa (1990) adopts a collapsing approach similar to Harrison and Stevens. The main difference is that while Harrison and Stevens collapse the same densities at each time, Kitagawa successively collapses the pair of densities which are "closest" in terms of Kulback-Leibler distance. Bruce and Martin (1992) combine a variety of pruning and collapsing methods in an adaptive tree growing algorithm.

Obtaining the seasonal decomposition

To obtain a seasonal decomposition, we need an estimate of the "smoothed" trend and seasonal. Most procedures yield the expected trend and seasonal: $E(T_t|Y_1, Y_2, \ldots, Y_N)$ and $E(S_t|Y_1, Y_2, \ldots, Y_N)$. Using a "Gaussian sum" smoother, Kitagawa (1990) shows how we can actually get an estimate of the densities $p(T_t|Y_1, Y_2, \ldots, Y_N)$ and $p(S_t|Y_1, Y_2, \ldots, Y_N)$. From the densities, we could obtain a point estimate using the expected value. However, in the non-Gaussian setting, a more natural point estimate is given by the *median.*

A particularly nice feature about this approach is that we can readily obtain confidence intervals. In addition, Bruce (1992) shows how to obtain estimates of the posterior probabilities of outliers, level shifts, and ramps.

Initialization of the Filters

To compute the likelihood (8) and to use the two-filter smoother of Kitagawa (1990), we need to handle densities such as $p(Y_t|Y_1, Y_2, \ldots, Y_{t-1})$. For $t < 14$ with monthly data and the models considered above, certain assumptions are needed about initial conditions. In the purely Gaussian case, the most natural approach is to assume a "diffuse prior": see Ansley and Kohn (1985), Bell and Hillmer (1987), and De Jong (1991). In the Gaussian mixture case, a "diffuse prior" could be used as well. The exact distribution, though, is a Gaussian mixture involving an intolerable number of terms for t bigger than 5 or 6. The various schemes for reducing the number of components do not work: the distributions are partially diffuse and it is not possible to determine which observations are likely outliers, level shifts, etc.. As a result, the current implementation of GAUSUM-STM does not properly handle the initialization. This causes problems in the seasonal adjustments with three series. Two possible solutions are available to overcome this problem. These are discussed in section 8.

3.4 The GAUSUM-STM program

A non-Gaussian seasonal adjustment method based on time series structural models, called "GAUSUM-STM", was developed for this study. It is based on a computer program developed by Kitagawa (1991). It permits trend models of the form (2) and a choice of either (3) or (4) for the seasonal model. The irregular is assumed to be white noise. Gaussian mixture distributions of the form (6) and (7) can be specified. The mixture distribution for structural changes, given by (7), can be extended to accommodate seasonal breaks as well. The method of Kitagawa (1990) is used for reducing the number of mixture terms in computing the likelihood and smoothed estimates. GAUSUM-STM is implemented in Fortran-77 as a function in the S language (Becker et al. (1988)).

Ease of Use

Potentially, GAUSUM-STM is easier for the naive seasonal adjuster to apply than X-12-ARIMA. In this study, only the transformation choice was not automatic (but it could easily be done so). By contrast, both an ARIMA model and the seasonal filters had to be specified for X-12-ARIMA.

Offsetting these advantages is the relative computational inefficiency of GAUSUM-STM. In general, GAUSUM-STM is an order of magnitude slower than X-12-ARIMA. However, much can be done to improve the speed of GAUSUM-STM. With the increasing computing power available to a broad spectrum of users, this should not be a major factor in the near future.

Continuity property of GAUSUM-STM

An important property of the GAUSUM-STM method is that it employs a continuous scheme for handling outliers and level shifts. GAUSUM-STM can adapt to different magnitudes of AO's or LS's using appropriate posterior probabilities. Large AO's or LS's are given probabilities close to 1 while small AO's or LS's are given probabilities close to 0. By contrast, the X-12-ARIMA outlier prior adjustment procedure declares observations as either AO's or LS's or neither. Essentially, X-12-ARIMA assigns a posterior probability of either 1 or 0. As a result, we can expect the X-12-ARIMA seasonal adjustment to change discontinuously as an observation passes the threshold and is declared as an AO or LS. While this discontinuity is mitigated by the outlier treatment intrinsic to X-11, we shall see that the GAUSUM-STM outlier method leads to more stable seasonal adjustments.

4 The Data, Filters, and Models

4.1 The Data

The empirical study involves 29 monthly U.S. macroeconomic time series, selected by time series staff at the Statistical Research Division, U.S. Bureau of the Census. These series have both outliers and structural changes (such as level shifts). Table 1 lists the series along with their abbreviations. Of the 29 series, 13 are retail trade series, 7 are housing starts series, and 9 are inventory series. The series exhibit a range of seasonal behavior. The retail trade series usually display very strong seasonal patterns. The construction series are often very erratic and quite difficult to adjust. The inventory series tend to have large level shifts but relatively stable adjustments. On the whole, this collection of economic time series gives a broad range of problems with which to assess and compare seasonal adjustment methods.

4.2 Log Additive Seasonal Decomposition

Time series staff at the Statistical Research Division identified a multiplicative decomposition in all of the series for the X-12-ARIMA seasonal adjustments. In th study, we transform the data by taking logarithms and multiplying by 1000

	Abbreviation	Series Description	SABL transform
B1	BAUTRS	Retail Sales of Automobiles	0
B2	BFRNRS	Retail Sales of Furniture	0
B3	BGMRRI	Retail Sales of General Merchandise	-0.5
B4	BGRCRS	Retail Sales of Groceries	0
B5	BHDWWS	Wholesale Sales of Hardware	0
B6	BLQRRS	Retail Sales of Liquor	0
B7	BMNCRS	Retail Sales of Men's Apparel	-0.25
B8	BSHORS	Retail Sales of Shoes	0
B9	BSPGWS	Wholesale Sales of Sporting Goods	0.25
B10	BTAPRI	Total Retail Sales of Apparel	-0.25
B11	BTNDRI	Retail Sales of Nondurables	-0.25
B12	BVARRS	Variety Store Retail Sales	0.25
B13	BWAPRS	Retail Sales of Women's Apparel	0
C14	CMW1HS	One Family Housing Starts in the Midwest	-0.25
C15	CMWTHS	Total Housing Starts in the Midwest	0
C16	CNE1HS	One Family Housing Starts in the Northeast	0.25
C17	CNETHS	Total Housing Starts in the Northeast	0.25
C18	CSOTHS	Total Housing Starts in the South	0
C19	CWETHS	Total Housing Starts in the West	0.25
C20	C24THS	Total Housing Starts - 2 to 4	0.25
I21	IBEVTI	Total Inventories of Beverages	-1
I22	ICMETI	Total Inventories of Communications Equipment	-1
I23	IFATTI	Total Inventories of Fats and Oils	-1
I24	IFMETI	Total Inventories of Farm Machinery and Equipment	0
I25	IGLCTI	Total Inventories of Glass Containers	-0.25
I26	IHAPTI	Total Inventories of Household Appliances	-0.25
I27	INEWUO	Unfilled Orders for Newspapers and Magazines	0.25
I28	ITVRTI	Total Inventories of TV's and Radios	0.25
I29	ITVRUO	Unfilled Orders for TV's and Radios	0.5

Table 1: List of abbreviations for the 29 series studied and power transformations used by the SABL seasonal adjustment procedure. The choice of powers is roughly consistent with a logarithmic transform (power = 0).

then apply an additive seasonal adjustment. The log-additive seasonal adjustment is used so that GAUSUM-STM can be fairly compared with X-12-ARIMA. While multiplicative and log-additive seasonal adjustments should be similar, there is a consistent downward bias in the trend from the log-additive procedure (see Ozaki and Thomson (1992)).

For simplicity and conceptual clarity, we have chosen not to consider other possible transformations (e.g., square root). Table 1 displays the transformation powers for each of the series chosen by the robust seasonal adjustment procedure SABL (Cleveland and Devlin (1980)). The powers range from -1 to 0.5, with the majority being ± 0.25. This is reasonably consistent with a logarithmic transformation.

4.3 X-12-ARIMA Filters

For the 29 series in this study, the options required by X-12-ARIMA were provided by the time series staff of the U.S. Bureau of the Census. This includes the choice of ARIMA models, filters, and trading day and Easter effects. For the retail trade and inventory series, the default filters are used (see Dagum (1980)). For the construction series, a 3×9 moving average is used, yielding smoother seasonal factors than the default filter. Trading day prior adjustment is done for all of the retail trade series and three of the construction series. Prior adjustment for the timing of Easter is done for five of the retail trade series.

4.4 GAUSUM-STM Models

Structural Models

Three types of seasonal models are fit to all 29 series: the "BSM" seasonal (3), the trigonometric seasonal (4) with the assumption that all of the noise terms $\omega_{j,t}$ have a common variance σ_ω^2, and the trigonometric seasonal (4) allowing different variances $\sigma_{\omega,j}^2$. These models will be denoted by BSM, TRIG-1, and TRIG-6.

The parameters for the structural models include:

- The variances of the mean and slope of the trend component σ_η^2 and σ_ξ^2.
- The variances of the seasonal component σ_ω^2 (or $\sigma_{\omega,j}^2$ in the case of TRIG-6).
- The variance of the irregular component $\sigma_{\rm I}^2$.
- The prior probabilities of an additive outlier and a level shift $\epsilon_{\rm I}$ and ϵ_η.

The variances of the outlier and level shift processes, $\tilde{\sigma}_{\rm I}^2$ and $\tilde{\sigma}_\eta^2$, are set to a large fixed value.

In addition, a constrained version of the BSM is fit, and will be denoted by BSM-CONS. Let y_t denote a time series for which the default filters of X-11 are "optimal". Then BSM-CONS corresponds to the basic structural model with an *acf* which closely matches the *acf* of $\Delta\Delta^{12}y_t$. This model optimizes over a single variance σ^2, which is related to the other variances as follows (Maravall (1985), Harvey and Valls Pereira (1989)):

$$\sigma_\epsilon^2 = \sigma^2 \quad \sigma_\eta^2 = 0.133\sigma^2 \quad \sigma_\xi^2 = 0.167\sigma^2 \quad \sigma_\omega^2 = 0.067\sigma^2 \tag{9}$$

This model is discussed further in section 7.5.

Outlier Model

A simplified version of the level shift and ramp model (7) is used in this study. The prior probability of a ramp ϵ_ξ is set to zero, reducing the Gaussian mixture in (7) to just two terms. This restricted model provides for most of what is desired in terms of modeling structural changes while significantly reducing the computational burden. Section 7.3 explores the more general ramp model.

4.5 Model Parameters

Table 2 gives the maximum likelihood estimates of the parameters for TRIG-6 for each of the series. See appendix A for details on the fitting procedure. The table displays the variances σ_η^2, σ_ξ^2, σ_I^2, the mean of the variances of the seasonal components $\bar{\sigma}_\omega^2 = \sum_{j=1}^6 \sigma_{\omega,j}^2/6$. and the prior probabilities ϵ_η and ϵ_I.

The type of convergence achieved by the optimizer is also given in Table 2. Five types of convergence are possible: (R)elative function convergence, (X)-convergence, (B)oth X- and relative function convergence, and (F)alse convergence (see appendix A for details). These are denoted in the table by the letters in the parentheses. A convergence code of R, X, or B indicates that the optimizer successfully found a local maximum. A convergence code of F means that the optimizer may be stuck at a non-critical value.

Note that a few series have false convergence for TRIG-6. By contrast, almost all of the fits for the BSM and TRIG-1 acheived successful convergence (Bruce and Jurke (1992)). This is probably due to the number of seasonal parameters and the relative flatness of the likelihood. Since the TRIG-6 uses the maximum likelihood estimates of TRIG-1 as starting values, we expect that reasonably good estimates are obtained in all cases. Furthermore, our experience with the optimizer indicates that many of the false convergences are at a local maximum.

Comparison of Seasonal Parameters

Table 3 compares the variances of the seasonal components for TRIG-1 and TRIG-6. The TRIG-6 variances are given relative to the TRIG-1 variance. Let $\sigma_{\omega,j}^2$ for

	Variances				Probabilities		Type of convergence
	mean σ^2_η	slope σ^2_ξ	irregular σ^2_ϵ	seasonal $\bar{\sigma}^2_\omega$	outlier ϵ_I	level shift ϵ_η	
BAUTRS	1488	1.228e-08	0.004631	1.657	0.008848	0.0001354	X
BFRNRS	285.9	9.559e-08	67.15	0.3738	5.943e-05	0.0005969	R
BGMRRI	81.22	1.514	0.005892	0.4867	0.0003415	0.0284	R
BGRCRS	4.624	0.5419	100	9.696e-14	9.294e-05	0.0004777	R
BHDWWS	471.2	7.394e-05	354.3	0.6132	5.643e-05	0.0001864	R
BLQRRS	137.4	0.05058	121.7	0.2775	0.006969	0.004569	F
BMNCRS	123.3	0.09107	335.6	4.963	8.679e-05	8.67e-05	R
BSHORS	408	0.001657	317.2	1.616	7.074e-05	0.008381	F
BSPGWS	1934	9.571e-05	843.3	4.77	0.006177	0.0002729	B
BTAPRI	132.6	2.174e-05	6.783e-06	0.1578	6.098e-05	0.002837	R
BTNDRI	41.84	0.002471	0.05045	0.2091	0.0001406	0.01906	F
BVARRS	134.1	0.0261	342.5	0.8838	0.000809	0.01346	F
BWAPRS	189.1	0.0149	217.6	2.153	0.000111	0.0001306	B
C24THS	8946	8.627e-05	11830	3.336	0.0007437	0.0006874	R
CMW1HS	5728	0.0008224	9063	4.575	0.028	0.0095	R
CMWTHS	11940	0.0001437	9377	9.969	0.008784	0.00821	
CNE1HS	4596	5.12e-05	11230	34.2	0.00503	0.000751	R
CNETHS	7683	0.001412	18370	17.77	0.01361	0.001678	X
CSOTHS	5014	0.001275	3083	0.4219	0.0004325	0.001477	R
CUSTHS	4276	3.132e-06	816.3	2.903	4.054e-05	4.344e-05	F
CWETHS	8078	0.000324	3979	2.895	0.0005571	0.0005949	R
IBEVTI	152.6	0.2058	5.47e-08	0.521	9.929e-05	0.001752	R
ICMETI	20.16	17.83	23.91	0.05469	1.052e-08	0.006383	R
IFATTI	4044	0.002189	2.959e-05	4.974	5.976e-05	0.002057	B
IFMETI	284.6	6.747	26.04	1.503e-09	0.0009216	0.01885	R
IGLCTI	517.3	0.1818	0.001409	0.1159	0.0001653	0.01075	R
IHAPTI	443.3	0.02705	1.685e-05	0.6787	0.0001042	0.005982	X
INEWUO	2221	0.9796	0.000118	4.472	7.64e-05	0.002101	R
ITVRTI	751.4	18.72	0.006499	0.5567	0.0008571	1.561e-05	R
ITVRUO	6574	0.01101	906.3	85.11	0.02277	0.003414	F

Table 2: Structural model parameters as estimated by GAUSUM-STM for the TRIG-6 model. Note that $\bar{\sigma}^2_\omega = \sum_{j=1}^{6} \sigma^2_{\omega,j}/6$. The individual seasonal variances are given in Table 3.

	TRIG-1	TRIG-6 variances						AIC		
	σ^2_ω	$\tilde{\sigma}^2_{\omega,1}$	$\tilde{\sigma}^2_{\omega,2}$	$\tilde{\sigma}^2_{\omega,3}$	$\tilde{\sigma}^2_{\omega,4}$	$\tilde{\sigma}^2_{\omega,5}$	$\tilde{\sigma}^2_{\omega,6}$	BSM	TRIG-1	TRIG-6
BAUTRS	0.84	0.0020	8.84	2.44	0.54	0.0025	0.0027	2748	2742	2736
BFRNRS	0.40	0.0011	0.86	2.37	1.51	0.48	0.37	2351	2350	2355
BGMRRI	0.19	12.09	2.90	0.88	0.041	0.43	0.015	1764	1751	1729
BGRCRS	7.19e-15	27.87	0.37	0.069	17.69	0.046	34.9	2049	2049	2059
BHDWWS	0.31	1.14	1.54	0.045	8.03	0.76	0.48	2566	2565	2572
BLQRRS	0.068	18.58	0.066	4.59	0.36	0.12	0.95	2290	2289	2296
BMNCRS	2.71	5.96	3.00	0.94	0.85	0.15	0.089	2577	2559	2544
BSHORS	1.61	2.93	0.0037	0.26	2.37	0.39	0.071	2607	2604	2603
BSPGWS	2.59	6.16	0.64	2.74	0.75	0.26	0.51	2888	2884	2888
BTAPRI	0.059	11.7	1.25	2.01	1.21	3.12e-05	7.92e-06	1667	1664	1663
BTNDRI	0.073	13.69	2.08	1.39	0.00041	0.025	0.0029	1519	1500	1484
BVARRS	0.76	3.81	1.65	3.27e-06	0.97	0.53	0.016	2506	2493	2489
BWAPRS	0.99	1.10	5.77	4.94	1.07	0.094	0.081	2441	2435	2434
C24THS	3.33	0.010	0.23	0.10	1.29	0.39	3.99	3855	3853	3860
CMW1HS	1.78	11.18	7.55e-07	3.98	3.12e-05	3.11e-05	0.2555	3880	3879	3883
CMWTHS	0.0067	8352	535.7	82.7	0.80	0.25	0.19	3858	3889	3865
CNE1HS	34.01	2.41	2.51	0.56	0.034	0.48	0.043	3877	3866	3868
CNETHS	5.54	10.86	5.81	1.22	0.0048	1.34	0.0036	3970	3968	3974
CSOTHS	3.25e-08	9437	42650	18040000	59730000	141800	250.6	3564	3564	3570
CWETHS	0.86	15.87	0.00042	0.037	0.00092	2.48	1.87	3686	3684	3691
IBEVTI	0.072	33.44	6.86	2.73	0.31	0.28	6.12e-05	2460	2455	2435
ICMETI	0.055	0.00036	3.28	1.07	1.34	0.26	0.016	1959	1958	1963
IFATTI	1.22	15.94	3.64	4.47	2.32e-07	0.34	0.052	3327	3323	3315
IFMETI	4.66e-10	1.91	10.19	0.0087	1.49	2.34	3.43	2879	2879	2889
IGLCTI	0.040	1.93	0.83	0.22	13.84	0.30	0.36	3003	3003	3011
IHAPTI	0.019	207.6	6.32	0.00031	0.00018	1.78	1.19	2913	2913	2911
INEWUO	2.0	7.09	2.76	2.95	0.34	0.225	0.071	3186	3177	3169
ITVRTI	0.25	5.97	6.52	0.44	0.24	0.024	1.84e-05	2895	2882	2877
ITVRUO	49.98	1.22	6.24	0.50	0.91	0.066	1.28	3818	3798	3782

Table 3: Seasonal parameters as estimated by GAUSUM-STM for the TRIG-1 and TRIG-6 models. The TRIG-6 variances are given as ratios to the TRIG-1 varaince: $\tilde{\sigma}^2_{\omega,j} \equiv \sigma^2_{\omega,j}/\sigma^2_\omega$. The AIC values are given for the BSM, TRIG-1, and TRIG-6.

$j = 1, \ldots, 6$ be the variances for the TRIG-6 model and let σ_ω^2 be the variance for the TRIG-1 model. Table 3 gives σ_ω^2 and $\tilde{\sigma}_{\omega,j}^2 \equiv \sigma_{\omega,j}^2/\sigma_\omega^2$.

Table 3 also gives Akaike's Information Criterion (AIC) for the BSM, TRIG-1, and TRIG-6 models. AIC is defined by

$$\mathrm{AIC} = -2 \times \log \mathrm{L} + 2p$$

where log L is the log-likelihood at the maximum and p are the number of parameters fit. AIC gives a guide towards selecting the "best" model, and models with lower AIC values are preferable.

5 Assessing and Comparing Seasonal Adjustments

Seasonal adjustments are assessed using a variety of criteria: see, for example, Dagum (1978), Granger (1978), Cleveland and Terpenning (1982), den Butter and Mourik (1990), and Findley et al. (1990). This empirical comparison is done mainly on the basis of a set of 9 diagnostic plots produced for each series. Naturally, only a small subset of these plots are given here. The complete "book" of plots is in Bruce and Jurke (1992). A brief description of the plots and statistics and what they are trying to assess is given below. Refer to Appendix B for detailed descriptions of each plot.

In addition to the plots, seven statistics were computed to measure various aspects of the seasonal adjustment: roughness of adjusted data (ADJ ROUGH), trend roughness (TREND ROUGH), orthogonality (ORTHOG), seasonal magnitude (SEAS MAG), seasonal flexibility (SEAS FLEX), seasonal roughness (SEAS ROUGH), and remaining seasonality (SEAS SIGNIF). These statistics are discussed in more detail below. Table 4(a) and 4(b) give the diagnostic statistics for X-12-ARIMA and TRIG-6 for each of the series.

Figures 1(a)-(f) display these statistics with boxplots for BSM, BSM-CONS, TRIG-1, TRIG-6, and X-12-ARIMA. The boxplots are broken down by series type (retail sales, construction starts, or inventories). The statistics ADJ ROUGH, TREND ROUGH, SEAS MAG, SEAS FLEX, and SEAS ROUGH are all "median corrected" within each plot (i.e., the median of all observations within a plot are subtracted).

5.1 Criteria for Assessment

Nature of seasonal factors

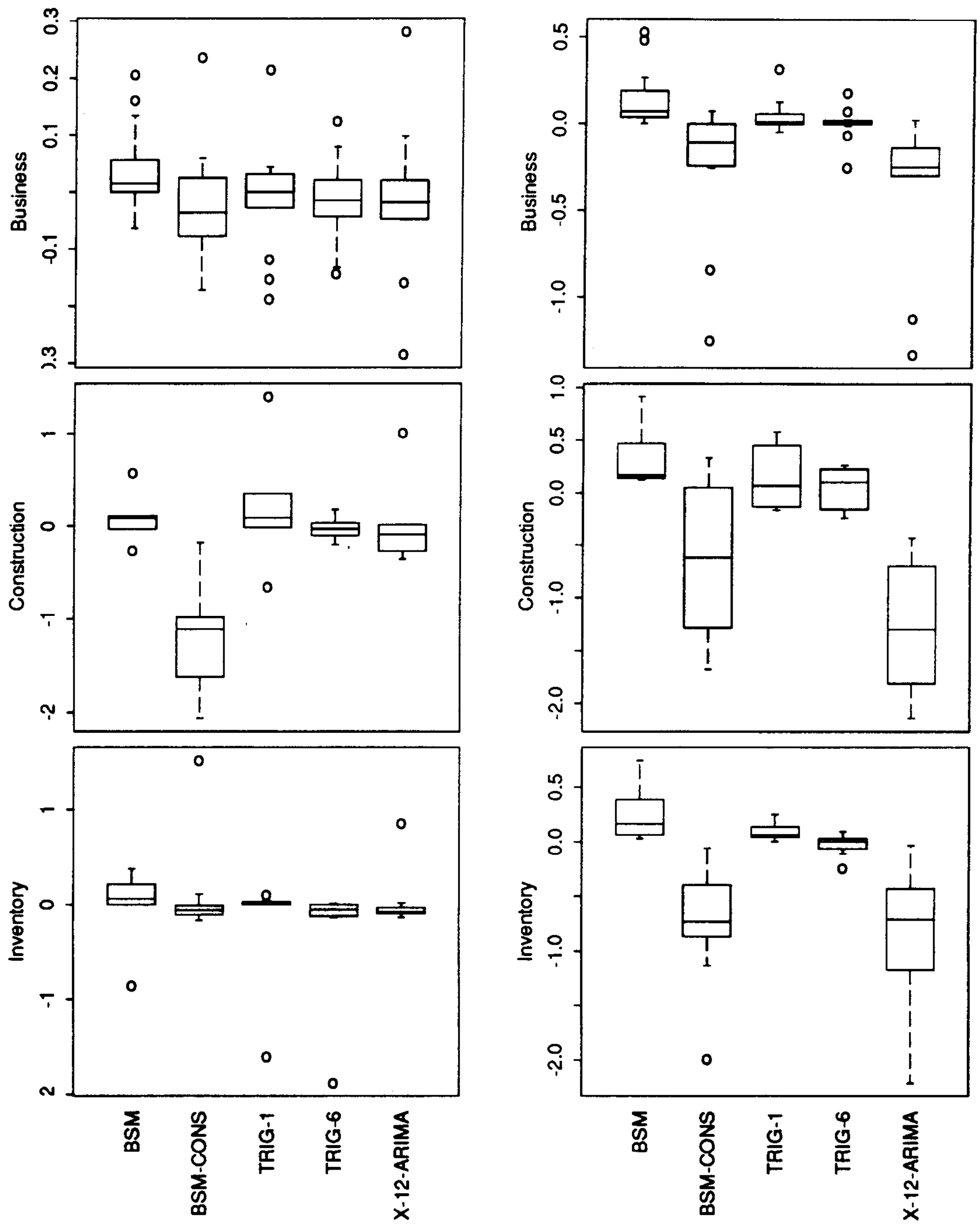
In this study, we focus on the "flexibility" and "smoothness" of the seasonal factors. Flexibility measures the amount that the seasonal effect for a given month is allowed

Series	ADJ ROUGH TRIG-6	ADJ ROUGH X-12	ORTHOG TRIG-6	ORTHOG X-12	TREND ROUGH TRIG-6	TREND ROUGH X-12
BAUTRS	3.21	3.21	-1.24	-1.51	2.68	1.17
BFRNRS	1.51	1.54	1.93	1.81	1.16	0.83
BGMRRI	1.08	1.17	3.46	3.99	1.08	0.84
BGRCRS	1.25	1.23	0.63	0.47	0.61	0.63
BHDWWS	2.56	2.20	0.01	-0.35	1.31	1.02
BLQRRS	1.46	1.44	2.06	2.15	0.69	0.55
BMNCRS	1.89	2.31	2.53	1.63	0.52	0.63
BSHORS	2.48	2.59	1.39	1.60	1.23	0.93
BSPGWS	4.36	4.48	3.12	2.58	2.39	1.27
BTAPRI	1.01	1.03	3.47	3.61	1.01	0.72
BTNDRI	1.01	0.83	2.76	3.87	1.01	0.81
BVARRS	2.37	2.21	2.06	2.05	0.69	0.69
BWAPRS	1.94	1.91	0.72	2.08	0.83	0.69
C24THS	13.60	13.40	1.56	-2.43	4.65	3.24
CMW1HS	13.80	13.70	6.00	6.54	4.37	3.52
CMWTHS	13.20	12.80	4.19	6.84	5.82	3.64
CNE1HS	12.40	13.30	1.12	1.28	2.71	2.52
CNETHS	16.70	16.80	3.58	6.82	3.41	2.05
CSOTHS	8.07	7.91	0.82	0.68	3.92	2.40
CWETHS	9.89	9.94	0.97	-0.14	5.28	3.20
IBEVTI	0.98	1.03	0.68	0.45	0.98	0.66
ICMETI	1.09	1.10	-1.57	-1.78	0.91	0.90
IFATTI	4.19	4.22	1.59	1.11	4.19	2.10
IFMETI	1.74	1.60	-1.37	-0.70	1.57	1.08
IGLCTI	1.93	1.82	-0.49	-0.81	1.94	1.20
IHAPTI	1.61	1.59	-1.48	-1.73	1.61	1.00
INEWUO	3.14	3.25	1.22	0.40	3.14	1.87
ITVRTI	2.32	2.23	-0.39	-1.15	2.23	1.45
ITVRUO	7.01	9.74	7.89	7.12	4.27	2.46

Table 4: ADJ ROUGH, given by (13), measures the roughness of the seasonally adjusted data. TREND ROUGH, given by (14), measures the rougness of the trend. ORTHOG is 100 times the correlation between the seasonally adjusted data and the seasonal component.

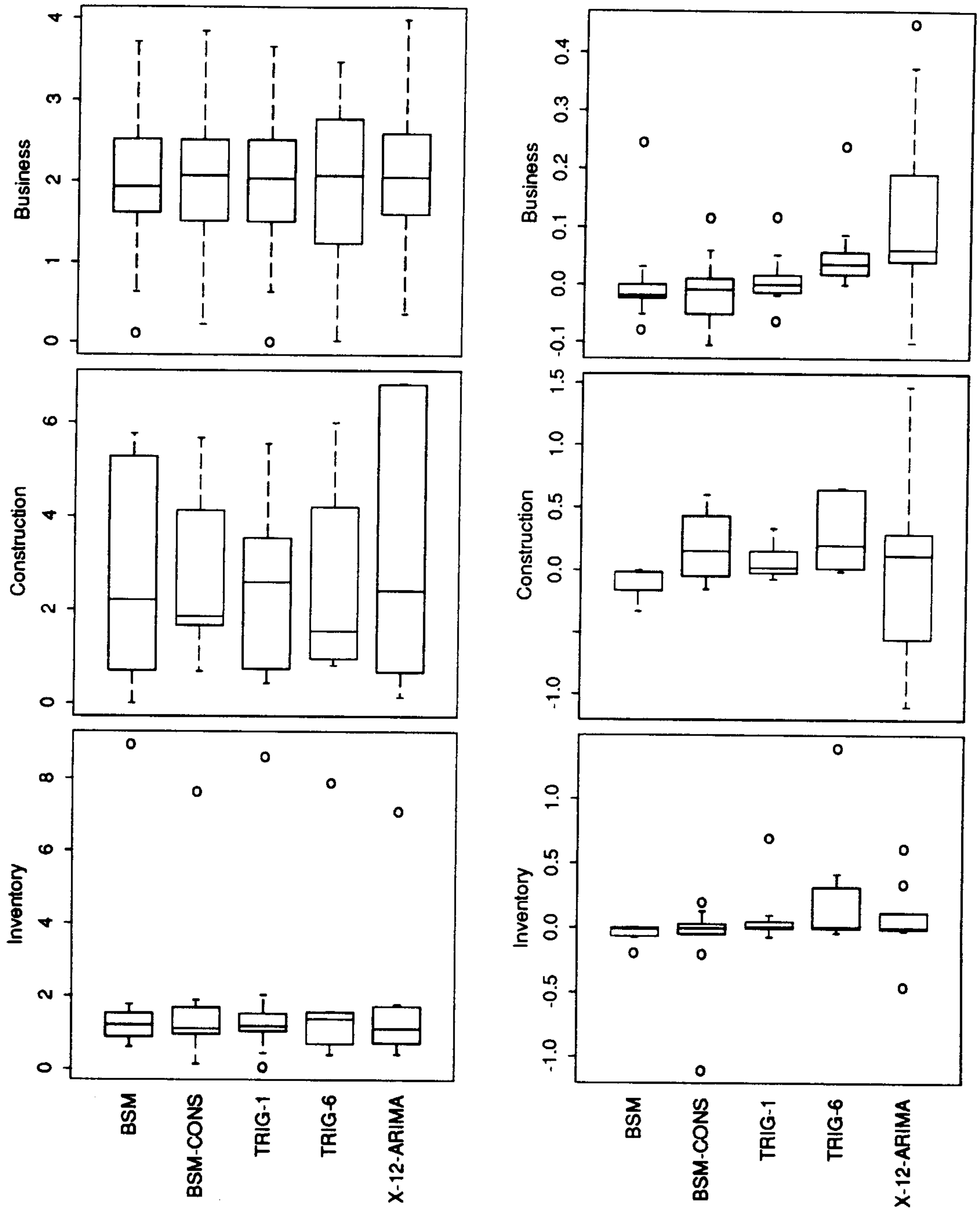
Series	SEAS FLEX TRIG-6	SEAS FLEX X-12	SEAS ROUGH TRIG-6	SEAS ROUGH X-12	SEAS SIGNIF TRIG-6	SEAS SIGNIF X-12	SEAS MAG TRIG-6	SEAS MAG X-12
BAUTRS	0.34	0.44	0.11	0.07	0.90	0.98	9.53	9.60
BFRNRS	0.15	0.22	0.04	0.03	0.98	1.00	6.07	6.18
BGMRRI	0.22	0.16	0.09	0.02	0.96	0.98	6.43	6.28
BGRCRS	0.00	0.08	0.00	0.02	0.98	0.96	3.37	3.37
BHDWWS	0.18	0.24	0.06	0.04	1.00	0.98	6.19	6.62
BLQRRS	0.10	0.16	0.03	0.03	0.97	0.98	8.93	8.97
BMNCRS	0.71	0.45	0.28	0.06	1.00	0.90	17.50	17.40
BSHORS	0.34	0.36	0.08	0.05	1.00	1.00	12.60	12.60
BSPGWS	0.52	0.67	0.12	0.10	1.00	1.00	9.33	9.65
BTAPRI	0.10	0.13	0.03	0.02	1.00	0.99	5.47	5.43
BTNDRI	0.18	0.10	0.08	0.01	0.00	0.91	3.63	3.84
BVARRS	0.23	0.23	0.05	0.03	0.99	1.00	16.10	16.10
BWAPRS	0.35	0.29	0.10	0.04	0.99	0.96	13.60	13.60
C24THS	0.41	0.76	0.08	0.11	1.00	0.98	16.30	17.80
CMW1HS	0.46	0.69	0.08	0.11	0.99	0.96	44.10	44.10
CMWTHS	0.71	0.74	0.11	0.10	1.00	0.99	40.80	39.60
CNE1HS	1.57	0.77	0.41	0.11	1.00	0.94	43.40	43.10
CNETHS	0.87	0.76	0.18	0.16	0.99	0.35	36.50	34.90
CSOTHS	0.10	0.31	0.02	0.06	1.00	0.96	14.80	14.90
CWETHS	0.28	0.48	0.04	0.09	1.00	0.97	15.80	15.80
IBEVTI	0.21	0.17	0.05	0.02	1.00	1.00	1.65	1.63
ICMETI	0.06	0.10	0.01	0.01	1.00	1.00	0.68	0.79
IFATTI	0.50	0.59	0.10	0.07	0.58	0.93	10.40	9.97
IFMETI	0.00	0.20	0.00	0.03	1.00	0.76	3.55	3.52
IGLCTI	0.06	0.22	0.01	0.03	1.00	0.99	1.71	2.09
IHAPTI	0.16	0.24	0.04	0.03	1.00	1.00	4.28	4.31
INEWUO	0.48	0.69	0.11	0.08	0.97	0.98	5.90	6.20
ITVRTI	0.16	0.30	0.03	0.04	1.00	0.98	4.89	4.89
ITVRUO	2.60	1.54	1.21	0.22	0.84	0.28	14.10	12.30

Table 4: SEAS FLEX, SEAS ROUGH and SEAS MAG measure the flexibility, roughness, and magnitude of the seasonal component as defined by (10), (12), and (11) respectively. SEAS SIGNIF represents the significance of seasonality in the seasonally adjusted data.



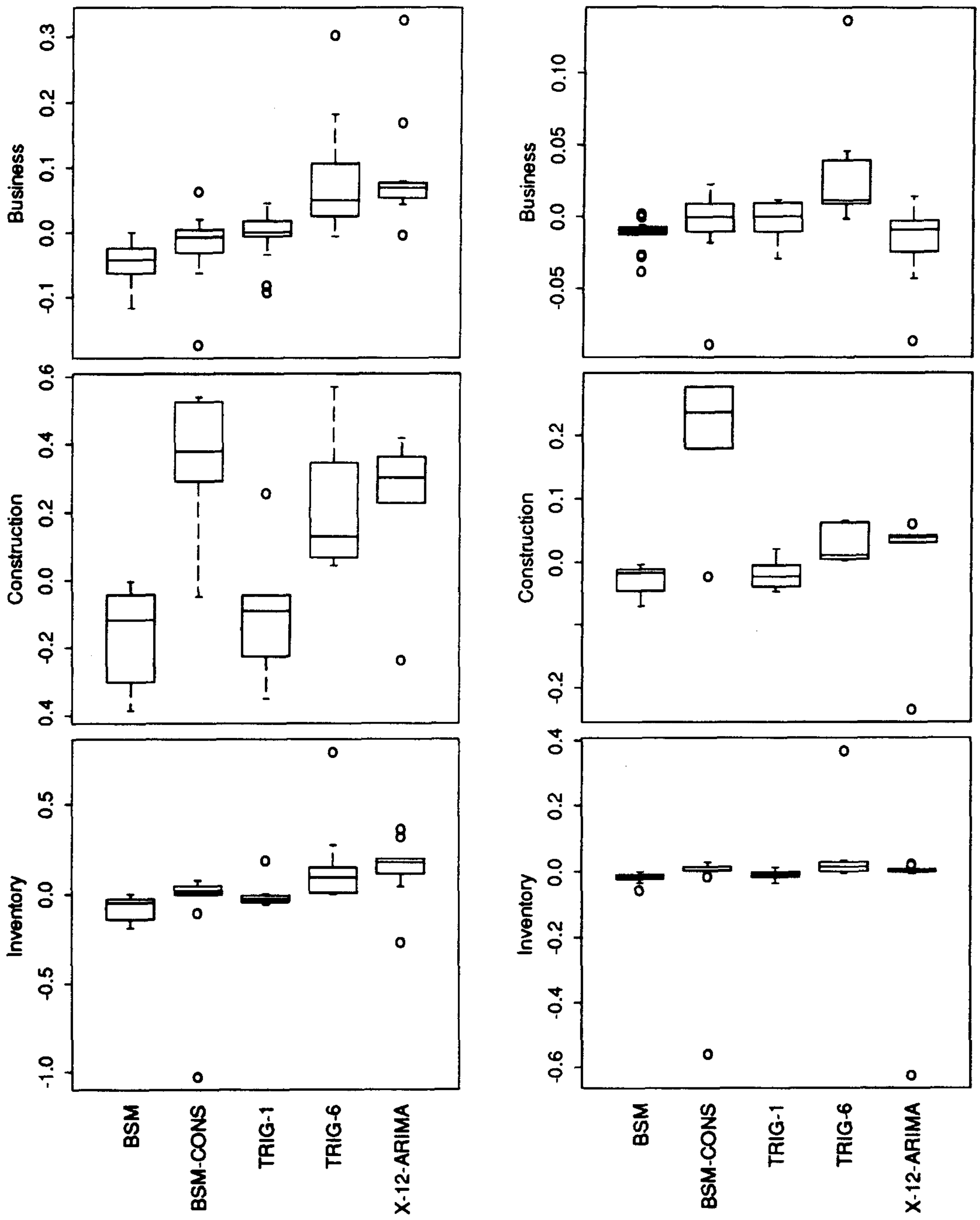
(a) (b)

Figure 1: (a) The diagnostic ADJ ROUGH, given by (13), and (b) the diagnostic TREND ROUGH, given by (14). The boxplots are "median" corrected within each plot (i.e., the median of all observations within a plot are subtracted).



(c) (d)

Figure 1: (c) The absolute value of the diagnostic ORTHOG, given by (15), and (d) the diagnostic SEAS MAG, given by (12). The boxplots for SEAS MAG are "median" corrected within each plot (i.e., the median of all observations within a plot are subtracted).



(e) (f)

Figure 1: (e) The diagnoistic SEAS FLEX, given by (10) and (f) the diagnostics SEAS ROUGH, given by (11). The boxplots are "median" corrected within each plot (i.e., the median of all observations within a plot are subtracted).

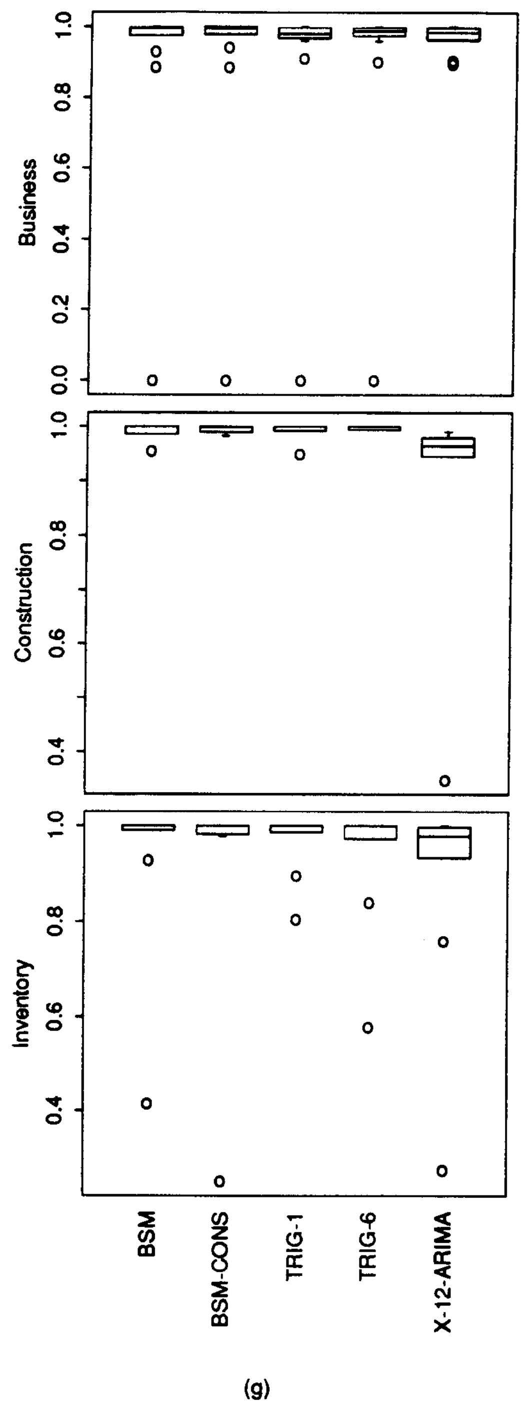


Figure 1: (g) The diagnostic SEAS SIGNIF, the significance of any remaining seasonality.

to "bend" or change from year to year. A certain amount of flexibility is desirable to allow the factors to adapt to the data.

Seasonal factors of a given flexibility can either be slowly varying or rapidly changing. This corresponds to a monthly effect which evolves smoothly or roughly from year to year. All things being equal, smoother seasonal factors are preferable. Hannan (1964) states

> ...there seems little point in allowing for anything more than the very slowest change in seasonal variation. It would seem wrong here to concern oneself too much with faithfully representing a possibly rapidly changing seasonal because of the consequent risk of seriously distorting the series.

A very effective diagnostic for displaying the nature of the seasonal factors is the "SSI-plot" (Cleveland and Terpenning (1982)). Figure 11 gives an example of an SSI-plot. The detrended transformed data is displayed for each month with the seasonal factors superimposed as a line. It is also useful to plot the evolution of the seasonal factors from year to year, as in Figure 10. This plot is an adaptation of similar plot in Cleveland and Terpenning (1982). The top plot shows the mean seasonal factor for each method and the subsequent plots show the deviations from the mean. See Appendix B for a complete description of these plots.

We measure seasonal flexibility by the year to year change in the smoothed seasonal factors. Denote the seasonal factors of the log-transformed data by S_t. Let $\tilde{S}_t$ be the result of smoothing S_t using a linear filter of length three with weights .25, .5 and .25. The diagnostic is given by

$$\text{SEAS FLEX} \equiv \frac{100}{N} \sum_{t=t_0}^{t_N} |\Delta^{12} \tilde{S}_t|. \tag{10}$$

SEAS FLEX corresponds roughly to the mean annual percentage change in the smoothed seasonal:

$$\frac{\exp \tilde{S}_t - \exp \tilde{S}_{t-12}}{\exp \tilde{S}_{t-12}} \approx \Delta^{12} \tilde{S}_t.$$

Seasonal roughness is measured by simply looking at the mean absolute residuals from the smooth:

$$\text{SEAS ROUGH} \equiv \frac{100}{N} \sum_{t=t_0}^{t_N} |S_t - \tilde{S}_t|. \tag{11}$$

Another useful concept is the overall magnitude of the seasonal effect. We measure this by

$$\text{SEAS MAG} \equiv \frac{100}{N} \sum_{t=t_0}^{t_N} |S_t|. \tag{12}$$

Nature of seasonal adjustments

A simple but important plot is to look at the decomposition produced by the adjustment procedure. An example of this plot is given by figure 8. While this plot is good at describing the overall nature of the decomposition, it is of little help for comparison of the adjustments. Figure 22 gives an example of a more direct comparison of seasonal adjustment methods. The top plot compares the seasonally adjust series given by TRIG-6 and X-12-ARIMA. The middle plot shows the ratio of these adjusted series, and the final plot compares the trends. See Appendix B for a complete description of these plots.

All things being equal, smoother seasonal adjustments are seen as desirable (den Butter and Mourik (1990)). Let A_t be the log transformed seasonally adjusted data. Roughness of the seasonally adjustments is measured by

$$\text{ADJ ROUGH} \equiv \frac{100}{N} \sum_{t=t_0}^{t_N} |\Delta A_t|. \tag{13}$$

This corresponds approximately to the mean percentage change in the untransformed seasonally adjusted data:

$$\frac{\exp A_t - \exp A_{t-1}}{\exp A_{t-1}} \approx \Delta A_t$$

Very smooth series have roughness close to zero.

A diagnostic for the roughness of the trend is similarly defined:

$$\text{TREND ROUGH} \equiv \frac{100}{N} \sum_{t=t_0}^{t_N} |\Delta T_t|. \tag{14}$$

where T_t is the trend of the log transformed data. While smoother trends are often visually more appealing, there is no consensus that a smoother trend is necessarily better.

Another desirable feature of a seasonal decomposition is for the seasonally adjusted data to be orthogonal to the seasonal factors. To measure this, we define

$$\text{ORTHOG} \equiv 100 \times \text{corr}(S_t, A_t) \tag{15}$$

where "corr" is the correlation between two variables.

Comparison of outlier treatments

It turns out that there is relatively little difference in the effect of outlier treatments on the adjusted data: see section 6.1. Hence, outlier treatments are compared mainly by examining their effect on the trend, as in figure 21. The top plot compares the trends with the location of outliers and level shifts as identified by X-12-ARIMA marked. The middle plot gives the ratio of the trends, and the final plot gives the posterior probabilities of outliers and level shifts as identifed by TRIG-6. See Appendix B for a complete description of these plots.

The stability of the outlier treatments is also considered in the sliding spans plots (see below).

Remaining seasonality in the seasonally adjusted data

There should not be any seasonality remaining the in the seasonally adjusted data. In a stable seasonal pattern, this is easily checked using a one-way ANOVA significance test for the presence of seasonality. The test is not very effective for testing the residual seasonality in a changing seasonal pattern. For this, we need to look at the periodogram of the detrended seasonally adjusted data, as in figure 12. It is probably desirable for a seasonal adjustment procedure to eliminate power in the periodogram at and near the fundamental frequency $\pi/6$ and perhaps its first harmonic (Hannan (1964)).

Let L_t be the prior adjustment for level shifts in the log scale as identified by X-12-ARIMA. Let $\tilde{T}_t$ be a smoothed version of the log transformed trend obtained by taking 23 point triangular moving average smooth of $T_t - L_t$. The detrended seasonally adjusted data is given by $\tilde{A}_t = A_t - (L_t + \tilde{T}_t)$, where A_t are the original log transformed seasonal adjustments. The diagnostic "SEAS SIGNIF" corresponds to the p-value of the one-way ANOVA test for the presence of a fixed seasonal effect in the series $\tilde{A}_t$. The one-way ANOVA test is computed by fitting the least squares regression. The regression matrix is composed of 12 dummy variables, each representing the effect of the jth month.

Other authors have used a test for "idempotency", which is defined by measuring the size of the seasonal factor obtained by reapplying the seasonal adjustment procedure to the seasonally adjusted data. A procedure is fully idempotent if no additional seasonality is found. Idempotency is distinct from detecting residual seasonality: if a procedure inadequately removed seasonality initially, it may still fail to do so in the second application. We have not looked at idempotency, but expect to achieve results similar to those obtained by den Butter and Mourik (1990): in contrast to X-12-ARIMA, the structural model procedures are nearly fully idempotent. It is natural that X-12-ARIMA, which is a semi-parametric procedure that yields more flexible seasonal factors, uncovers more seasonality in the second application.

Stability of the seasonal adjustments procedure:

Another crucial aspect of seasonal adjustment is the stability of the procedure. Seasonal adjustment is often done on official statistics, and it is especially important that the procedure does not contain any inherent volatility. The sliding spans of

Findley et al. (1990) are one measure of volatility involving adjustment of the data using four overlapping spans of approximately eight years in length. This plot shows the volatility of the month-to-month percentage change in the seasonal adjustments. For time t and span k, these are defined as

$$MM_t(k) = \frac{\exp A_t(k) - \exp A_{t-1}(k)}{\exp A_t(k)}$$

where $A_t(k)$ is the log transformed seasonally adjusted value at time t in span k. Using the notation of Findley et al. (1990), the value plotted for a given time t is

$$MM_t^{\max} = \max_k MM_t(k) - \min_k MM_t(k) \tag{16}$$

where k varies over those spans that contain both months t and $t-1$. Seasonal adjustments with more than 25% of the months with $MM_t^{\max} > 0.03$ are almost never acceptable.

- The sliding spans concept is used to construct a "Sliding Spans Plot", such as figures 23 and 24. The top plot shows the seasonally adjusted data for the four spans The second and fourth plots display the statistic $MM_t^{\max}$. The third plot compares the outlier treatments over the spans.

6 Main Results

6.1 Comparison of Seasonal Factors

Smoothness and Flexibility of X-12-ARIMA

X-12-ARIMA and structural models yield significantly different seasonal decompositions and seasonal adjustments. This is fundamental to the nature of the methods, and has nothing to do with the difference in outlier treatments. For the business and inventory series, X-12-ARIMA produces smoother seasonal factors than TRIG-6 as defined by the year to year change: see figure 1(f). At the same time, the X-12-ARIMA seasonal is as or more flexible than any of the structural models: see figure 1(e). Correspondingly, X-12-ARIMA often removes significantly more power in the periodogram around the fundamental frequency $\pi/6$ (this is reflected in the periodogram plots).

The X-12-ARIMA and TRIG-6 seasonal factors clearly differ in 15 of the 29 of the series (B1, B2, B3, B5, B6, B7, B10, B13, C20, I21, I23, I24, I25, I27, and I28). The distinction is still noticeable, though less significant, in 10 others (B8, B9, B11, B12, C14, C15, C20, I22, I26, and I29). In general, in series with little structure or linear changes in seasonal patterns, the differences between the methods is small. Both procedures can capture basically linear evolution in a seasonal cycle.

The differences are the greatest in series which exhibit strong, non-linear changing seasonal patterns.

As discussed in section 5, smoothness is a desirable feature. Increased flexibility is also important for adapting to changing seasonal patterns. Hence, disregarding other important considerations such as stability, we feel that the X-12-ARIMA seasonal factors are "preferable" in the above series. The *short* term evolution of seasonal patterns for GAUSUM-STM are much "rougher". We are probably better off putting this "excess" local variation of a seasonal pattern into the trend or irregular component. Seasonal patterns which are allowed to evolve into very different shapes in a short time frame, such as those produced by the GAUSUM-STM method, are probably not capturing what "most users" think of as a seasonal effect.

Consider, for example, the IGLCTI series. Figures 8 and 9 give the seasonal decompositions obtained by the X-12-ARIMA and TRIG-6 methods (see appendix B for a description of these and other plots). The seasonal factors for X-12-ARIMA are much more flexible. This is reinforced by figure 10, which compares the seasonal factors for X-12-ARIMA, BSM, TRIG-1 and TRIG-6. The SSI-Plot, figure 11, indicates that much is gained from this additional flexibility. The slow variation of the August effect is captured by X-12-ARIMA. By contrast, the structural models estimate a constant August effect. The power near the fundamental frequency is considerably reduced for X-12-ARIMA: see figure 12.

A less dramatic but equally revealing example is given by the BMNCRS series (figures 13-15). In this case, the structural models are as or more flexible than X-12-ARIMA. However, this flexibility is achieved at a significant increase in roughness. The seasonal factors for structural models exhibit a great deal of seemingly undesirable local variation. The X-12-ARIMA seasonal factors are intuitively more appealing.

These conclusions are in contrast to the study by den Butter and Mourik (1990), who found no consistent difference in the flexibility between the two methods. The criteria used by den Butter and Mourik (1990) is apparently not sufficient to distinguish between the performance of seasonal factors. Not only should flexibility be distinguished from roughness, but it is also necessary to see how well the seasonal factors fit the data.

These results are not surprising. X-12-ARIMA is semi-parametric while GAUSUM-STM is based on parametric structural models. Hence, X-12-ARIMA has significantly more degrees of freedom to devote to the seasonal factors. The smoothness of X-12-ARIMA seasonal factors has been explicitly incorporated into the procedure based on practical considerations. By contrast, the seasonal factors for structural models are based on the maximum likelihood estimates. This does not guarantee seemingly desirable features such smoothness.

Seasonal factors for structural models which mimic those of X-12 might obtained by constraining the variances of the model (see section 7.5). Inclusion of a local AR component may help some by sopping up local variability. A more fundamental solution probably lies in development of a better model for the seasonal component.

Rigidity of the BSM

In all series, the trigonometric seasonal models (TRIG-1 and TRIG-6) provide as or more flexible seasonal factors than the BSM seasonal model. TRIG-1 adapts significantly better to changing seasonal patterns than the BSM in 8 series (B3, B7, B8, B10, B12, I21, I27, I28). For these series, TRIG-1 removes more power around the fundamental frequency, although the seasonal factors are generally rougher.

The BGMRRI series provides a dramatic example of the inadequacy of the BSM seasonal model. Figure 16 gives the SSI-Plot for BGMRRI: the movements in the BSM seasonal factors bear little resemblance to the data! Considerable power is left in the periodogram both at and near the fundamental frequency (see figure 17).

Recall that TRIG-1 and BSM have the same number of parameters. If only one model is to be considered for seasonal adjustment, on the basis of these results, we would prefer TRIG-1.

TRIG-6 is more flexible than TRIG-1

Optimizing over all six variances instead of just one for the trigonometric model makes a significant difference. As to be expected, TRIG-6 is as or more flexible than TRIG-1 in all series. In 15 of the series, this leads to significantly better fits to changing seasonal patterns and removal of power near the fundamental frequency (B1, B3, B6, B9, B10, C14, C15, C17, C19, I21, I23, I25, I26, I27, I28). A good example of the flexibility gained by optimizing over all six variances is given by the BGMRRI series. Often the seasonal factors progress in flexibility from BSM to TRIG-1 to TRIG-6. The X-12-ARIMA seasonal looks like a smoothed version of the TRIG-6 one. See Figures 16 and 17 for a typical example of this.

In regards to goodness of fit, Harvey (1989) argues that it is rarely necessary to optimize over all six variances of the trigonometric seasonal model. The AIC's of Table 3 favor TRIG-6 in several instances. This indicates that optimizing over all variances may often lead to a significantly better fit. In regards to seasonal adjustment, the diagnostic plots show that optimizing over additional parameters is quite important.

Inadequacy of Simple Diagnostics

These results also point out the need for a range of diagnostics and plots to evaluate the effectiveness of a seasonal adjustment procedure. Consider, for example, the simple one-way ANOVA test for residual seasonality, discussed in section 5.1. This test has very little power and is sometimes misleading. For the IFMETI series, the test has a "p-value" of essentially 1 for all structural models, indicating no seasonality remaining in the residuals. On the other hand, the p-value for X-12-ARIMA is 0.761. On the surface, this would indicate that the structural models are quite adequate, and do a better job of removing seasonality from the data. Examination of figures 19 and 20 tells otherwise: X-12-ARIMA seasonal factors adapt to the data in an appealing manner and considerable power is reduced in the periodogram around the fundamental frequency.

The results also indicate that a better fitting model, according to AIC, doesn't mean a more appealing seasonal factor. For example, compared with TRIG-1 or BSM, TRIG-6 has a much lower AIC value for the ITVRUO series. However, examination of the seasonal factors with the SSI Plot does not show a strong preference for TRIG-6. The ITVRUO series is difficult to fit, and in that sense is atypical (the series undergoes a variance shift in the latter portion). A more typical example is given by BLQRRS, for which TRIG-6 has a slightly higher AIC value but significantly more flexible seasonal factors.

6.2 Comparison of outlier treatment methods

Advantages of the GAUSUM-STM method

For many of the series, GAUSUM-STM detects fewer outliers/level shifts or the same number with lower probability (B1, B2, B4, B5, B7, B9, B13, C16, C17, C19, I21, I23, I26, I27, I28). This is illustrated by figure 25, which compares the outlier detection schemes for IFATTI. Eight outliers and ten level shifts are detected by X-12-ARIMA. By contrast, TRIG-6 detects only 3 major level shifts. This is partly a reflection of the rather arbitrary level at which the outlier threshold is set for the X-12-ARIMA procedure. Setting it to a higher level would obviously reduce the number of series for which X-12-ARIMA detects more outliers/level shifts.

In some series, GAUSUM-STM detects numerous low probability level shifts or outliers not identified by X-12-ARIMA (B3, B6, B8, B12, C14, C15, C20, I22, I24, I25, I29). For example, GAUSUM-STM picks up several small level shifts for ICMETI not identified by X-12-ARIMA (see figure 26). These level shifts are barely visible in a plot of the seasonally adjusted data. For the series BLQRRS, GAUSUM-STM models a "ramp" using two successive level shifts of moderate to high probability while X-12-ARIMA uses one large level shift. The GAUSUM-STM approach seems intuitively more appealing in these cases.

For the series listed above, the two procedures handle major outliers and level shifts in a similar manner. The only difference for these series is the way in which the methods handle small outliers and level shifts. The GAUSUM-STM procedure has two apparent advantages. First, it has an automatic way to adapt the "cutoff" level based on the likelihood. Second, it can incorporate small level shifts or outliers by giving them low probability. However, these advantages are more theoretical than practical: the data does not give strong evidence supporting the GAUSUM-STM outlier procedure over the X-12-ARIMA method.

Problems with GAUSUM-STM

In several series, GAUSUM-STM has a problem with detecting moderate outliers if level shifts are present (B1, B3, B12, I21, I23, I24, I26). For the BVARRS series, both X-12-ARIMA and GAUSUM-STM pick up several major level shifts (see figure 29). GAUSUM-STM also detects several additional level shifts with low to moderate probability. The X-12-ARIMA procedure identifies outliers at 12/74 and 3/86. The plot of the seasonal adjusted data strongly supports this: 12/74 and 3/86 stand out as outliers. However, GAUSUM-STM detects 12/74 as a moderate probability level shift and 3/86 as a low probability level shift and outlier. GAUSUM-STM fails to detect the "obvious" in this case.

This an illustration a general problem with the GAUSUM-STM procedure. When the series has mostly level shifts with a couple of outliers, the estimated posterior probability of the occurrence of an outlier tends to be very small. The converse holds as well: when a series has many outliers and a single level shift, the estimated posterior probability of a level shift is small (this is illustrated by BAUTRS).

For BVARRS, there are several obvious level shifts and two moderate outliers. The procedure gives a relatively high probability to level shifts. These give a small degree of protection against the two outliers in this series (by increasing the variance). By giving a low probability to outliers, the likelihood is optimized since it does not incur the "penalty" of modeling outliers when there aren't any (as is the case for all but two observations).

This problem is analogous to that of determining the smoothness of the trend. The trend which is estimated by the likelihood is much rougher than is intuitively appealing. One possible solution to ensure a smooth trend is to constrain the optimization (Gersch and Kitagawa (1983)). In this case, the prior probabilities could be constrained to be greater than a certain value. This could be justified in a Bayesian framework. Setting or constraining the priors is reasonable since the likelihood will be relatively flat in this particular case.

Local level shift or an outlier patch?

X-12-ARIMA and GAUSUM-STM occasionally use very different approaches to modeling non-Gaussian behavior in a number of series (B3, B10, B11, C18, I24, I29). Whereas the GAUSUM-STM tends to treat these series using one or more level shifts, X-12-ARIMA identifies an outlier "patch". A good example of this is given by figure 18, which compares the outlier procedures for the BGMRRI series. The series exhibits a small hump encompassing three observations: 10/75, 11//75, and 12/75. X-12-ARIMA models this by declaring 10/75 and 1/76 as outliers, while GAUSUM-STM identifies these as level shifts. According to the data, both modeling approaches are plausible. In general, when X-12-ARIMA and GAUSUM-STM differ in this way, neither method is demonstrably superior.

Difference in outlier treatments not important

As discussed above, the GAUSUM-STM and X-12-ARIMA outlier procedures differ in significant (although not major) ways. This difference, however, is *usually* not very important in terms of the estimated seasonally adjusted data (some exceptions are given below). The difference in the seasonally adjusted data is primarily due to the different estimation methods of the seasonal factors (see section 6.1). The relative effect of the outlier methods on the seasonal adjustments is generally small. For example, the outlier treatments for the series IFMETI are substantially different: see figure 21. Figure 22 illustrates that the dominant differences in the estimates of the seasonal factors are clearly due to the different way in which the seasonal factors are estimated (see also figure 19).

In a few cases, the outlier treatments lead to quite notable differences in the seasonal adjustments (B1, B3, B6, B10, B12, C16, C17, I23). For CNETHS, the X-12-ARIMA procedure identifies 5 outliers in January and 4 in February. TRIG-6 also identifies most of these outliers, but generally with probability less than one. As a result of the outlier identifications, the X-12-ARIMA seasonal factors for January are elevated: see figure 34. According to the periodogram (figure 35), there is considerable seasonality remaining in the X-12-ARIMA decomposition. For this series, it would seem that exceptionally low January values are part of the seasonal effect. The outlier procedure of X-12-ARIMA is perhaps adjusting too much for these values.

Another example is given by the series BVARRS. Recall that GAUSUM-STM has difficulty in picking up a fairly major outlier in this series at time 3/86 (see the above discussion and figure 29). This outlier clearly seems to have leaked into the seasonal pattern for the GAUSUM-STM decomposition: see figure 30. A less dramatic, and more typical, example is given by the BAUTRS series. Some of the largest differences between the seasonal adjustment are at times in which the outlier treatments are different. However, the difference in the seasonal adjustments for BAUTRS are still mainly due to the different seasonal factors.

For seasonal adjustment, the crucial thing is to deal with moderately large outliers and level shifts in an adequate manner. It is not crucial to handle small outliers or level shifts. The way in which the procedure deals with non-Gaussian behavior is not especially important: e.g., either a local level shifts or an outlier patch may suffice.

Problems with doublets and triplets

Often economic data has two or three adjacent aberrant values, which we call a "doublet" or "triplet". These are due to strikes, weather, or any condition which has a temporary effect on the economy. For example, the CMW1HS series has a doublet at 1/79 and 2/79 (see figure 33), presumably caused by unusually cold weather. X-12-ARIMA models this using a local level shift. This leads to very unnatural looking trends. GAUSUM-STM is not very satisfactory either. It uses a combination of outliers and level shifts to model the patch. This results in a trend which chases after the peaks and valleys.

Both of the procedures can be modified to handle this situation better. An *ad hoc* solution exists for X-12-ARIMA: the identification procedure can automatically search for adjacent or near adjacent level shifts of opposite sign, and replace them with an outlier patch. GAUSUM-STM can be modified by extending the outlier model underlying the procedure to to incorporate outlier patches: see section 7.4.

6.3 Stability of GAUSUM-STM

X-12-ARIMA is less stable according to sliding spans

According to sliding spans statistics, the X-12-ARIMA seasonally adjusted data is less stable than GAUSUM-STM with the BSM fit in 16 series (B1, B2, B7, B8, B9, B12, C14, C15, C18, C19, C20, I24, I25, I27, I28). Note that the BSM fit is used for comparison in this case rather than TRIG-6. X-12-ARIMA is significantly more stable for only two series, and only then because of problems with the initialization of the GAUSUM-STM method (B10 and B11; see section 6.5).

Some tradeoff between stability and flexibility is inevitable. The BSM seasonal factors are much less flexible than those of X-12-ARIMA: see figure 1(e). Hence, it is not surprising that X-12-ARIMA is less stable. For example, the X-12-ARIMA seasonal factors for IFMETI are quite flexible relative to those of the structural models, capturing the changing seasonal pattern (see figures 19 and 20). Correspondingly, the seasonal adjustments are significantly less stable: compare figures 23 and 24. In this case, the decrease in stability is probably a price worth paying for the increase in sensitivity.

A more worrisome cause of instability in X-12-ARIMA seasonal adjustments is the instability in its outlier identification procedure. For example, X-12-ARIMA identifies completely different outliers and level shifts for each span of the series BSPGWS. Figure 27 shows the consequences of an unstable outlier procedure: the sliding spans statistic $MM_t^{\max}$ exceeds 0.25 in one case. By contrast, the maximum value of $MM_t^{\max}$ for the GAUSUM-STM procedure is about 0.09: see figure 28. This problem is discussed in more detail below.

GAUSUM-STM outlier identification procedure is more stable

X-12-ARIMA outlier identification procedure is less stable than the GAUSUM-STM method for 13 series (B1, B2, B8, B9, B12, B13, C15, C17, C18, C19, C20, I25, and I27). In all but one of these series, this instability leads to significantly less stable seasonal adjustments (as measured by the diagnostic $MM_t^{\max}$). One example was shown above with the series BSPGWS. Another example is given by the series BVARRS. In spans 2 and 3, an outlier is identified at 2/86. In span 4, no outlier is identified at 2/86. The corresponding seasonally adjusted data for span 4 is much lower since the influence of the data value leaks into the seasonal (see figure 31). The GAUSUM-STM procedure is quite stable for BVARRS, including both the estimates of posterior probabilities of outliers and the seasonal adjustments (see figure 32).

GAUSUM-STM is less stable for 6 series (B6, B10, B11, C16, I24, I29). Three of these (B10, B11, and I29) are due to a problem with the initialization procedure (see below). In two others (B6 and I24), GAUSUM-STM estimates several additional level shifts with moderate or low probability in one span. This has very little effect on the seasonal adjustments. For CNE1HS, an outlier is only identified in two spans by GAUSUM-STM, leading to quite a large value of $MM_t^{\max}$ for that month (this it the type of instability which is more typical of X-12-ARIMA).

It seems reasonably safe to conclude that the X-12-ARIMA outlier identification procedure is less stable. Furthermore, the instability of X-12-ARIMA has a greater effect on the instability of the seasonal adjustments. These results are not surprising. X-12-ARIMA uses a discontinuous outlier detection and estimation method. This means that fitting to subsets, such as with the sliding spans procedure, is more likely to lead to large changes in the estimates. By contrast, GAUSUM-STM uses a continuous scheme. Similar results should hold with respect to perturbations in the data. Hence, we can expect that GAUSUM-STM is a more robust-resistant procedure than X-12-ARIMA.

It is interesting to note that the outlier identification for X-12-ARIMA in the spans is sometimes quite different than for the entire series (I21, I23, I29). This is due to the different estimate of variance when a shorter segment of the series is used. It is not clear, though, why GAUSUM-STM does not exhibit the same behavior for these series.

6.4 X-12-ARIMA trends are smoother

X-12-ARIMA has smoother trends in all but 3 series (the exceptions are B4, B7, I22). The decompositions of IGLCTI, given in figures 8 and 9, give a typical example of this. The X-12-ARIMA trends are visually more appealing. This is a well known feature of the time series structural model seasonal decomposition. Harvey and Valls Pereira (1989) defend the rough trends yielded by structural models.

Smoother trends for structural models can be obtained simply by constraining the variances (see section 7.5). Including a local AR trend in the STM's may also produce more stable trends, although the fitting would become more difficult and perhaps unstable.

6.5 Problems with GAUSUM-STM Initialization

As discussed towards the end of section 3.3, the current method for initializing the filter for the GAUSUM-STM procedure is inadequate. Problems with initialization crop up in three series: BTAPRI, BTNDRI, and ITVRUO. These series have poor or unstable estimates of the seasonal factors. The problems with ITVRUO stem from having large outliers within the first thirteen observations. With the current initialization procedure, GAUSUM-STM identifies adjacent level shifts in the fourth span. In all other spans, these observations are identified as outliers.

The seasonal pattern for BTAPRI and BTNDRI drops dramatically from November to December. Instead of capturing this drop in the seasonal, GAUSUM-STM partially models it using a level shift. As a result, the estimated seasonal factors for BTNDRI are quite poor: see figure 36. Note that this effect is worst at the beginning of the series. Considerable power is left in the periodogram for BTNDRI.

7 More on GAUSUM-STM

7.1 Estimation of standard errors

A big advantage of a model based procedure over X-12-ARIMA is in the availability of standard errors for the seasonally adjusted data. GAUSUM-STM is especially good in this regard, since it incorporates outliers and structural changes within the model. GAUSUM-STM actually produces an estimate of the posterior density, not just standard errors. Kitagawa (1987) gives several nice examples of the advantages of non-Gaussian confidence intervals. See Kitagawa (1988) for examples in the context of seasonal adjustment.

In the parametric outlier procedure used by X-12-ARIMA, the estimate of standard errors is not as realistic. This is partly because the standard errors are calculated under the assumption that the location of the outlier or level shift is *known*. This can make a big difference, since the timing of a local level shift or outlier patch is often in doubt, especially towards the ends of the series. In addition, the intervals produced by the ARIMA outlier identification procedure are purely Gaussian, and cannot capture the long tailed nature of the densities.

Figure 2 gives 99% confidence intervals for the seasonal adjustments for the BVARRS series. The top plot shows the seasonally adjusted data with the intervals and the bottom plot indicates the width of the intervals. The intervals widen towards the ends of the series, reflecting "end effects". The intervals also tend to get wider near level shifts and outliers.

The importance of non-Gaussian intervals is better exhibited by the 99% confidence intervals for the trend, which are given in figure 3. Each of the level shifts identified in figure 29 are associated with a significantly wider confidence interval. Note that the greatest uncertainty is not near the large level shift in 1976, which is easy to identify and model. Rather, it is near the series of smaller level shifts in 1986. The confidence interval is over twice as wide near the patch of level shifts in 1986. The GAUSUM-STM confidence regions do a good job of reflecting the difficulty of tracking the trend. One could not expect the parametric outlier identification procedure of X-12-ARIMA to perform as well in this regard.

7.2 Modeling calendar effects

In the fits done for GAUSUM-STM, trading day and Easter effects were handled by prior adjustment based on X-12-ARIMA. Optimizing over trading day and Easter regression variables is not very critical, and is unlikely to lead to significantly different results. This is mainly because both GAUSUM-STM and the ARIMA model underlying X-12-ARIMA give reasonable fits to the data and adequately deal with outliers and level shifts. Hence, fitting a fixed effects regression variable such as for trading day should be roughly equivalent with either procedure.

For example, the trading day coefficients were optimized for the BAUTRS series. The top plot in figure 4 shows the original "REGCMPNT" calendar effect. The ratio of the calendar effect obtained by optimizing over the regression coefficients to the REGCMPNT calendar effect is given in the second plot. Clearly further optimization in this case makes little difference.

7.3 Building in Ramps and Other Outlier Models

For the sake of parsimony, simplicity, and computational efficiency, the model used to fit the series only accommodated level shifts. A slight generalization of this model can be obtained by inflating both the variances of η_t and ξ_t in the second component of the Gaussian mixture model (7):

$$\begin{pmatrix} \eta_t \\ \xi_t \end{pmatrix} \sim \begin{cases} \mathrm{N}\left(\mathbf{0}, \begin{pmatrix} \sigma_\eta^2 & 0 \\ 0 & \sigma_\xi^2 \end{pmatrix}\right) & \text{with probability } 1 - \epsilon_{\eta,\xi} \\ \mathrm{N}\left(\mathbf{0}, \begin{pmatrix} \tilde{\sigma}_\eta^2 & 0 \\ 0 & \tilde{\sigma}_\xi^2 \end{pmatrix}\right) & \text{with probability } \epsilon_{\eta,\xi} \end{cases}.$$

This more general "ramp" yields very similar estimates for the trend or seasonal in a representative subset of eight series. The series experiencing the largest change in the estimated trend is IFMETI, for which the maximum difference is only ±0.5%.

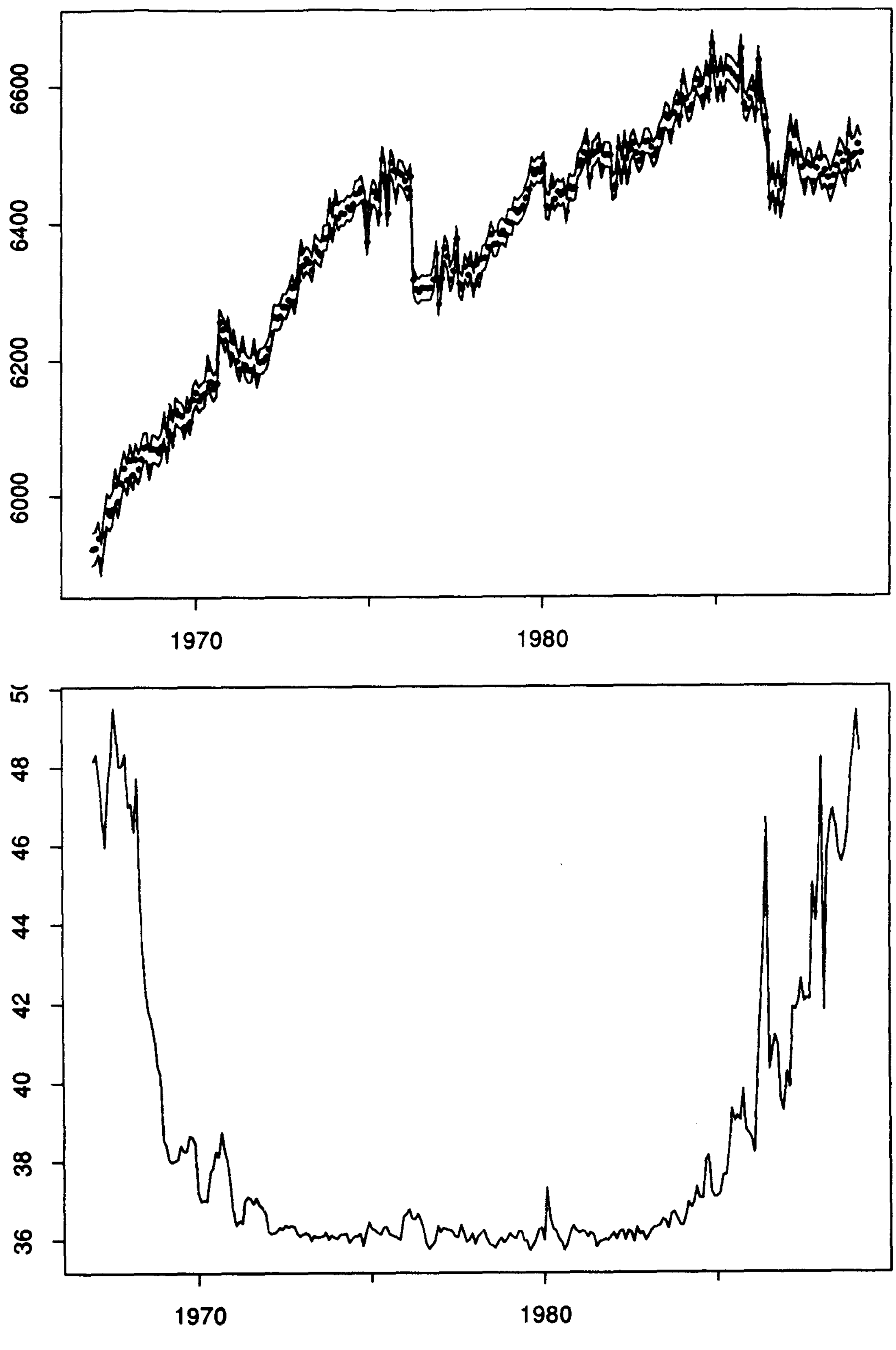


Figure 2: Seasonally adjusted data for BVARRS with 99% confidence intervals (top plot) and the width of the confidence intervals (bottom plot).

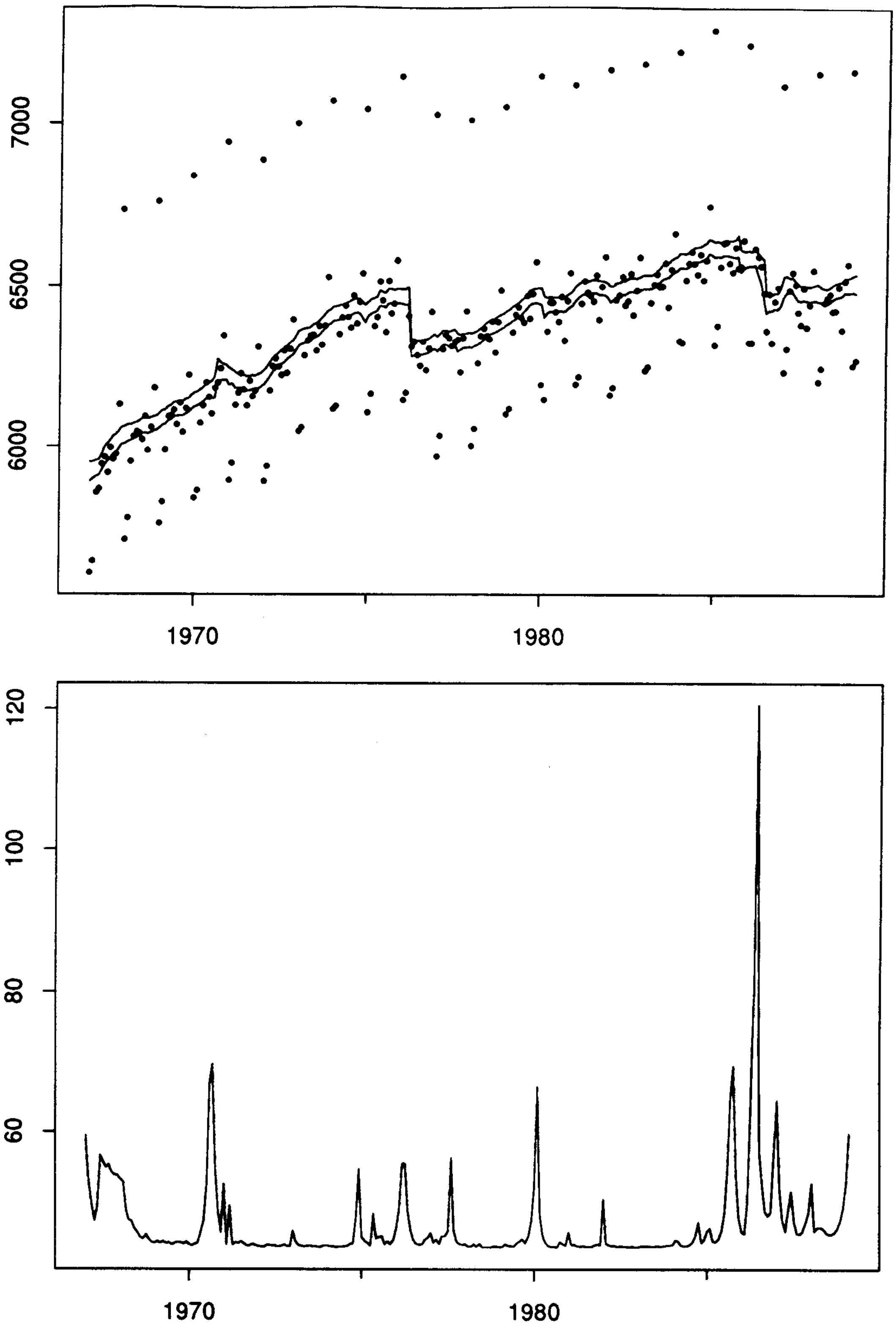


Figure 3: The BVARRS data with 99% confidence intervals for the trend (top plot) and the width of the confidence intervals (bottom plot).

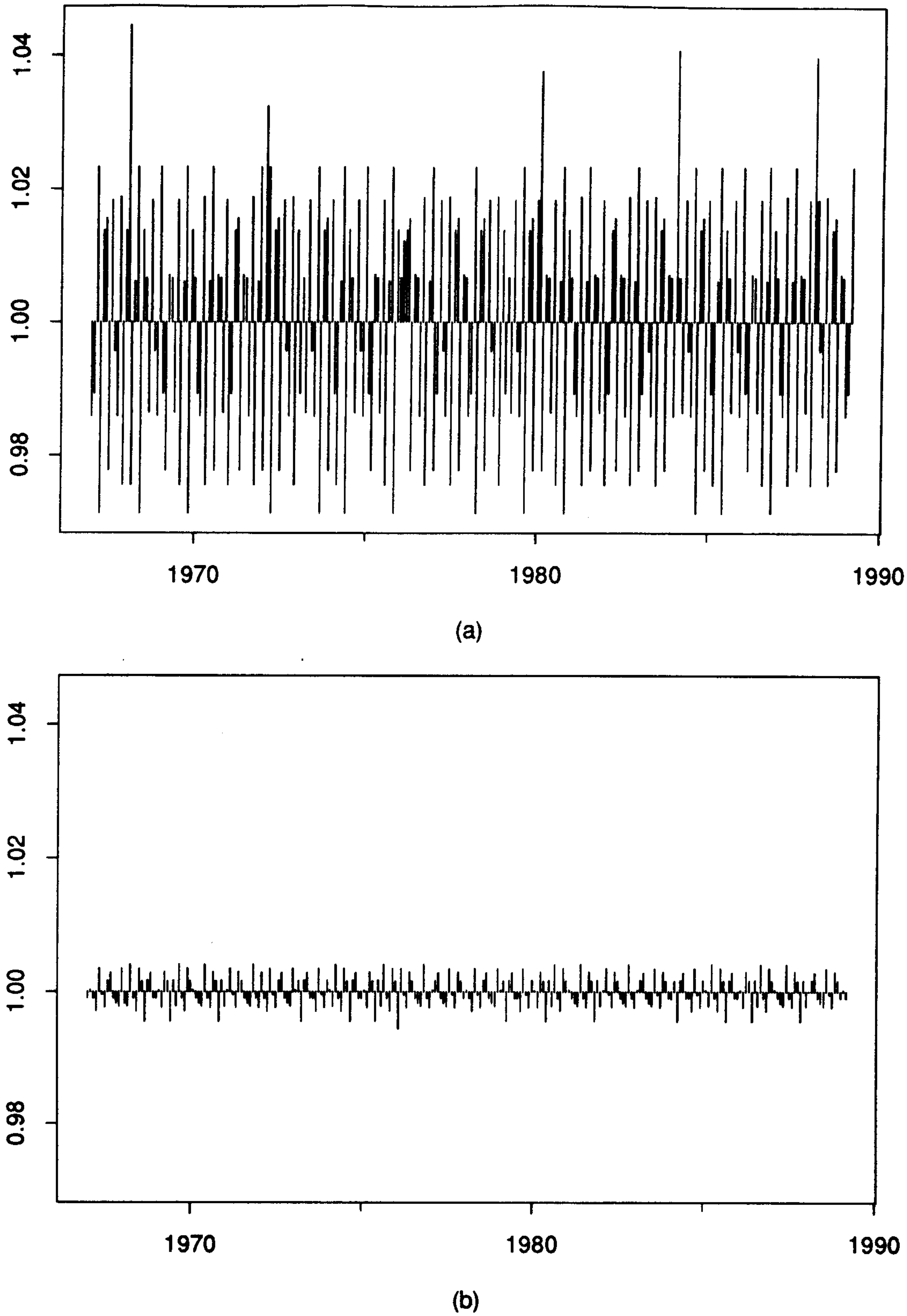


Figure 4: (a) The original REGCMPNT calendar effect for BAUTRS and (b) the ratio of the optimised calendar effect to the original.

Figure 5 compares the trends for the model with and without the "ramp" effect. The top plot displays the log-transformed seasonally adjusted data along with both trends. Note that the trends are virtually indistinguishable. The second plot displays the differences between the transformed seasonally adjusted series. The third plot compares the probability of detection of a level shifts and outliers, with the dashed lines corresponding to the "ramp" model.

The interesting thing to note here is that the ramp model detects a patch of very high probability structural changes in the beginning of 1975. This corresponds to the sudden change in the slope of the series at that time. This indicates that the main benefit from including ramps may be to improve the overall fit from the model.

7.4 Handling Doublets

GAUSUM-STM does not produce very appealing decompositions in the presence of adjacent outliers (called "doublets" or "triplets"). A fairly unappealing spike appears in the trend, which "chases" after the outliers. An example of this is given by the CMW1HS series (see figure 33).

This behavior of GAUSUM-STM stems from shortcomings in the model: the occurrence of an outlier is assumed to be independent of whether an outlier occurred at the previous observation. This is counter to what we know about economic (and many other) time series: outliers often come in patches. Indeed, the number of outliers in a patch often depends on the sampling interval. A sensible generalization of the outlier model is to allow Markov behavior in the outlier generating process.

Let Z_t be a 0-1 process which indicates whether an outlier has occurred at time t. We assumed in section 3.2 that $p(Z_t = 1|Z_1, \ldots, Z_{t-1}) = p(Z_t = 1) = \epsilon_{\mathrm{I}}$. A more natural assumption is

$$p(Z_t = 1|Z_1, \ldots, Z_{t-1}) = p(Z_t = 1|Z_{t-1}) = \begin{cases} \epsilon_{\mathrm{I}} & \text{if } Z_t = 0 \\ \epsilon_{\mathrm{I}}^0 & \text{if } Z_t = 1 \end{cases} \tag{17}$$

where $\epsilon_{\mathrm{I}}^0 \gg \epsilon_{\mathrm{I}}$. Hence, if an outlier occurs at time $t-1$, than an outlier is much more likely to occur at time t.

Figure 6 compares the trends obtained by fitting the "doublet" outlier model (17) and the usual model (6) for the CMW1HS series. In the doublet model the outlier pair in 1979 is excluded completely from the trend. In addition, although masked in plots due to the size of the outlier, the trend for the doublet model smooths several sharp peaks prominent in the original trend. For this example, ϵ_{I}^0 is set to 0.25.

7.5 Constraining the Variances

To find a model "similar" to the X-11 decomposition, a constrained version of the basic structural model was fit to the series (BSM-CONS). To match *acf*'s, the

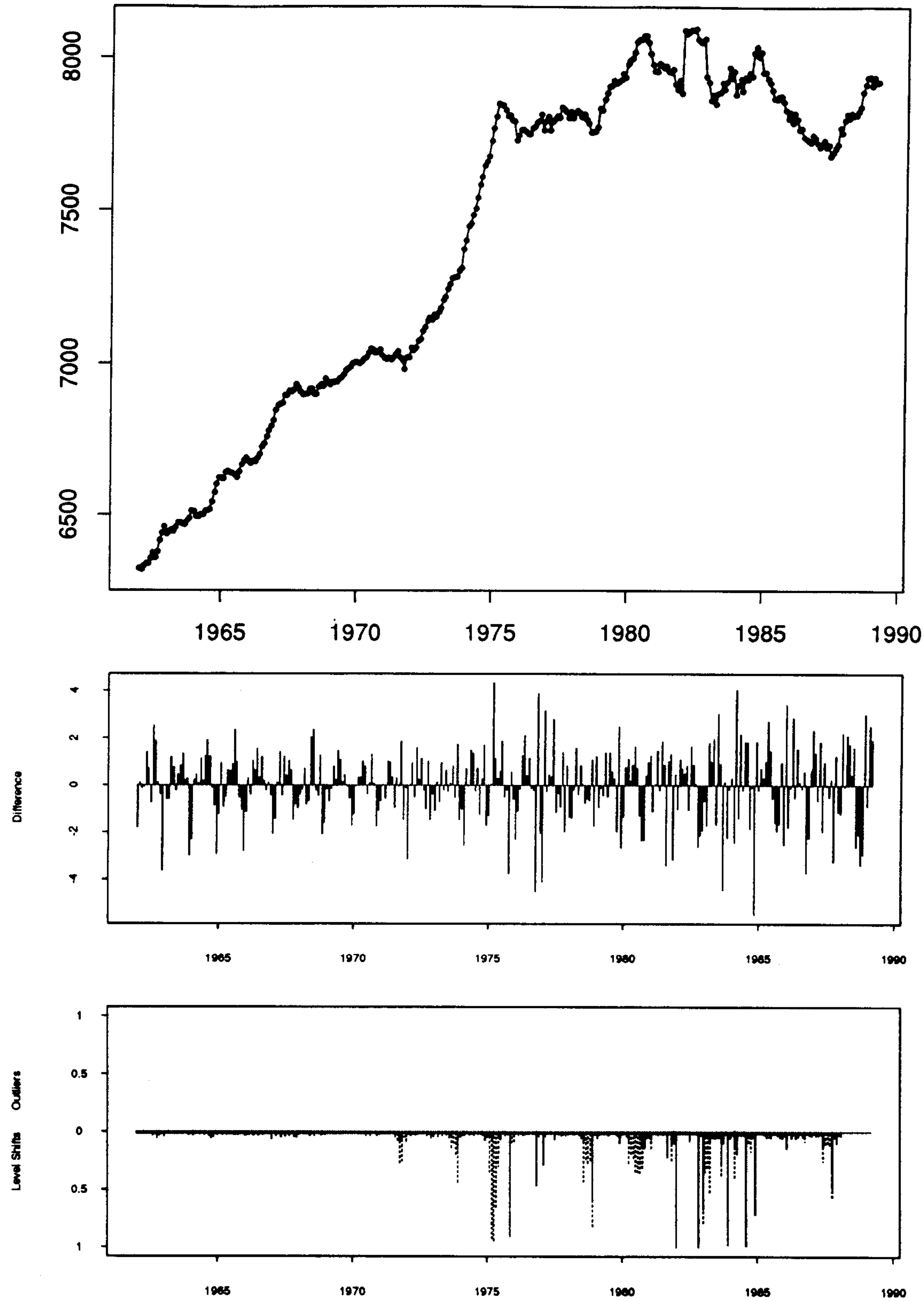
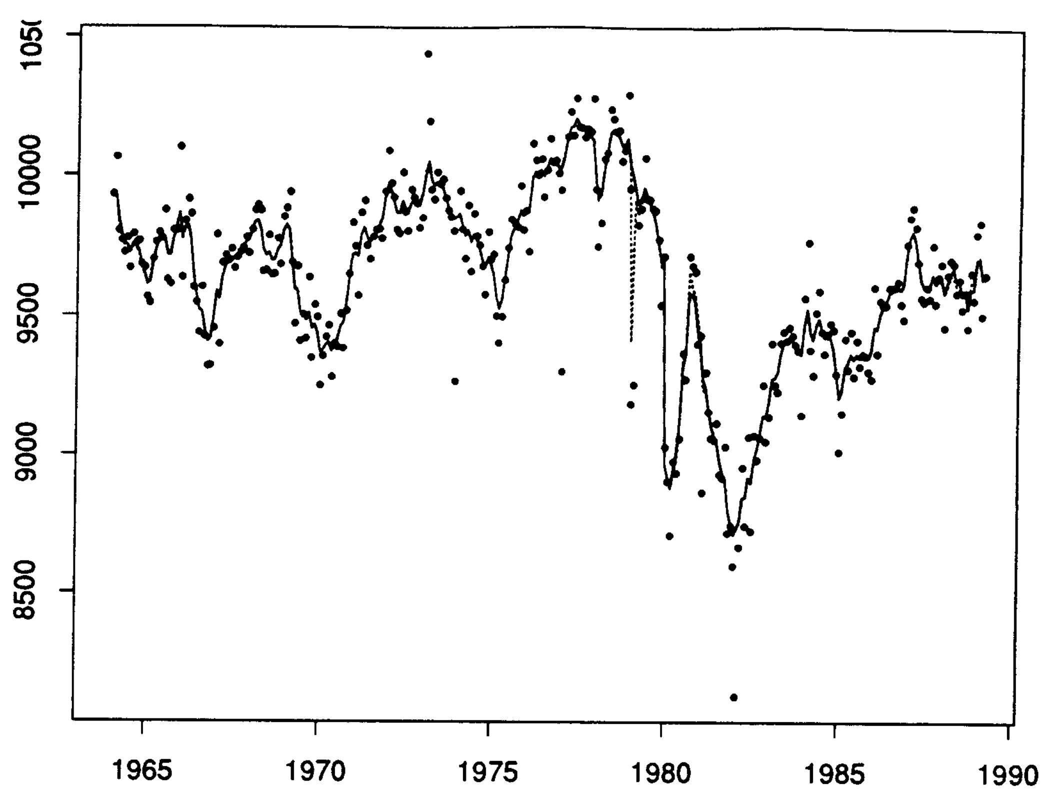


Figure 5: The top plot shows log transformed trend (line) and the seasonally adjusted data for IFMETI (points). The second plot shows the difference between trends with and without the "ramp" effect (indistinguishable in the top plot). The bottom plot gives the posterior probability of AO's and LS's with and without the ramp effect (dashed and solid lines).



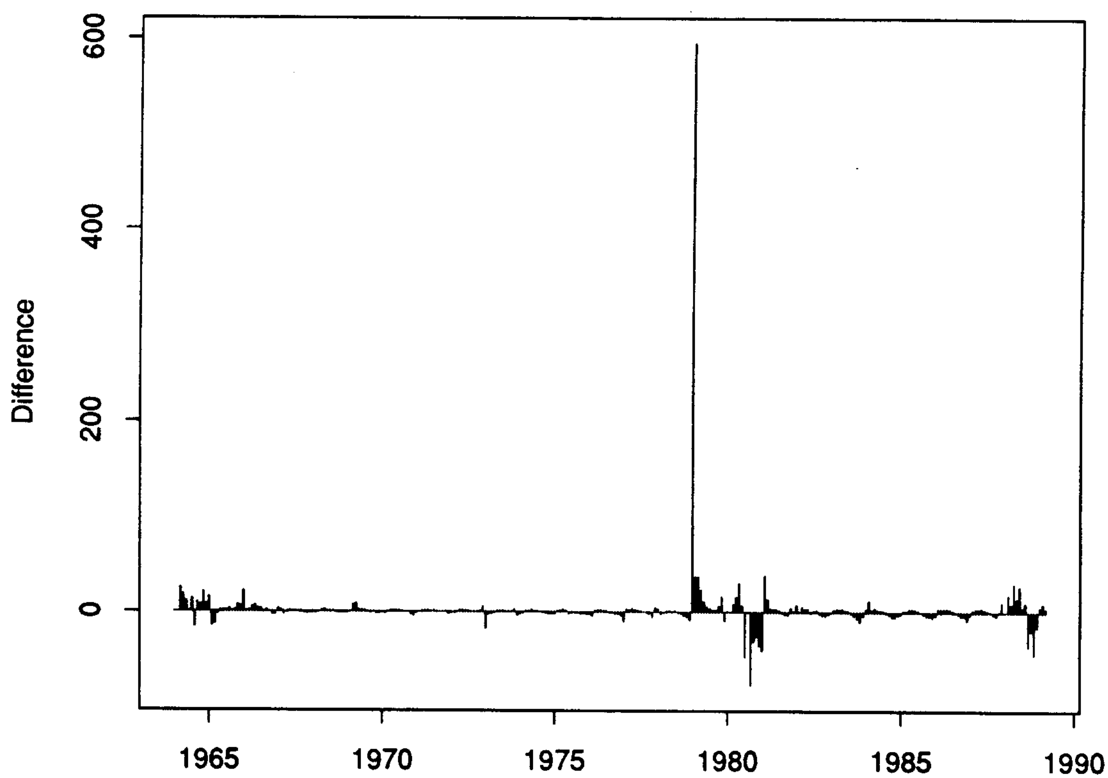


Figure 6: The top plot compares of the trends from the "doublet" model (solid line) and the BSM model (dashed line) for the CMW1HS series. The bottom plot gives the difference between the trends.

variances of BSM-CONS are set as in (9) (see Maravall (1985)). This results in much smoother trends: see the boxplots of figure 1(b). However, for the business and inventory series, the seasonal factors are only slightly more flexible and rougher (see figure 1(e)-(f)).

For the construction series, the seasonal factors are much more flexible and rougher. These series are noisier, and the longer filters of X-12-ARIMA produce smoother seasonal factors than the default filters. Hence, it is not surprising that the constrained model leads to very flexible but rough factors for the construction series.

A typical example of the difference constraining the parameters makes on the seasonal factors is given in figure 7. The constrained BSM is slightly more flexible than the BSM, but is also significantly rougher. Neither the BSM nor BSM-CONS match the seasonal factors of X-12-ARIMA in terms of smoothness and fit to the data.

It seems likely that simply constraining the parameters of the BSM will not lead to a decomposition very similar to X-12-ARIMA. Maravall (1985) acknowledges this by noting that equivalent *acf*s do not translate into equivalent decompositions. Increased flexibility for the BSM seems inevitably to require increased roughness. A more promising but more complicated approach is to constrain the variances in the TRIG-6 model.

8 Open Problems and Conclusions

On one level, this study can be viewed as an endorsement of X-12-ARIMA. The procedure adequately handles most of the series with both outliers and structural changes. The decompositions would appear to be more appealing than those generated by an STM based method.

However, X-12-ARIMA has some significant shortcomings, such as the discontinuous nature of the outlier identification procedure. In addition, a procedure such as GAUSUM-STM offers several potential advantages, including estimates of standard errors, generalization to multivariate seasonal adjustment, and an appealing underlying methodology. Hence, we should not give up on alternatives to X-12-ARIMA.

Several issues need attention if GAUSUM-STM is to become a serious competitor to X-12-ARIMA.

- We feel that the seasonal models used in this paper are inadequate. The assumptions underlying these models need to be investigated. Alternative models should be explored: see, for example, Harvey and Valls Pereira (1989) and Hannon et al. (1970).

- Inclusion of a local AR component may help in a number of ways. By sopping up local variability, it may cause both the trend and the seasonal term to

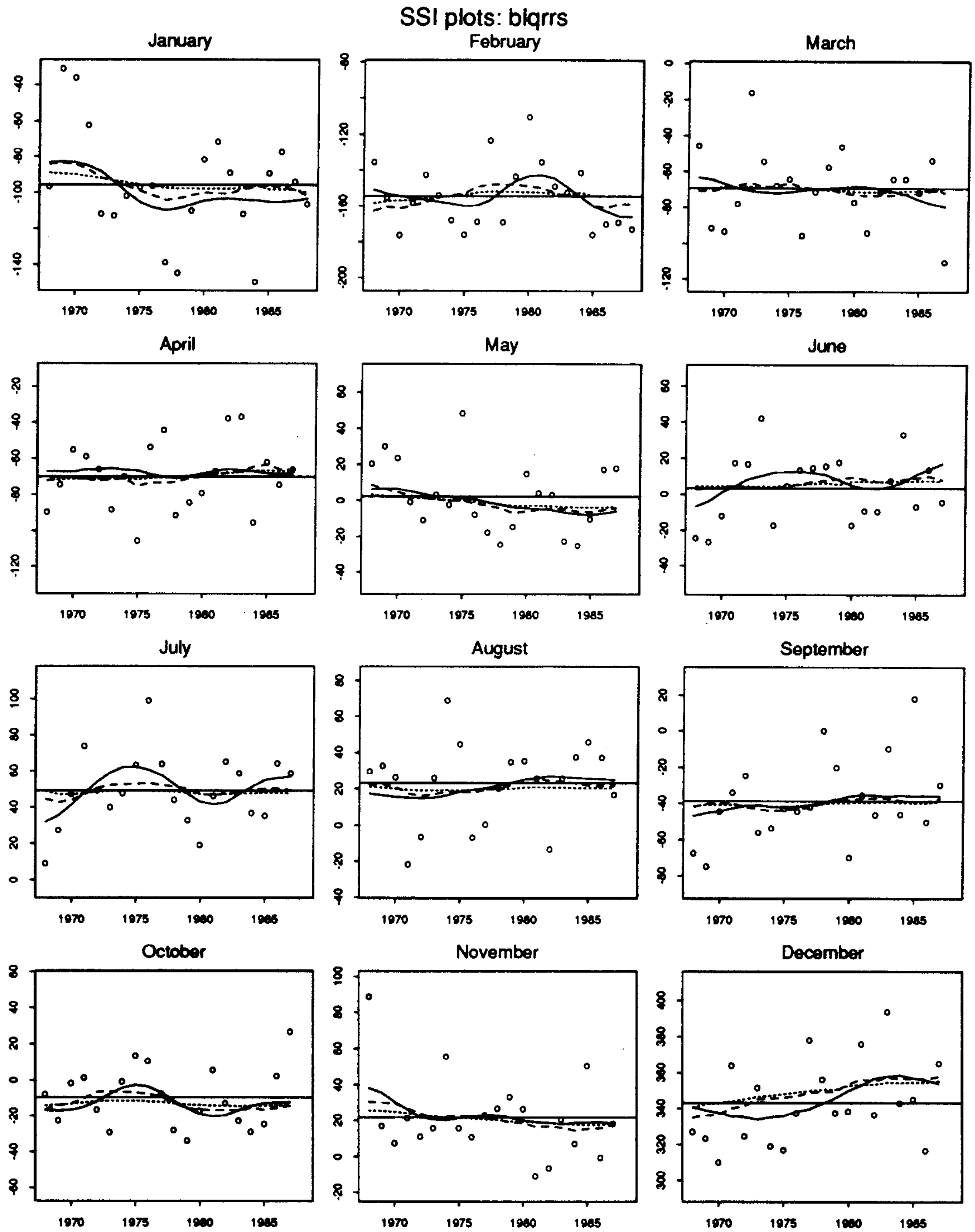


Figure 7: Comparison of the seasonal factors for the BSM model (short dashed line), BSM-CONS model (medium dashed line), and X-12-ARIMA (solid line) for the BLQRRS series. BSM-CONS offers only a slight improvement in flexibility over BSM, and only at the cost of "rougher" factors. Neither matches the flexibility of the X-12-ARIMA seasonal factors.

be smoother. Note that inclusion of an AR term will involve more difficult optimizations.

- The initialization method used for GAUSUM-STM is clearly inadequate, as exhibited by several examples. One possible solution is to use an "EM approach", estimating the posterior probabilities of outliers and level shifts in the beginning of the series using a backwards filter. This method is relatively time consuming to program. A simpler, but computationally expensive, approach is to estimate the initial conditions (see De Jong (1988) for the Gaussian case). See Bruce (1992) for further discussion.

- GAUSUM-STM is relatively slow. Several changes could be made to speed up the likelihood evaluations. In particular, use of a different criteria to determining which densities should be collapsed could lead to substantial improvements. More dramatic computation savings could be achieved if an adaptive tree approach, as in Bruce and Martin (1992), is adopted.

- Only a couple of possible outlier models were used in this study. It is worthwhile investigating whether more complex models offer any significant improvement. In particular, experimentation needs to be done in terms of modeling seasonal breaks. Several of the series (e.g., BGMRRI) seemed to exhibit a seasonal break, and this appeared to cause problems in fitting the models.

- The initial parameters for the optimization of GAUSUM-STM were obtained by first running X-12-ARIMA and then running REGCMPNT. For GAUSUM-STM to be useful as a stand-alone routine, a good robust but fast method needs to be developed to estimate initial parameters for the optimizer.

- The seasonal factors for structural models are based on the maximum likelihood estimates of the models discussed in section 3.1. It is possible more appealing seasonality can be obtained by simply constraining the variances in the existing models. However, as section 7.5 indicates, this approach is not promising with the BSM.

- Fitting time series structural models by maximum likelihood is often a difficult task due to the flatness of the likelihood. It may be possible to "tune" the optimizer to obtain successful convergence. A more fundamental solution to this problems lies in finding alternative transformations of the parameters for easier optimization.

A Fitting Details for GAUSUM-STM

The likelihoods are maximized using the quasi-Newton nonlinear optimizer of Gay (1979) (see also Dennis et al. (1981)). The optimizer uses a trust region approach with a double dogleg step. Finite difference gradients are used with the BFGS secant update to the hessian.

The initial values to the optimizer for the BSM are obtained by fitting the BSM Gaussian model. To ensure robust initial estimates, the outlier identification scheme of X-12-ARIMA are used to first identify AO's and LS's. These are included in the model as fixed regression effects. The fits are done using the program REGCMPNT (Monsell and Otto (1991)), which is more efficient than GAUSUM-STM for purely Gaussian models. The TRIG-1 model was fit with the initial values derived from the maximum likelihood estimates for the BSM. The TRIG-6 model was fit with the initial values derived from the maximum likelihood estimates for TRIG-1.

Prior adjustment is done for trading days based on the REGCMPNT procedure. While GAUSUM-STM accommodates fitting trading day variables, this involves nonlinear optimization over six parameters, greatly increasing the computations. Some examples indicate that further refinement of the estimates for trading day effect is not important (see section 7.2).

Convergence Criteria

The optimizer is considered to have converged successfully if little improvement has been achieved in the objective function from the previous iteration. This is known as "relative function convergence", and is satisfied if

$$\frac{|\log \mathrm{L}(j) - \log \mathrm{L}(j-1)|}{|\log \mathrm{L}(j)| + |\log \mathrm{L}(j-1)|} \leq 0.00005 \tag{18}$$

where $\log \mathrm{L}(j)$ is the log-likelihood on the j-th iteration. Alternatively, the optimizer converges if the change in the estimated parameters is small. This is known as "relative X-convergence" and is satisfied if

$$max_{i=1,\ldots,p} \left\{ \frac{|\hat{\alpha}_i(j) - \hat{\alpha}_i(j-1)|}{|\hat{\alpha}_i(j)| + |\hat{\alpha}_i(j-1)|} \right\} \leq 0.005. \tag{19}$$

where $\hat{\alpha}$ is a vector of the scaled parameters.

B Description of plots

The empirical comparison was done mainly on the basis of a set of 9 diagnostic plots which were generated for each of the 29 series. Naturally, only a small subset of these plots are shown here for illustrative purposes. The complete "book" of plots is given in Bruce and Jurke (1992). The following is a description of the diagnostic plots.

(a) X-12-ARIMA Decomposition

The first plot shows the X-12-ARIMA log-additive seasonal decomposition of the series. An example of this plot is given in figure 8. The untransformed data, trend, seasonal, calendar (if present), and irregular components are all plotted. The seasonal, calendar and irregular components are factors centered on 1. Multiplying these factors by the trend gives the original data. To the right of the plots of the seasonal, calendar, and irregular is a bar which portrays the relative scaling of that plot. The bar spans the same number of units in each plot, so a shorter bar implies a more variable component.

(b) GAUSUM-STM Decomposition

This plot is identical to (a) except that it is for the GAUSUM-STM decomposition with the TRIG-6 model (the trigonometric seasonal model optimized over six variances). The median of the distributions are used to estimate the seasonal, trend, and irregular. An example of this plot is given in 9.

(c) Seasonal Adjustment Comparison

This consists of three plots as in figure 22. The top plot compares the seasonally adjusted series (trend$\times$irregular) given by TRIG-6 (dashed line) and X-12-ARIMA (solid line). The data points plotted are the original series. The data and seasonally adjusted data are untransformed.

The middle of the three plots shows the ratio of the two untransformed seasonally adjusted series from the top plot. Tick marks on the interior of the top axis of the plot show the times at which X-12-ARIMA identifies additive outliers. Tick marks on the interior of the bottom axis show times at which levels shifts are identified by X-12-ARIMA. These are given to see if there is any association between outliers/level shifts and large discrepancies in the seasonally adjusted data.

The third plot compares the untransformed trends yielded by the two procedures. The X-12-ARIMA trend is the solid line. As in the middle plot, tick marks on the top and bottom axes show the locations of additive outliers and level shifts respectively. In addition, the identified outliers are plotted as circles, and a triangle is plotted on each end of the X-12-ARIMA trend line connecting the two points on either side of a level shift.

(d) Outlier Treatments

This consists of three plots as in figure 21. The first plot is similar to the bottom plot of (c), except that it compares the 1000 $\times$ log transformed trends from the two procedures. The points correspond to seasonally adjusted data from the TRIG-6 fit.

The second plot shows the difference between the transformed trends from the two procedures. Outliers and level shifts identified by the X-12-ARIMA are displayed using circles and triangles respectively.

The third plot shows the probabilities of outliers and level shifts for the GAUSUM-STM procedure. The probability of an outlier at each point in the series is plotted as a vertical line extending upwards from the zero line. Similarly, level shift probabilities extend downwards. Times at which outliers and level shifts are identified by X-12-ARIMA are shown by dashed lines of unit length in the appropriate direction.

(e) Month Plot

A useful plot for comparing the seasonal factors of procedures is the "Month-Plot" (Cleveland and Terpenning (1982)). An example of a Month-Plot is given in figure 10. In the version of the Month-Plot used in this paper, the average seasonal cycle for each procedure is displayed in the top plot. The procedures plotted are the BSM (short dashed line), TRIG-1 (medium dashed line), TRIG-6 (long dashed line) and X-12-ARIMA (solid line). In the subsequent plots, the deviations from the average cycle for each year are displayed by month for each procedure. This plot is especially good at contrasting the difference in the range of the seasonal factors.

(f) SSI Plot

The best diagnostic for determining the nature of the seasonal factors is the "SSI-Plot" (Cleveland and Terpenning (1982)). An example of an SSI-Plot is given in figure 11. For each month, the SSI-Plot displays the detrended transformed data with the seasonal factors superimposed as a line. The procedures and line types are the same as those used in plot (e). The horizontal solid line corresponds to the mean.

The data is detrended by first removing the level shifts and outliers as estimated by X-12-ARIMA from the data and trend. Then, a smoothed X-12-ARIMA trend is subtracted off the data. The reason for smoothing the X-12-ARIMA trend is that different procedures tend to lead to different trends. Smoothing the trend will minimize these differences. The trend is smoothed using a simple 23 point triangular moving average. As a result of the smoothing procedure, a years worth of data is dropped from each end.

Usually the data for the SSI-Plot is detrended by simply subtracting the trend obtained from the seasonal decomposition. The above method was adopted in this case since we are comparing several decompositions with possibly sharp breaks in the trend.

(g) Periodogram Plot

This consists of a sequence of periodograms, each one corresponding to a seasonal adjustment. An example of this plot is given in figure 12. If the seasonal adjustment procedure is successful, then the periodogram should have very little power at or near the fundamental frequency $\omega = \pi/6$ and its harmonics $\omega_j = j\pi/6$ (indicated by the dashed lines).

The top periodogram is of the transformed and detrended data. The power at the fundamental frequency and its harmonics is set to zero. This corresponds roughly to what one would obtain by fitting a fixed seasonal pattern. The data is detrended using a smoothed version of the X-12-ARIMA trend. As in figure (f), a 23 point triangular moving average smooth of the trend is used after adjusting for any level shifts identified by X-12-ARIMA. The subsequent periodograms are of transformed and detrended seasonally adjusted data from the BSM, TRIG-1, TRIG-6 and X-12-ARIMA procedures. The seasonally adjusted data is detrended as above is used except that the appropriate trend is substituted for the X-12-ARIMA trend.

(h) Sliding Spans Plot for GAUSUM-STM

This figure consists of four plots relating to sliding spans applied using the GAUSUM-STM procedure for the BSM. An example of the sliding spans plot for GAUSUM-STM is given in figure 23. The first plot shows the untransformed seasonally adjusted data obtained when the procedure is applied to each of four 8 year spans.

The second plot displays $MM_t^{\max}$ of (16). $MM_t^{\max}$ is the maximum difference in the month-to-month percentage changes in the seasonally adjusted data . As a benchmark to judge stability, dashed horizontal lines are drawn at $k \times 0.01$ for $k = 3, 4, \ldots$. The cutoff of 0.03 comes from Findley et al. (1990). Adjustments with more than 25% of the months with $MM_t^{\max} > 0.03$ are almost never acceptable. Good seasonal adjustments seem to have less than 15% greater than the cutoff. This criteria is designed for the X-11-ARIMA procedure and may not be suitable for the structural models. The histogram at the right of this plot shows the distribution of the $MM_t^{\max}$. Boxplots of the $MM_t^{\max}$ by month are shown in the fourth plot.

The third plot of this figure is in fact four plots, each corresponding to a different span. The plots show the probability of outliers and level shifts, as in the third plot of figure (d). As with the third plot of figure (c), tick marks on the interior of the upper and lower axes show times at which X-12-ARIMA identified outliers and levels shifts respectively.

(i) Sliding Spans Plot for X-12

This figure gives the sliding spans diagnostics for X-12-ARIMA. An example of this plot is given in figure 24. The first, second and fourth plots are analogous to those given figure (h). The third plot shows the outliers and level shifts detected by X-12-ARIMA in each span. Level shifts are indicated by "steps". Outliers are indicated by points, and connected to the level by vertical lines. The size of the step or vertical line indicates the relative magnitude of the level shift or outlier.

List of Figures

The figures on the following pages are selected from the "book" of plots given in Bruce and Jurke (1992). The table below provides a summary of the figures that have been selected and the gives pages in this report on which they occur.

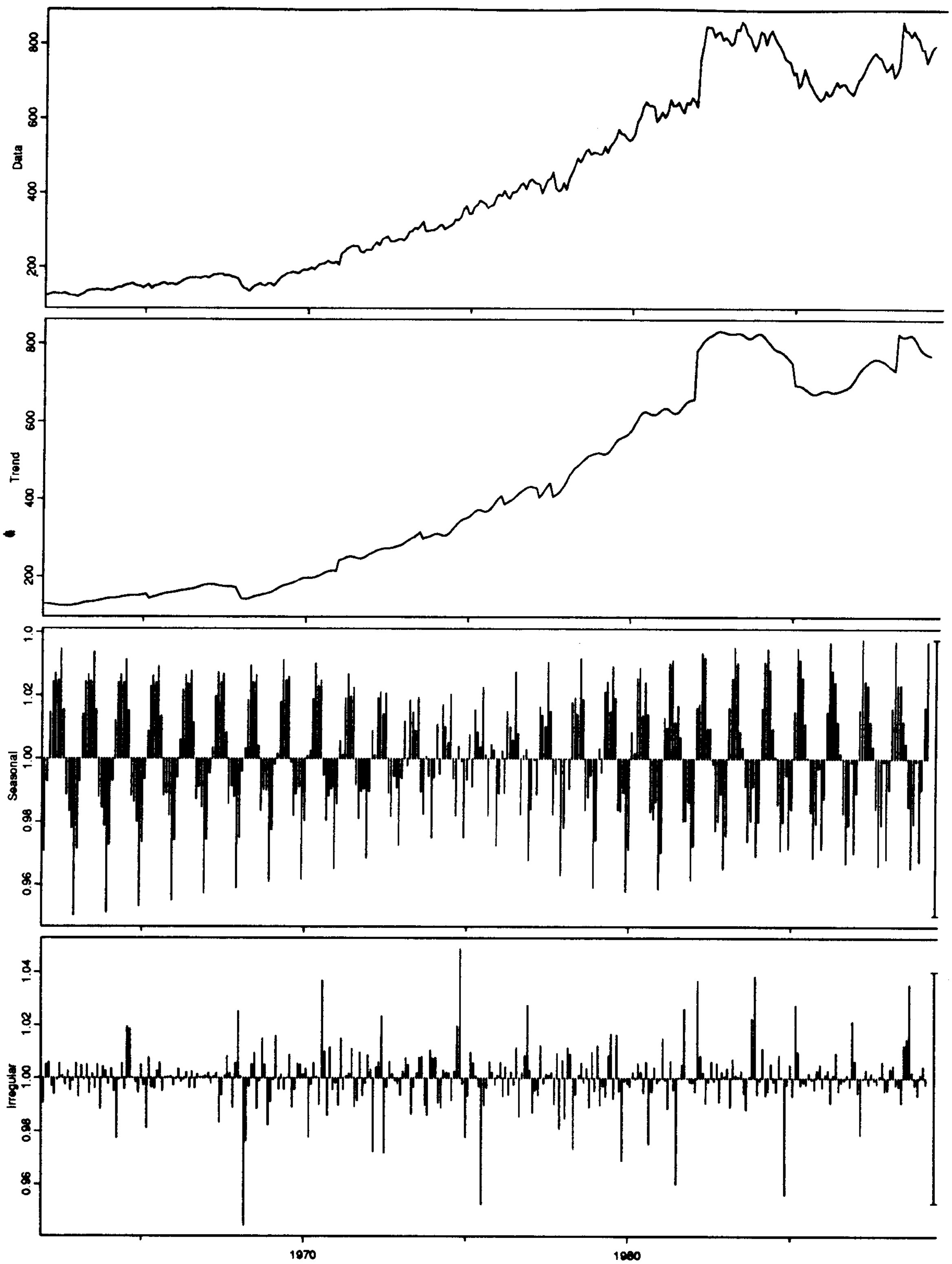


Figure 8: X-12-ARIMA Decomposition for IGLCTI.

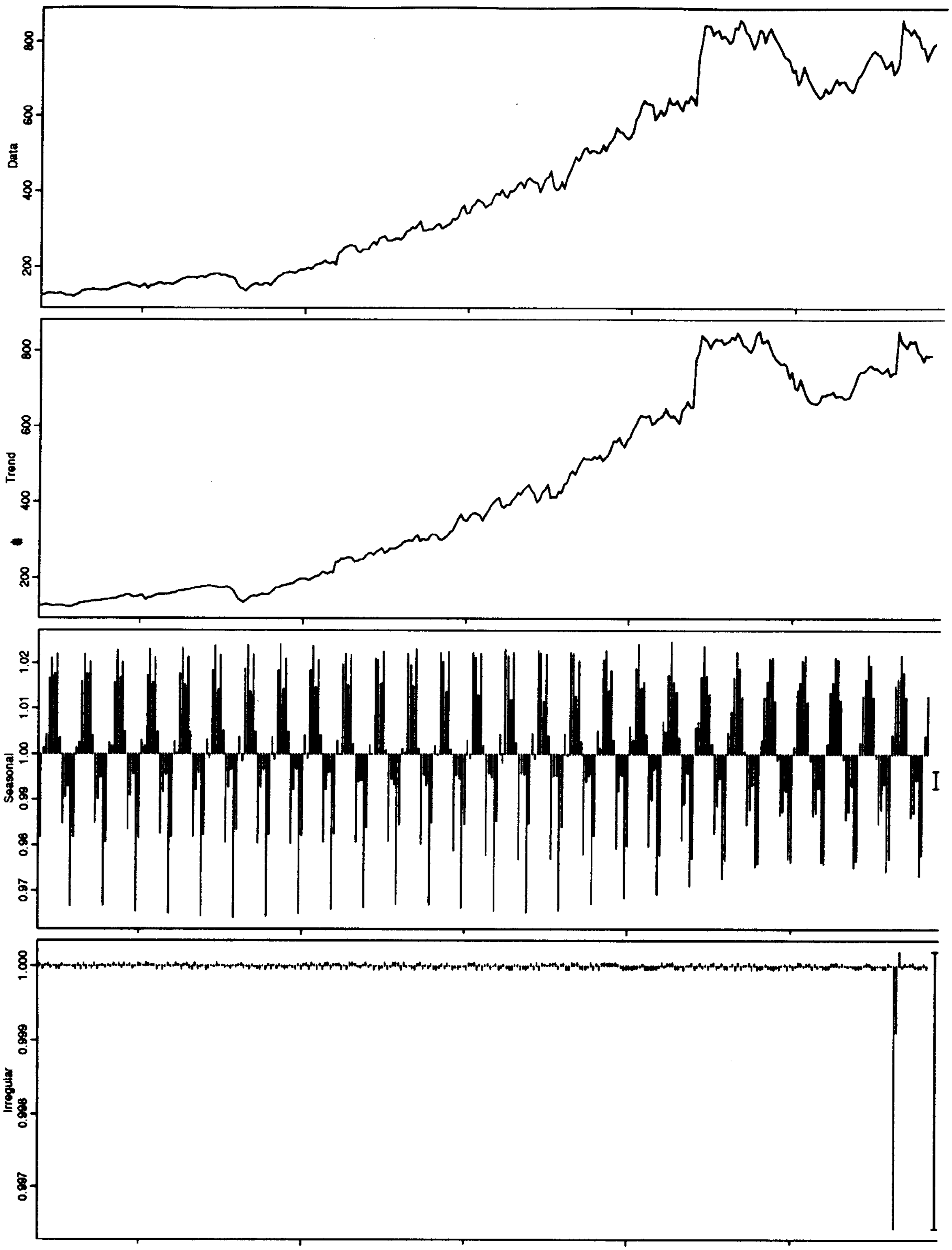


Figure 9: Gaussian Sum Decomposition (TRIG-6) for IGLCTI.

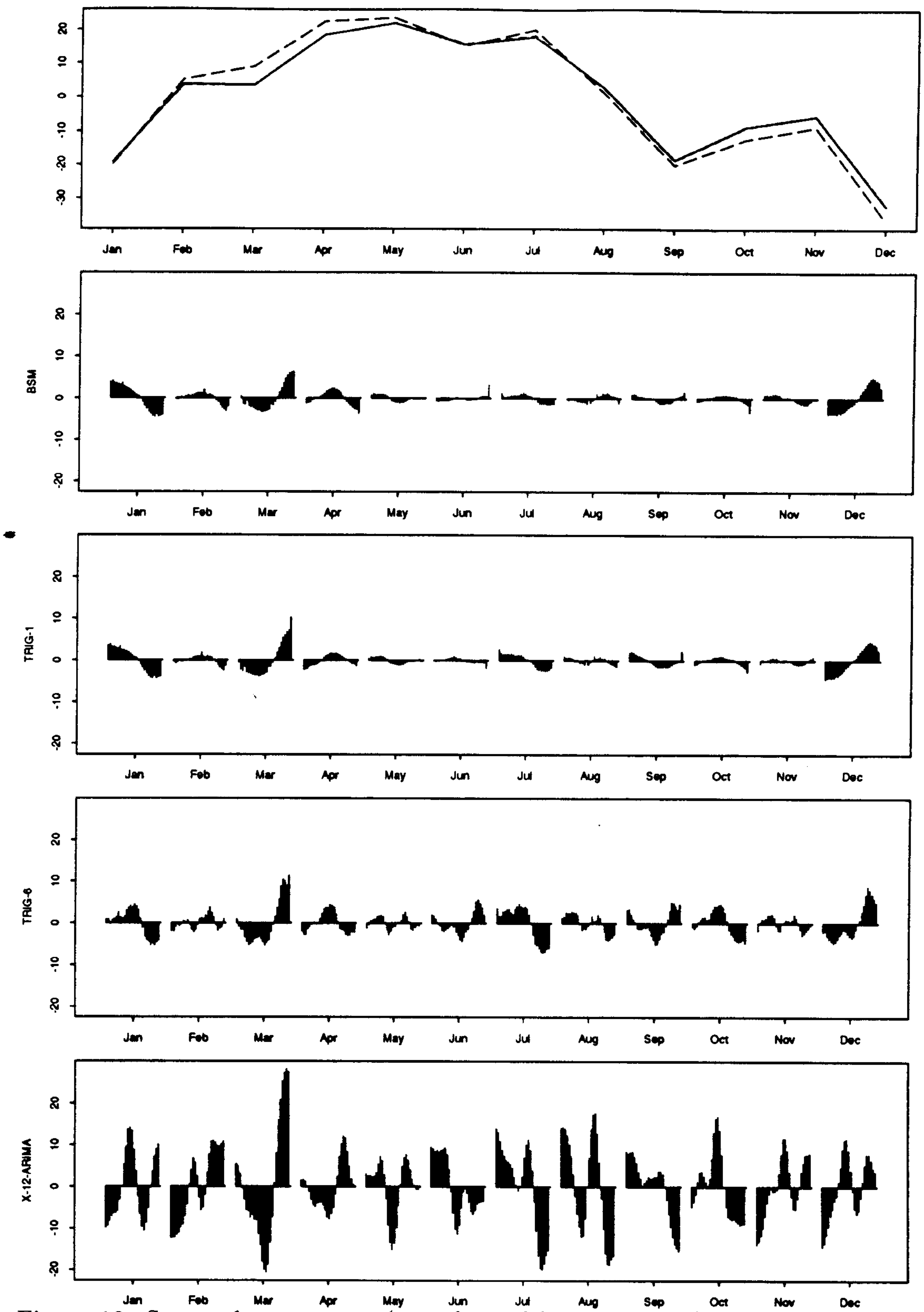


Figure 10: Seasonal component (transformed by 1000 $\times$ log) for IGLCTI. The top plot shows the mean for the BSM (short dashed line), TRIG-1 (medium dashed line), TRIG-6 (long dashed line) and X-12-ARIMA (solid line). Subsequent plots show the deviation from the mean.

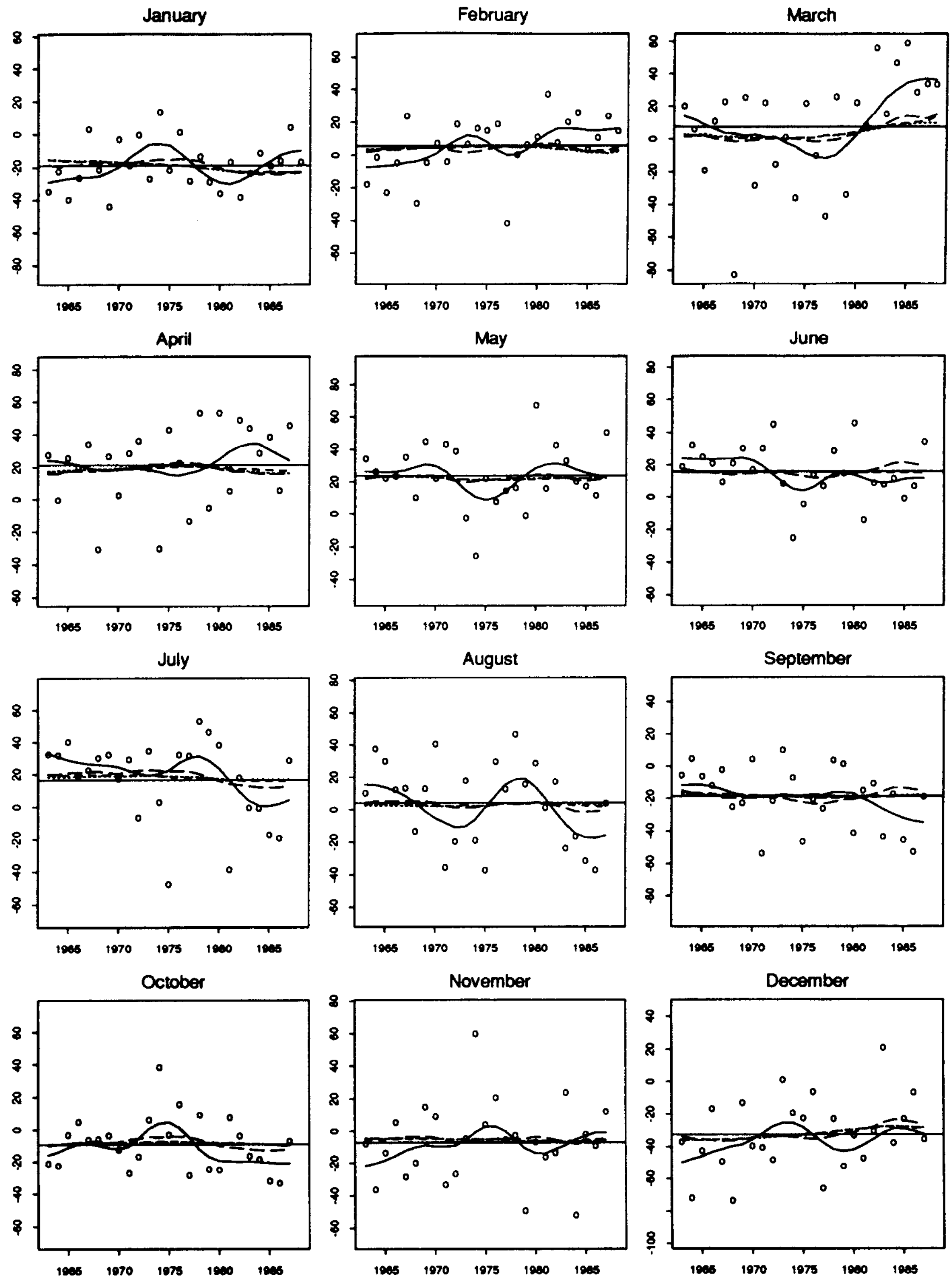


Figure 11: SSI plots for IGLCTI. The procedures compared include the BSM (short dashed line), TRIG-1 (medium dashed line), TRIG-6 (long dashed line) and X-12-ARIMA (solid line). The horizontal solid line corresponds to the mean.

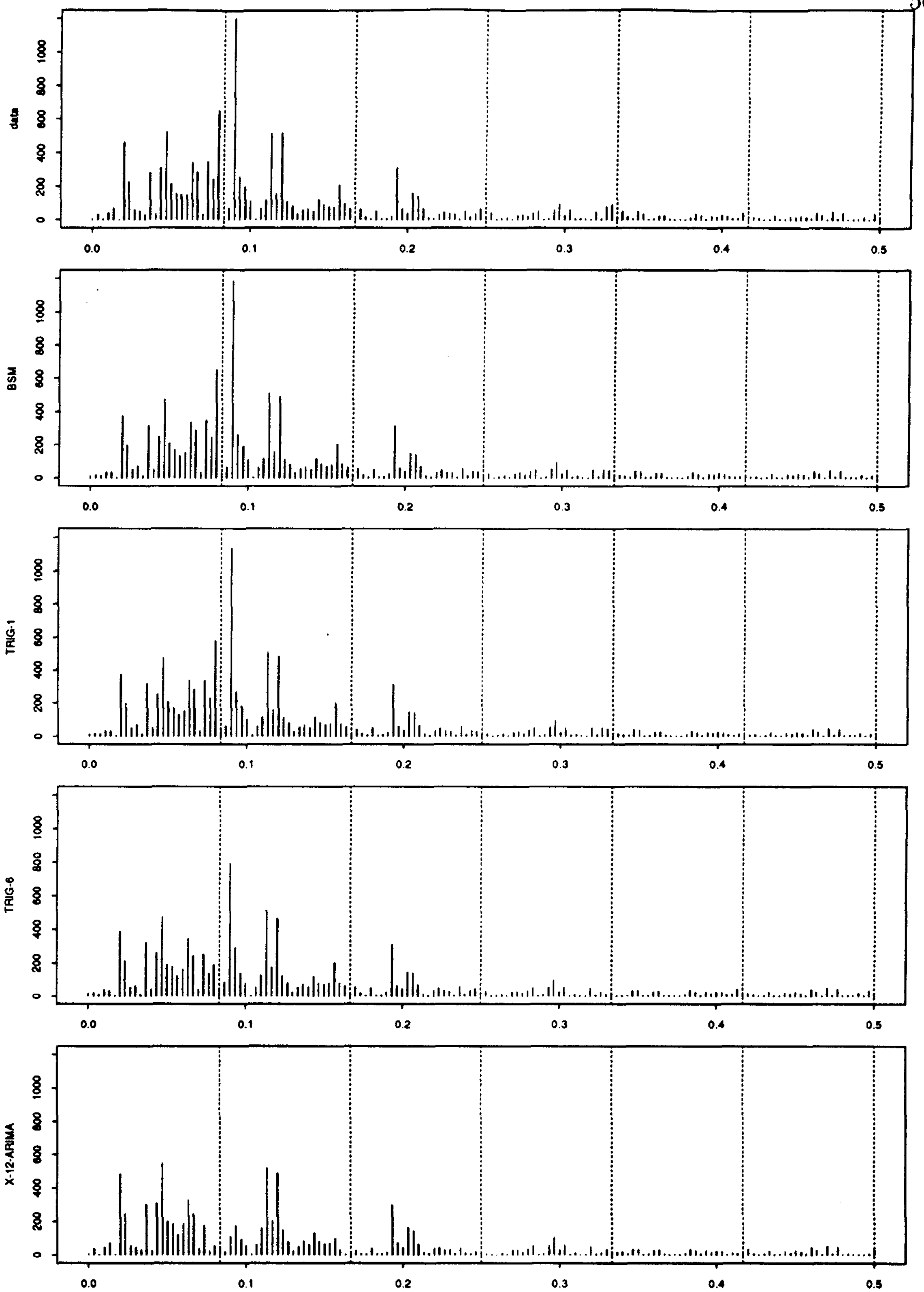


Figure 12: The top plot is the periodogram of the detrended data for IGLCTI (with the fundamental and its harmonics suppressed). The other plots give the periodograms for the detrended seasonally adjusted data for BSM, TRIG-1, TRIG-6, and X-12-ARIMA.

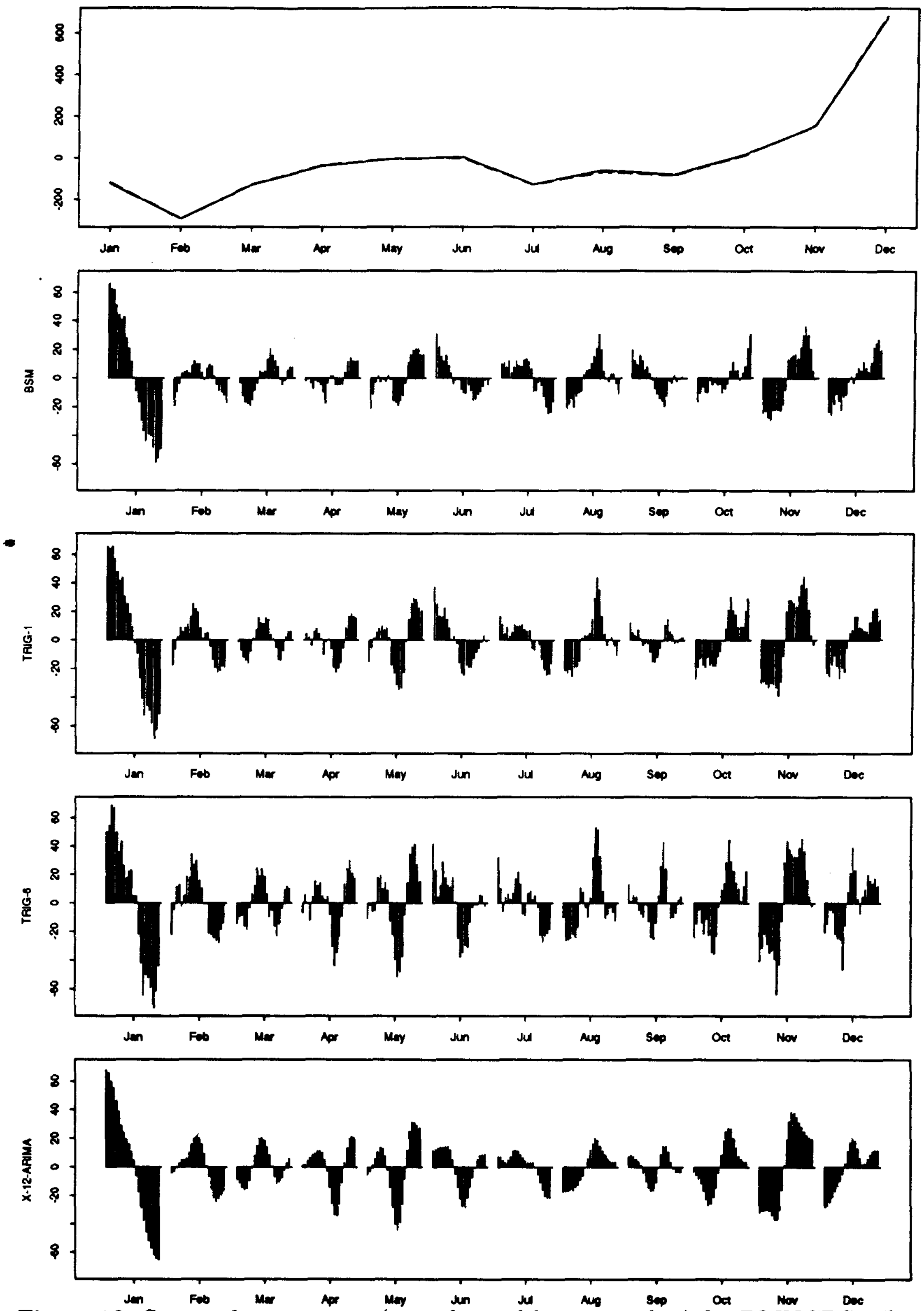


Figure 13: Seasonal component (transformed by 1000 × log) for BMNCRS. The top plot shows the mean for the BSM (short dashed line), TRIG-1 (medium dashed line), TRIG-6 (long dashed line) and X-12-ARIMA (solid line). Subsequent plots show the deviation from the mean.

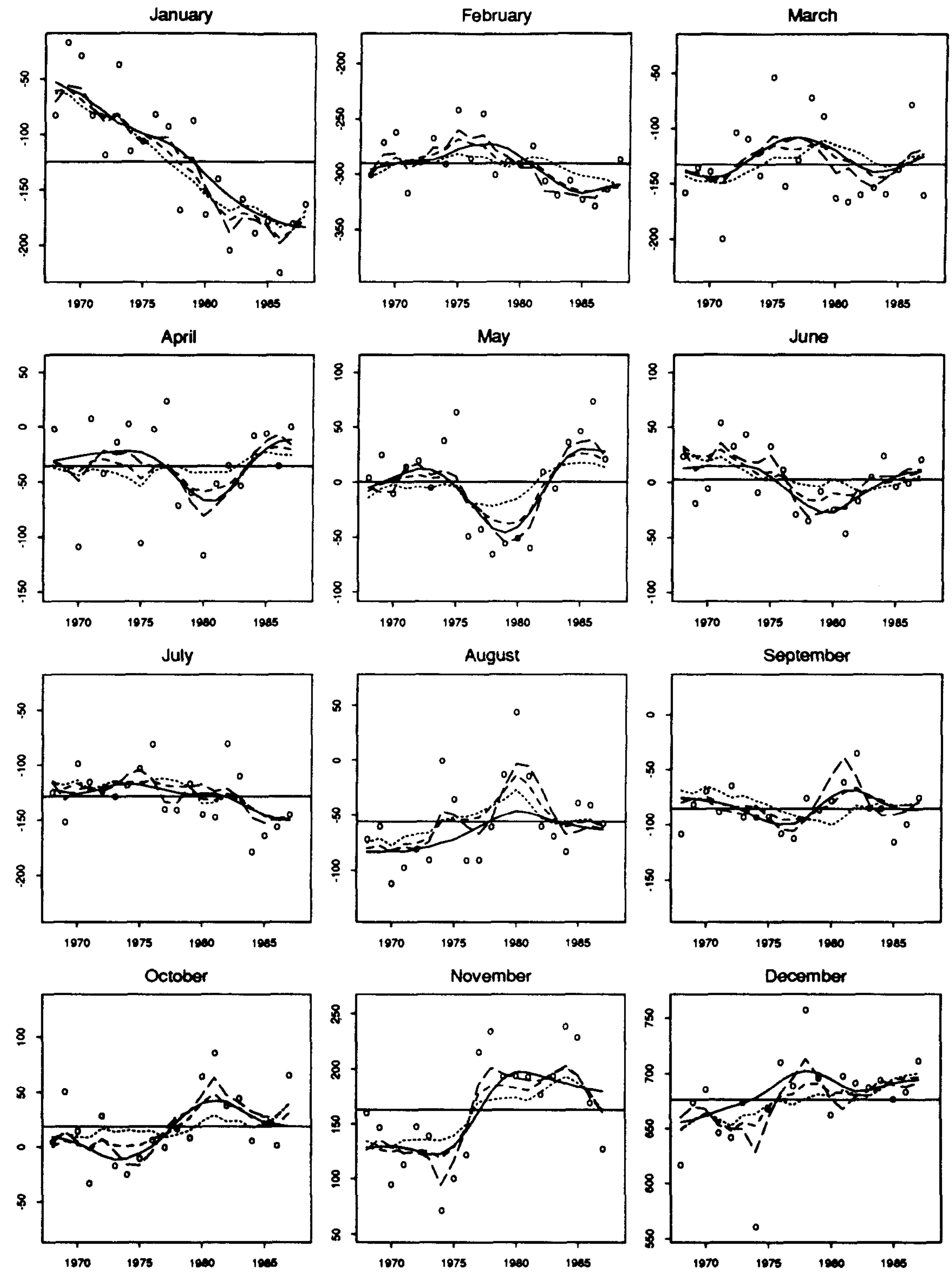


Figure 14: SSI plots for BMNCRS. The procedures compared include the BSM (short dashed line), TRIG-1 (medium dashed line), TRIG-6 (long dashed line) and X-12-ARIMA (solid line). The horizontal solid line corresponds to the mean.

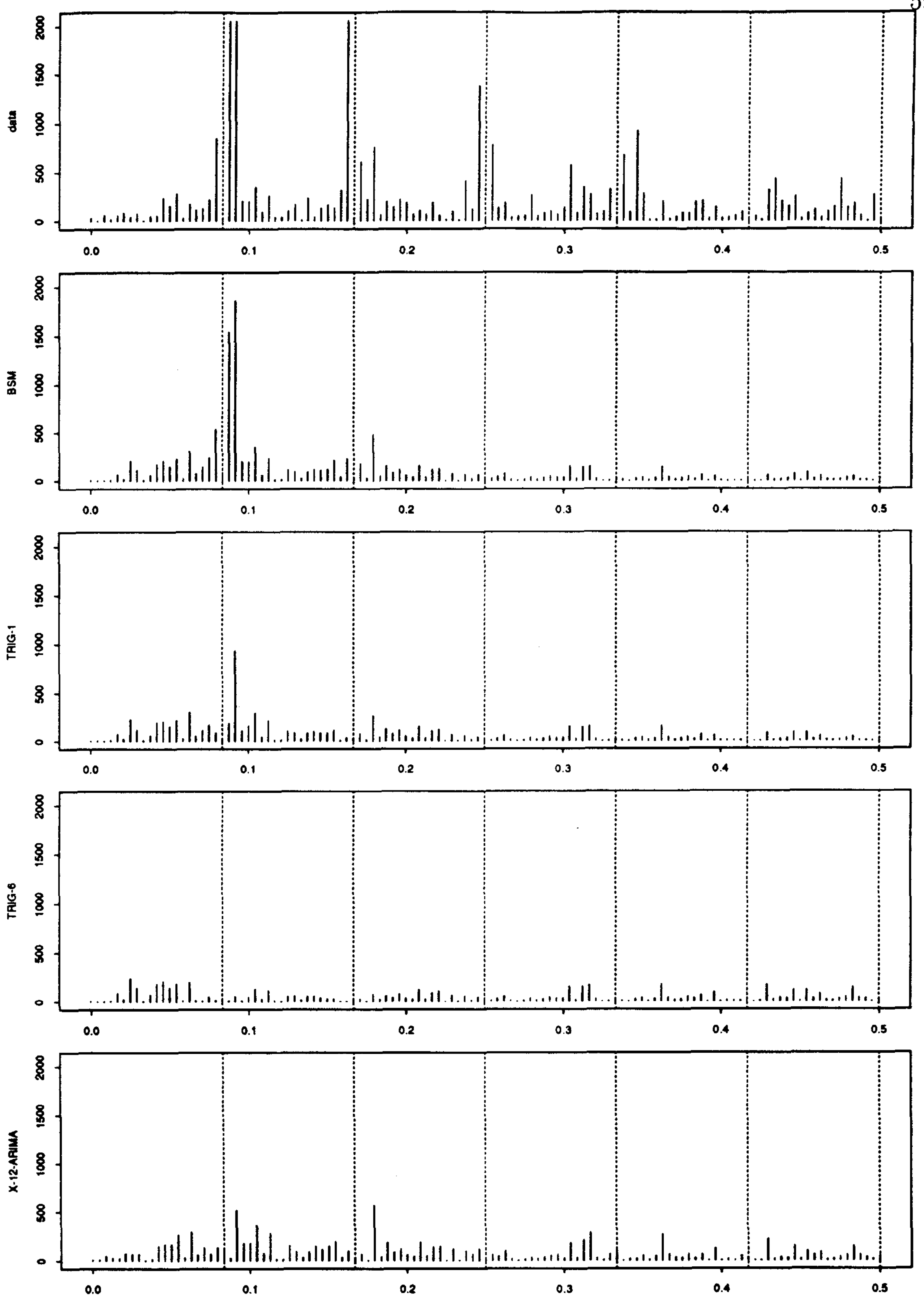


Figure 15: The top plot is the periodogram of the detrended data for BMNCRS (with the fundamental and its harmonics suppressed). The other plots give the periodograms for the detrended seasonally adjusted data for BSM, TRIG-1, TRIG-6, and X-12-ARIMA.

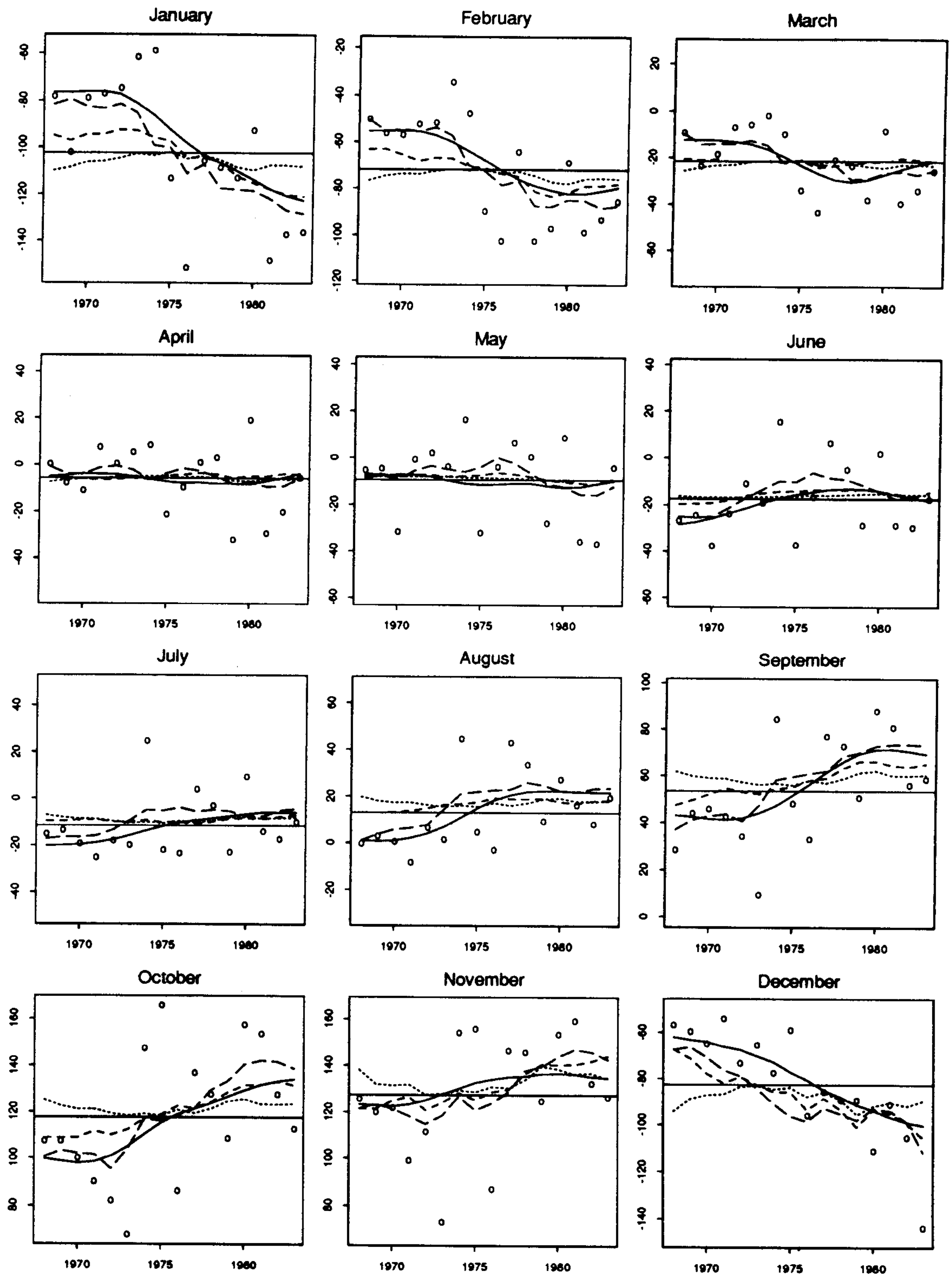


Figure 16: SSI plots for BGMRRI. The procedures compared include the BSM (short dashed line), TRIG-1 (medium dashed line), TRIG-6 (long dashed line) and X-12-ARIMA (solid line). The horizontal solid line corresponds to the mean.

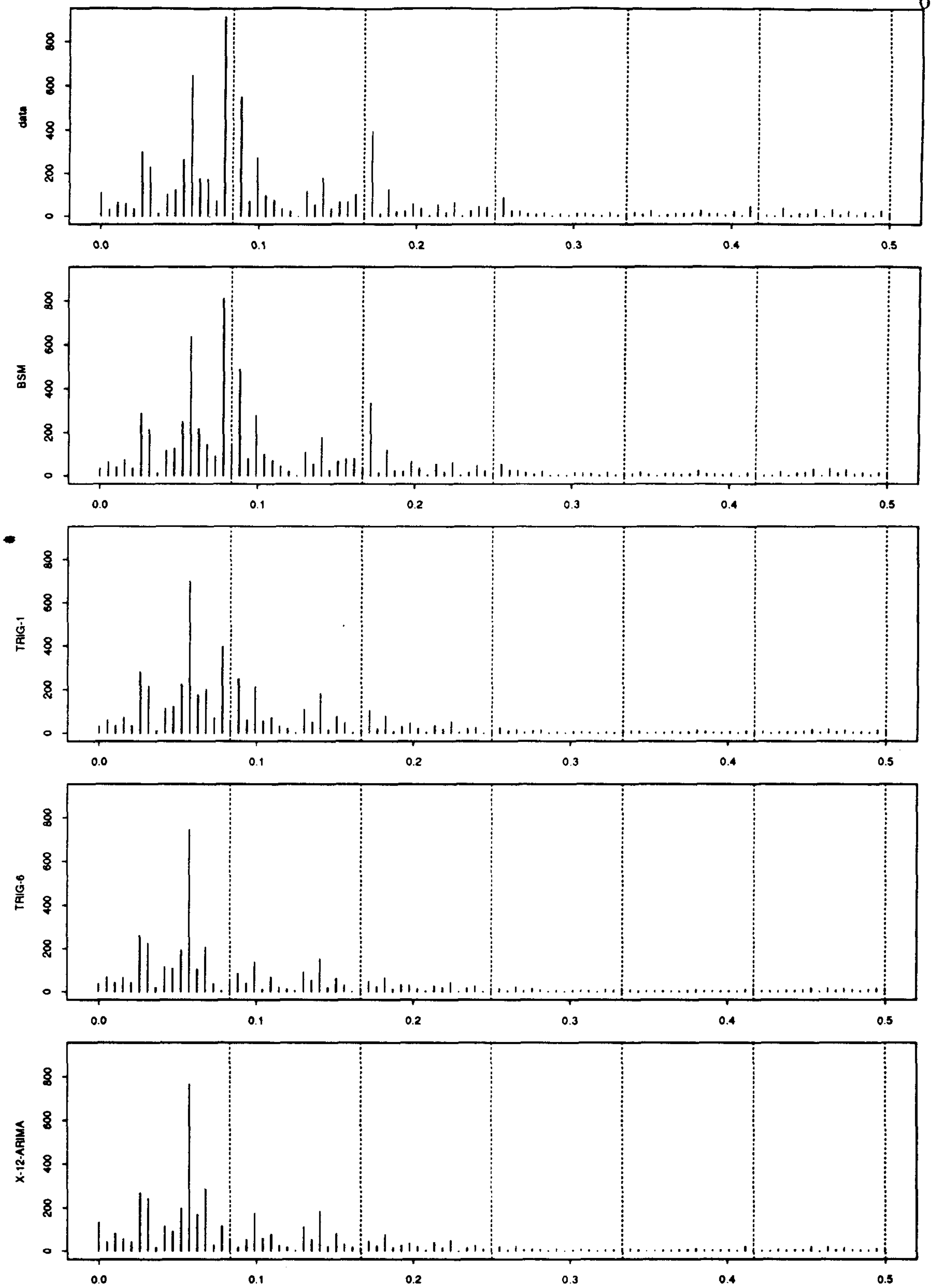


Figure 17: The top plot is the periodogram of the detrended data for BGMRRI (with the fundamental and its harmonics suppressed). The other plots give the periodograms for the detrended seasonally adjusted data for BSM, TRIG-1, TRIG-6, and X-12-ARIMA.

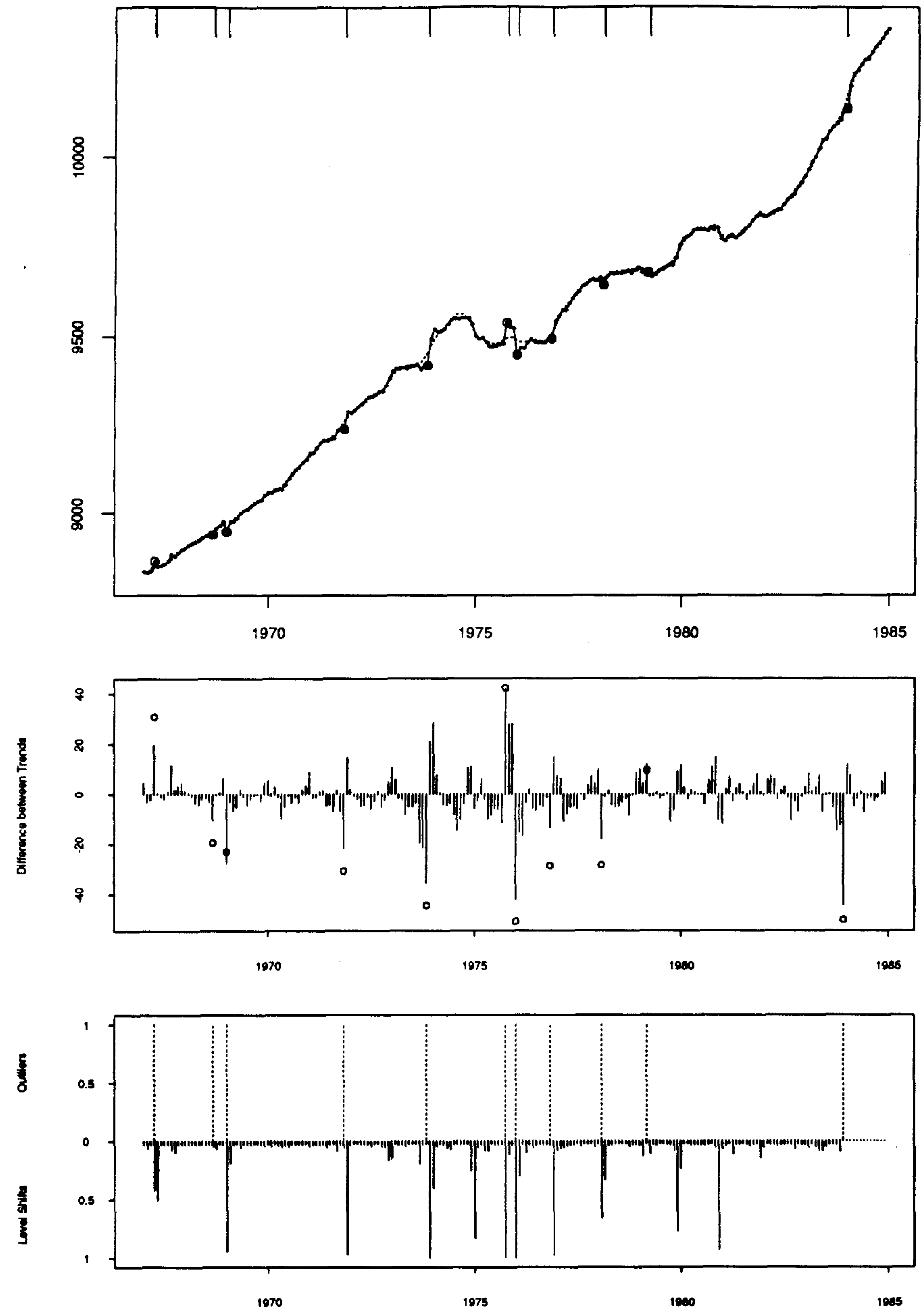


Figure 18: The top plot compares the trends for TRIG-6 and X-12-ARIMA for BGMRRI. The middle plot displays the ratio of the trends. The bottom plot gives the posterior probability of AO's and LS's for TRIG-6 (solid lines) the X-12-ARIMA (dashed lines - either 0 or 1).

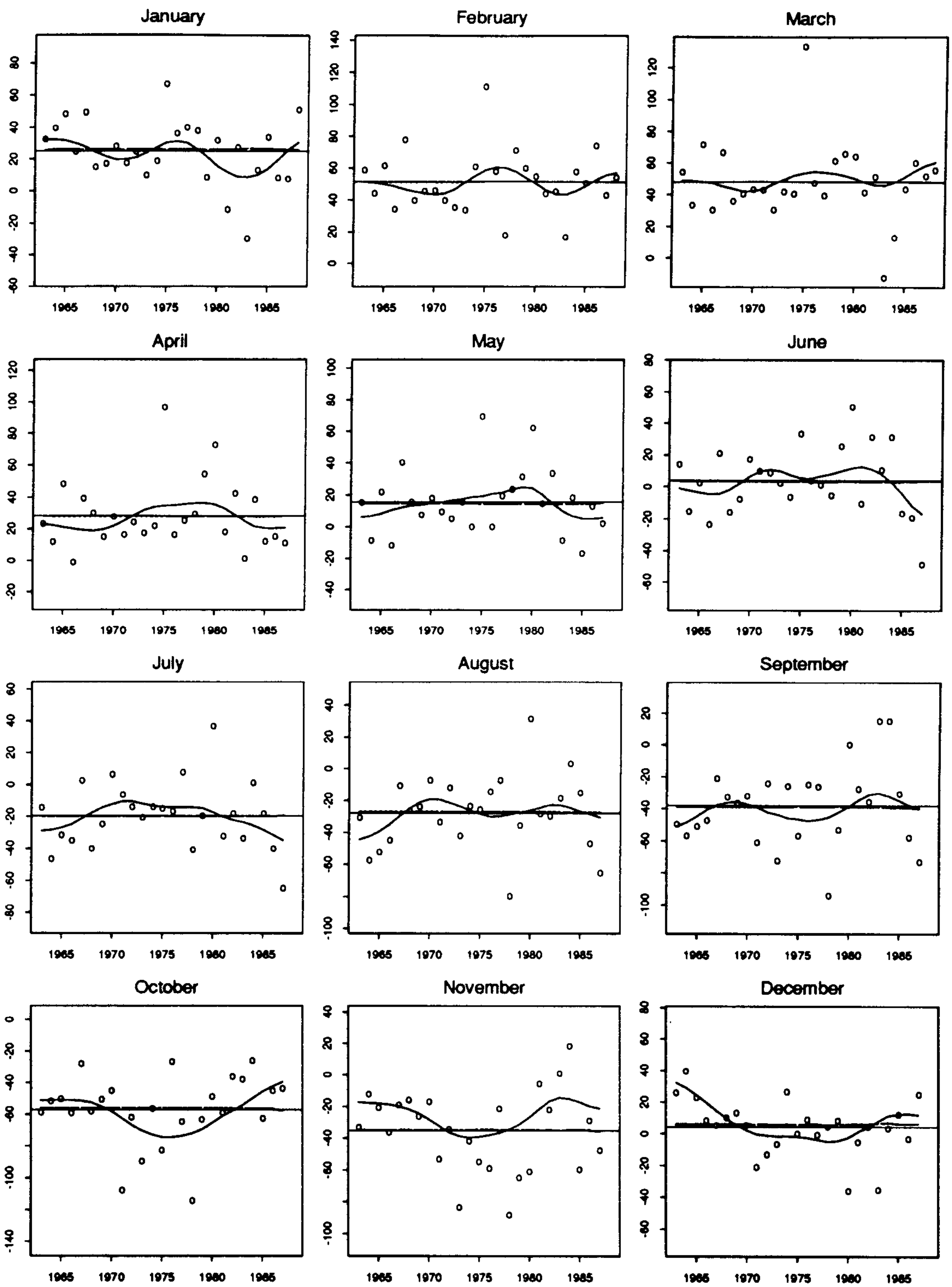


Figure 19: SSI plots for IFMETI. The procedures compared include the BSM (short dashed line), TRIG-1 (medium dashed line), TRIG-6 (long dashed line) and X-12-ARIMA (solid line). The horizontal solid line corresponds to the mean.

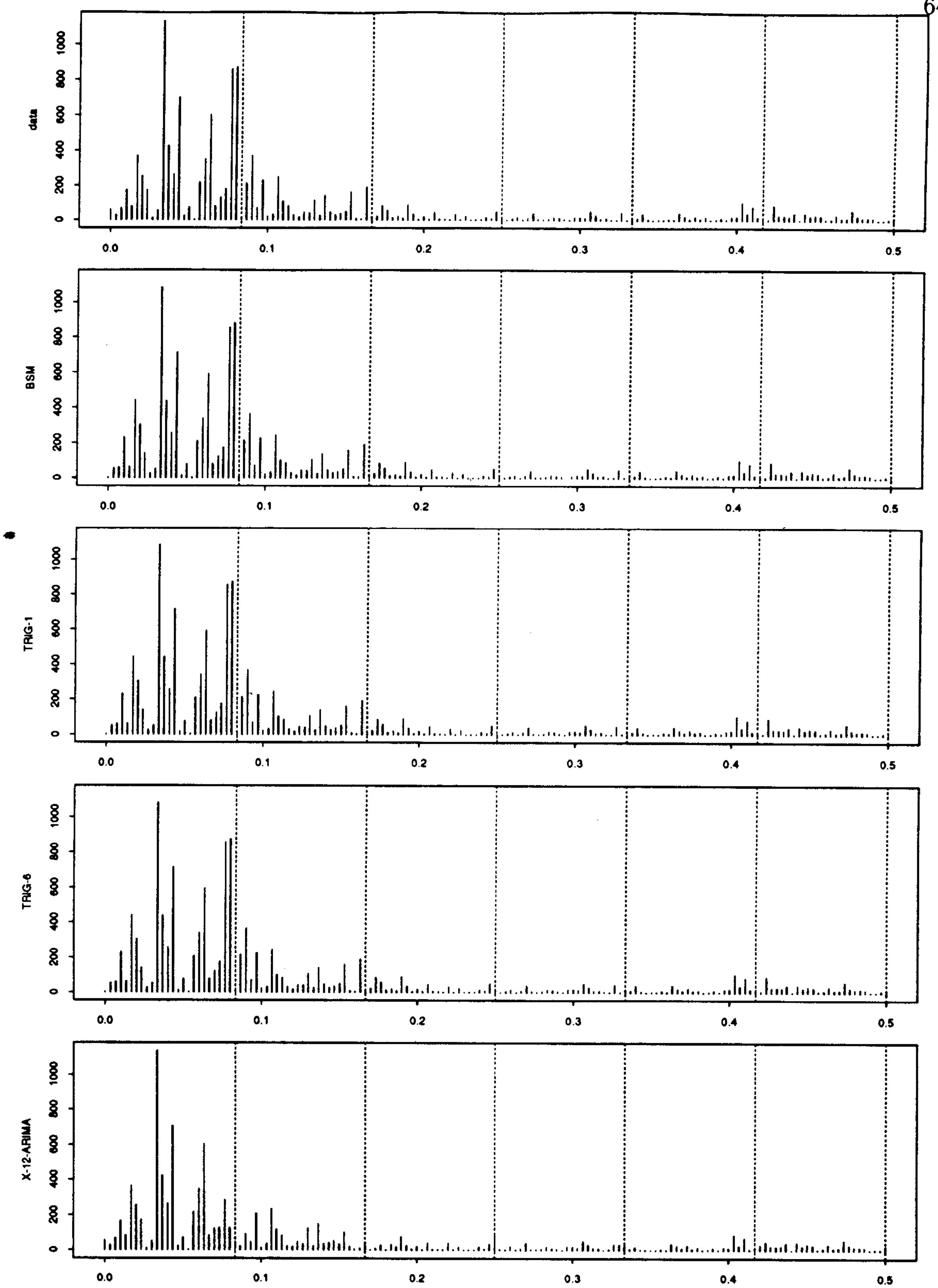


Figure 20: The top plot is the periodogram of the detrended data for IFMETI (with the fundamental and its harmonics suppressed). The other plots give the periodograms for the detrended seasonally adjusted data for BSM, TRIG-1, TRIG-6, and X-12-ARIMA.

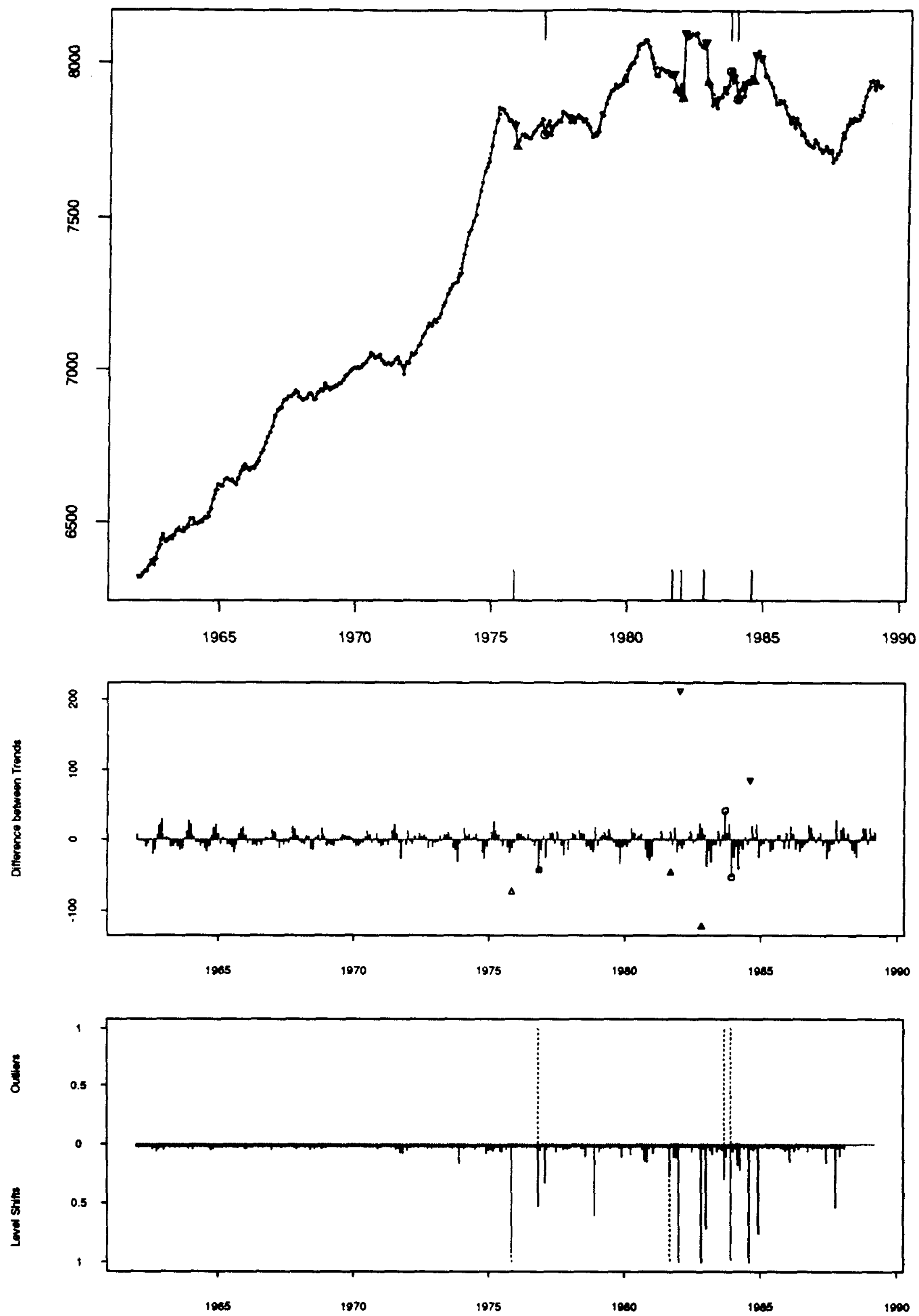


Figure 21: The top plot compares the trends for TRIG-6 and X-12-ARIMA for IFMETI. The middle plot displays the ratio of the trends. The bottom plot gives the posterior probability of AO's and LS's for TRIG-6 (solid lines) the X-12-ARIMA (dashed lines - either 0 or 1).

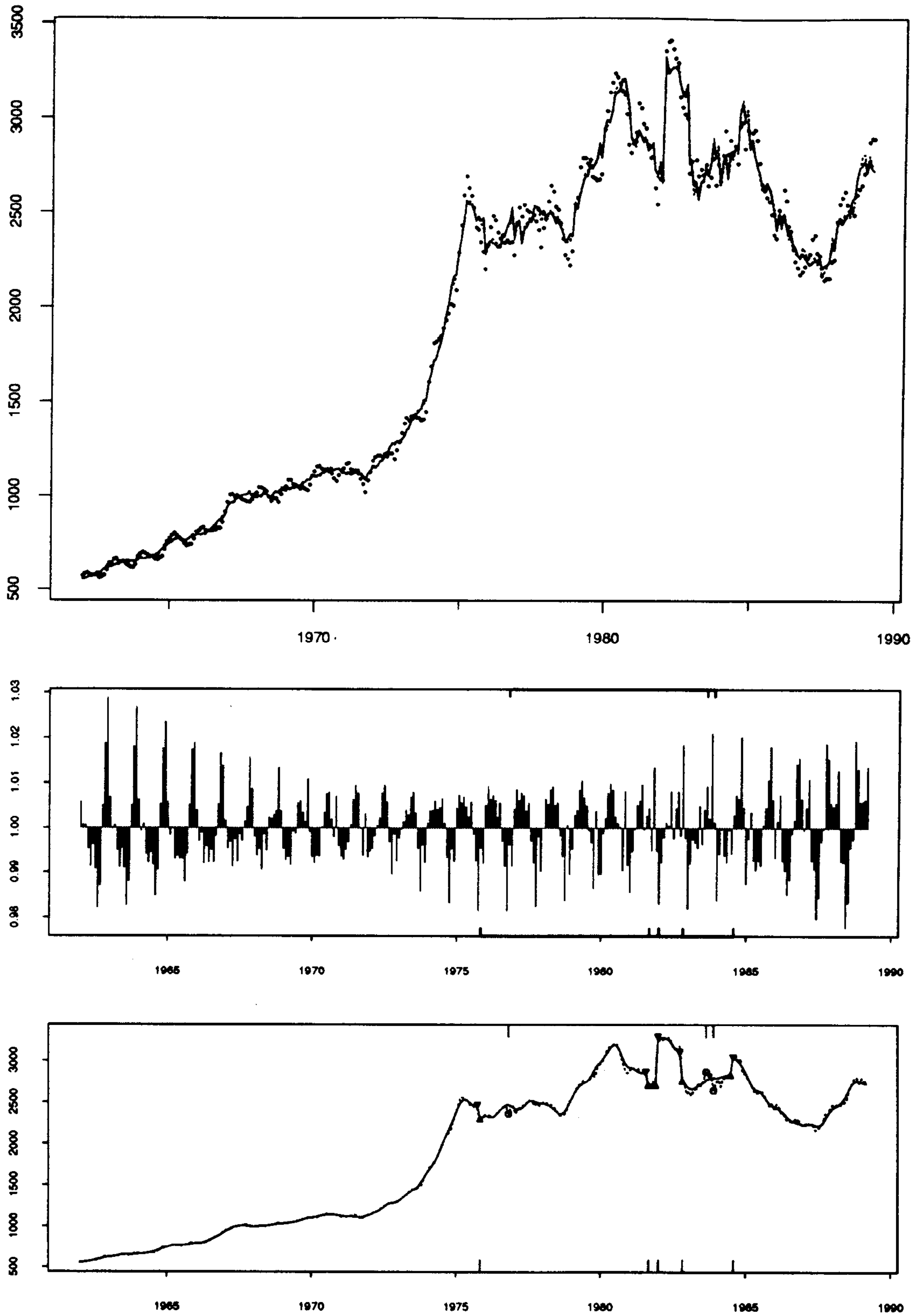


Figure 22: The top plot shows the data for IFMETI (points), the TRIG-6 seasonally adjusted data (dashed line), and X-12-ARIMA seasonally adjusted data (solid line). The middle plot shows the ratio of the seasonally adjusted data and the bottom plot compares the trends.

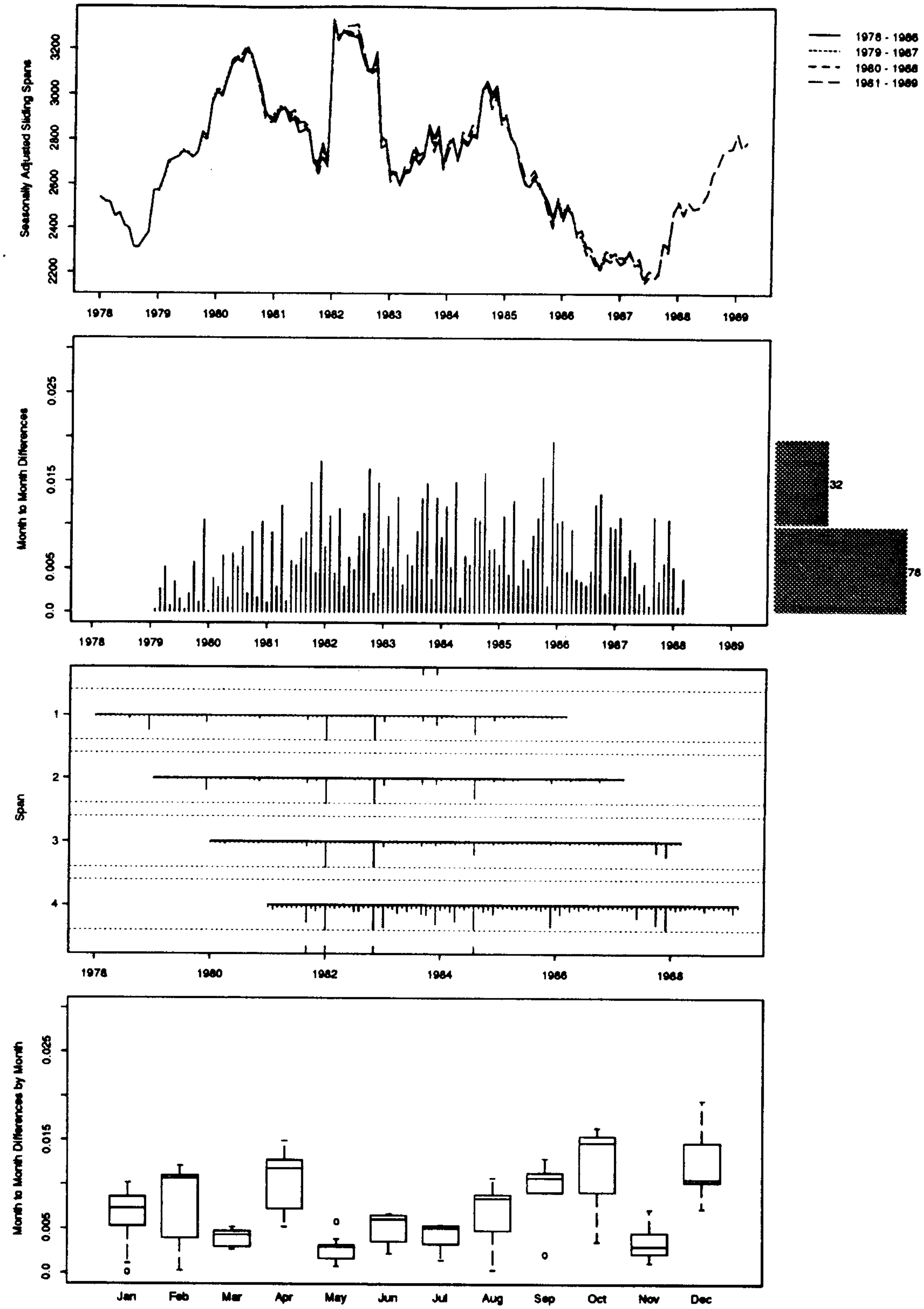


Figure 23: BSM sliding spans for IFMETI. The top plot shows the seasonally adjusted data for the four spans The second and fourth plots display the statistic $MM_t^{\max}$. The third plot compares the outlier treatments over the spans.

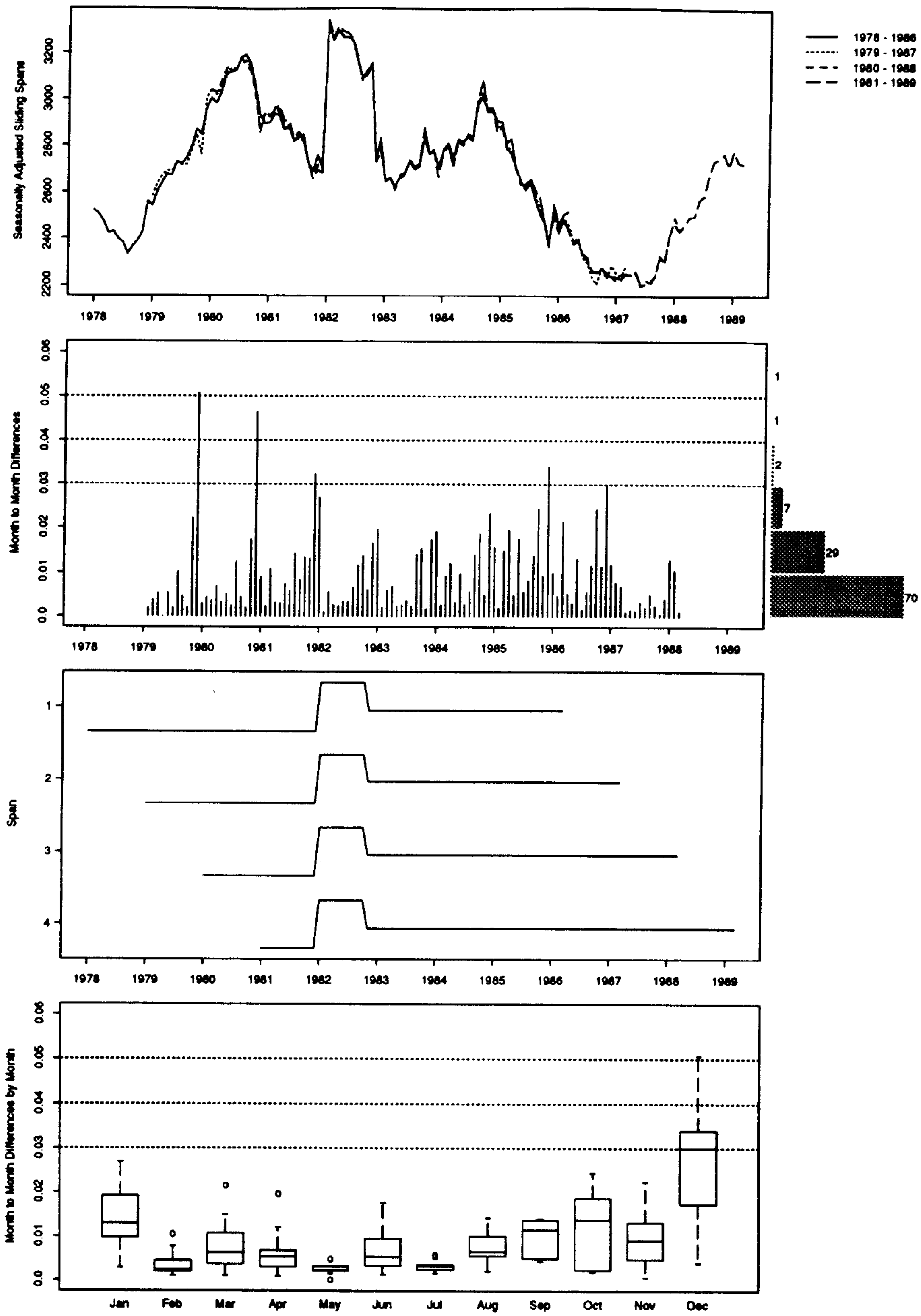


Figure 24: X-12-ARIMA sliding spans for IFMETI.

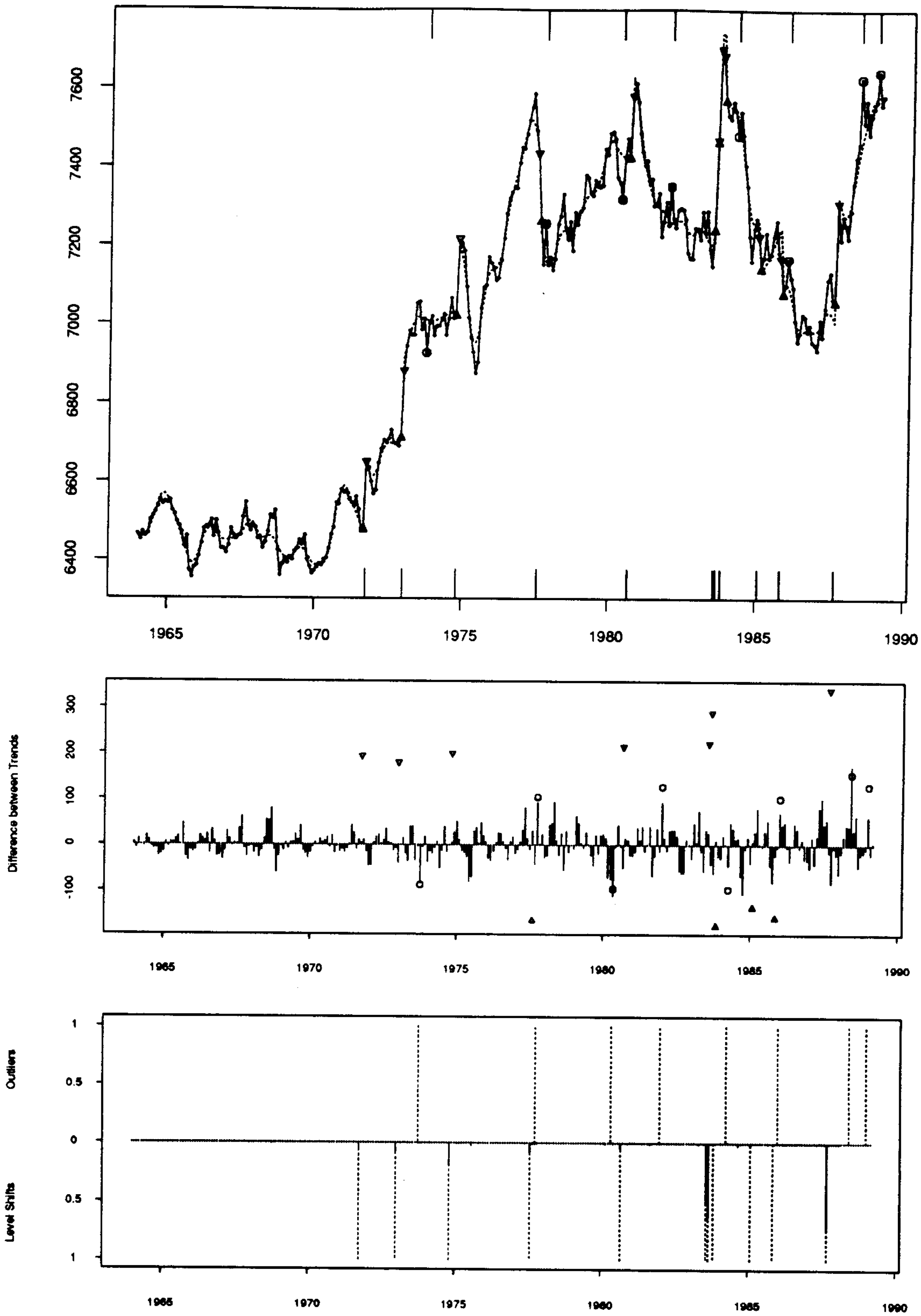


Figure 25: The top plot compares the trends for TRIG-6 and X-12-ARIMA for IFATTI. The middle plot displays the ratio of the trends. The bottom plot gives the posterior probability of AO's and LS's for TRIG-6 (solid lines) the X-12-ARIMA (dashed lines - either 0 or 1).

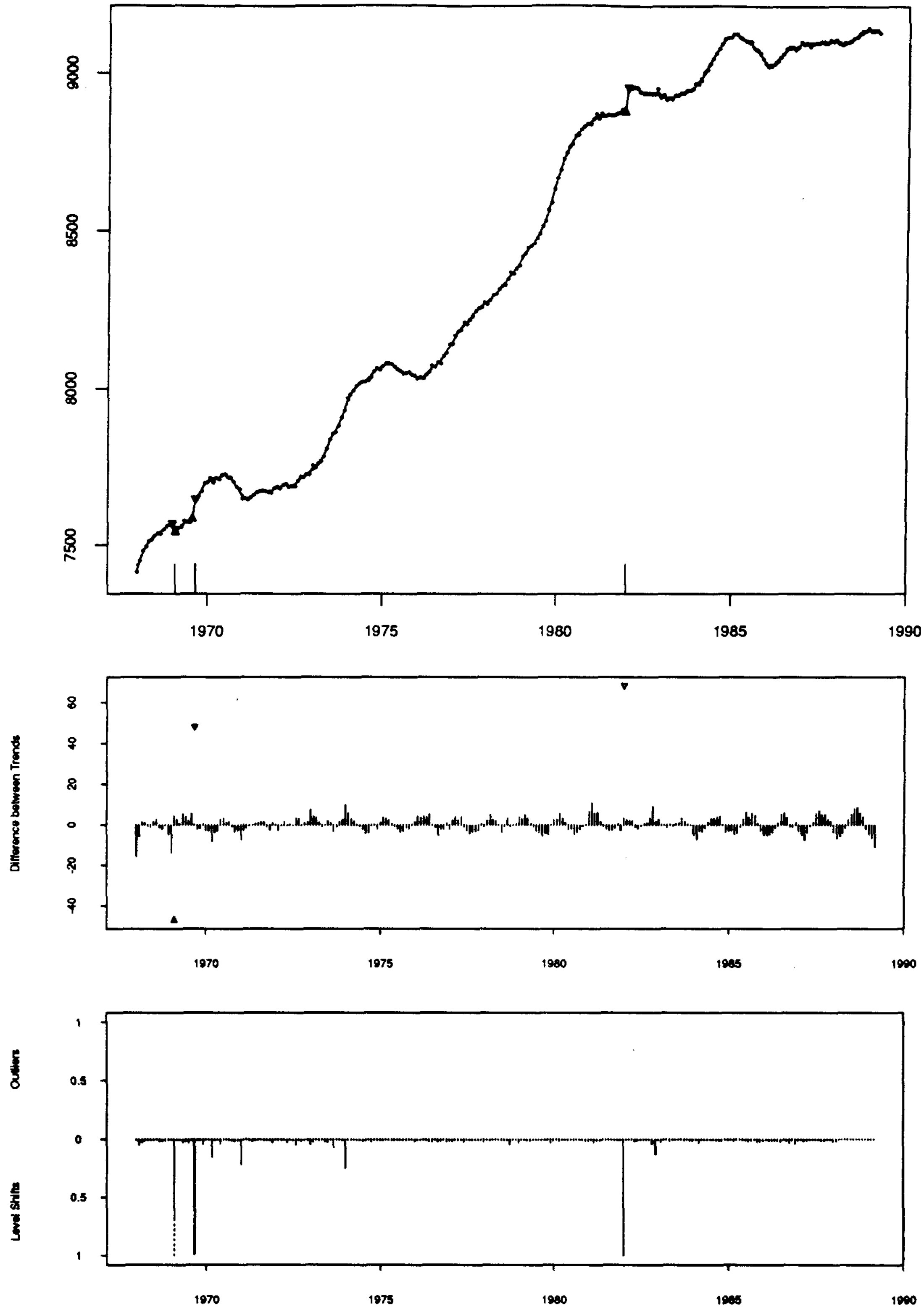


Figure 26: The top plot compares the trends for TRIG-6 and X-12-ARIMA for ICMETI. The middle plot displays the ratio of the trends. The bottom plot gives the posterior probability of AO's and LS's for TRIG-6 (solid lines) the X-12-ARIMA (dashed lines - either 0 or 1).

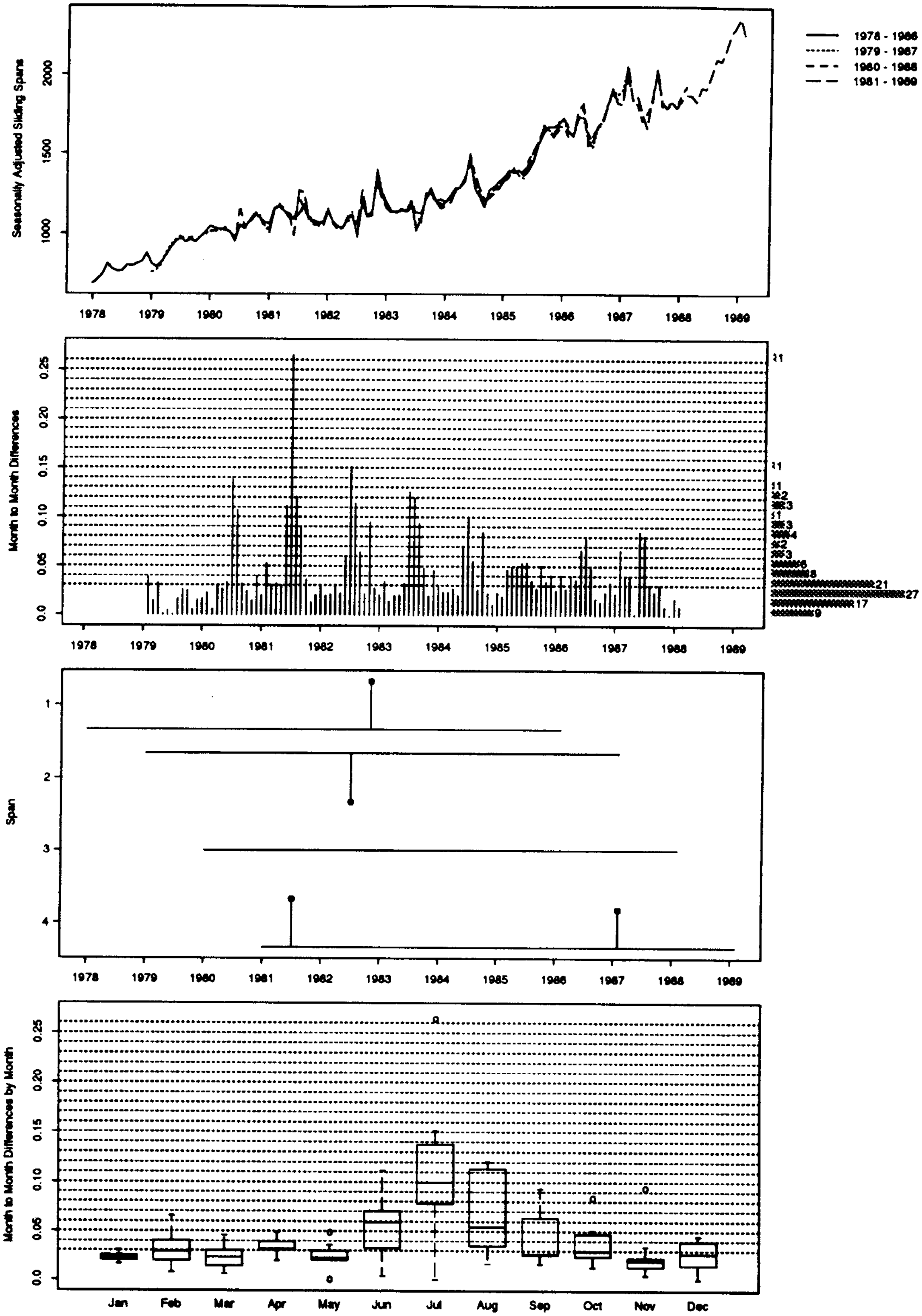


Figure 27: X-12-ARIMA sliding spans for BSPGWS. The top plot shows the seasonally adjusted data for the four spans The second and fourth plots display the statistic $MM_t^{\max}$. The third plot compares the outlier treatments over the spans.

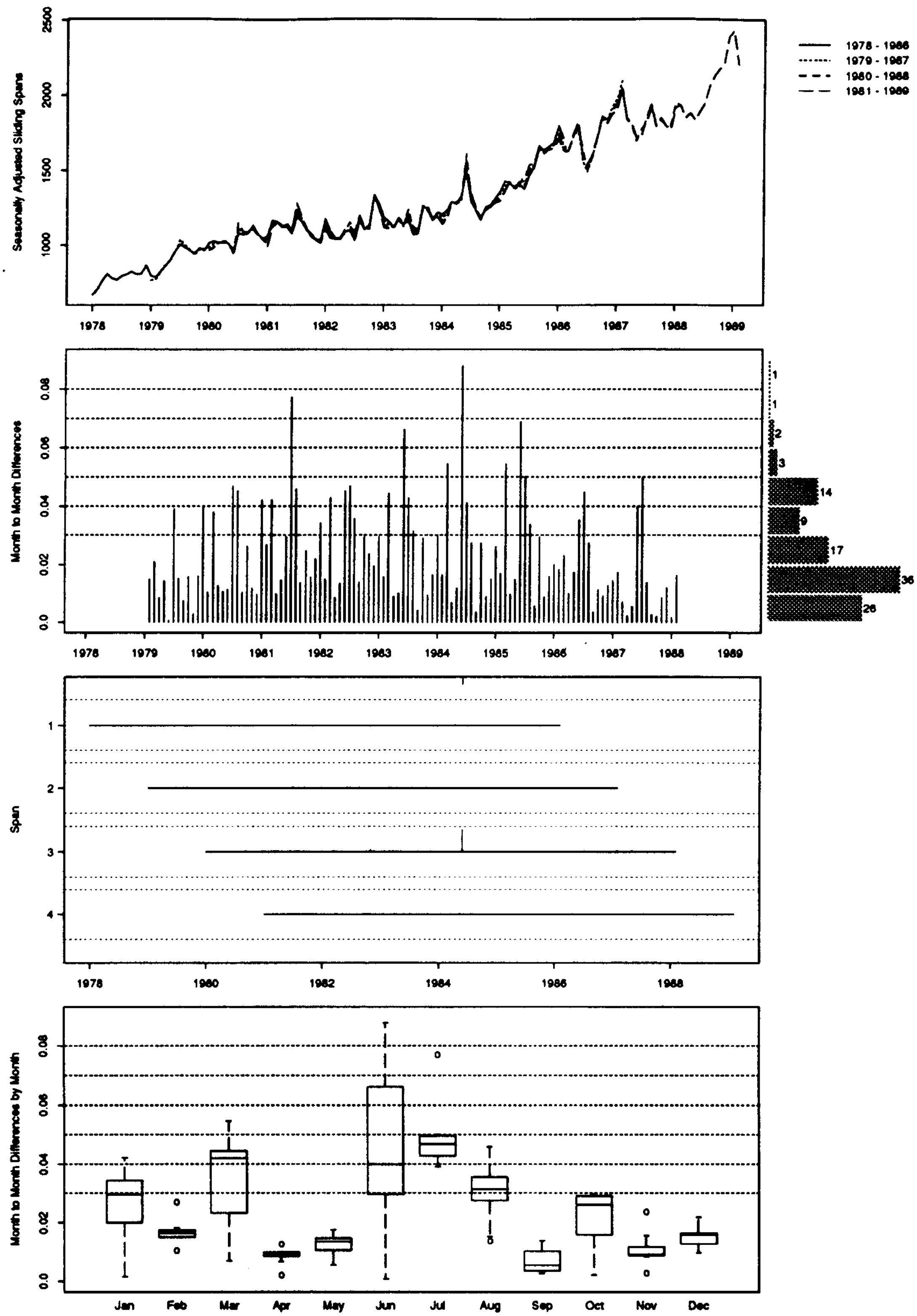


Figure 28: BSM sliding spans for BSPGWS. The top plot shows the seasonally adjusted data for the four spans The second and fourth plots display the statistic $MM_t^{\max}$. The third plot compares the outlier treatments over the spans.

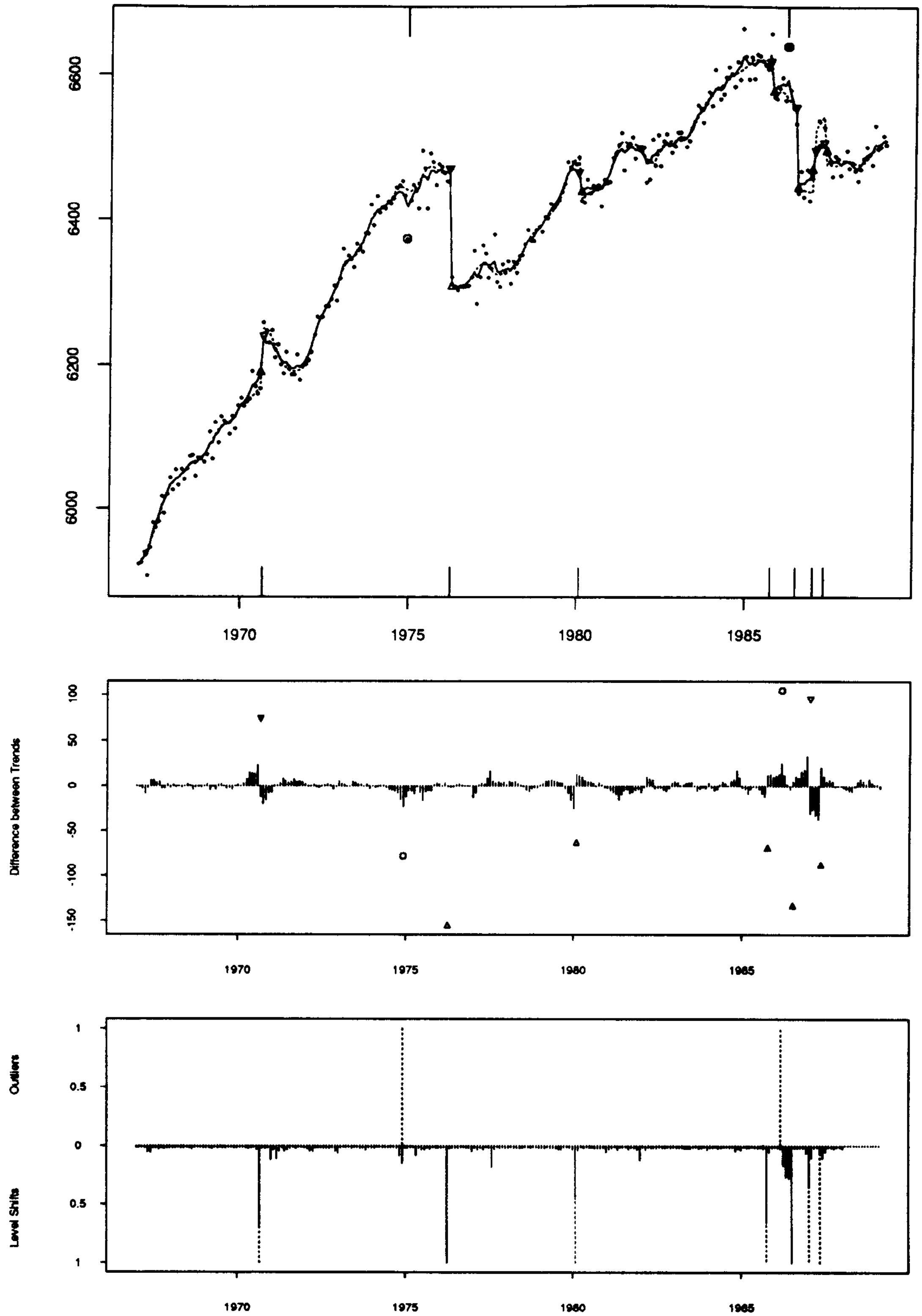


Figure 29: The top plot compares the trends for TRIG-6 and X-12-ARIMA for BVARRS. The middle plot displays the ratio of the trends. The bottom plot gives the posterior probability of AO's and LS's for TRIG-6 (solid lines) the X-12-ARIMA (dashed lines - either 0 or 1).

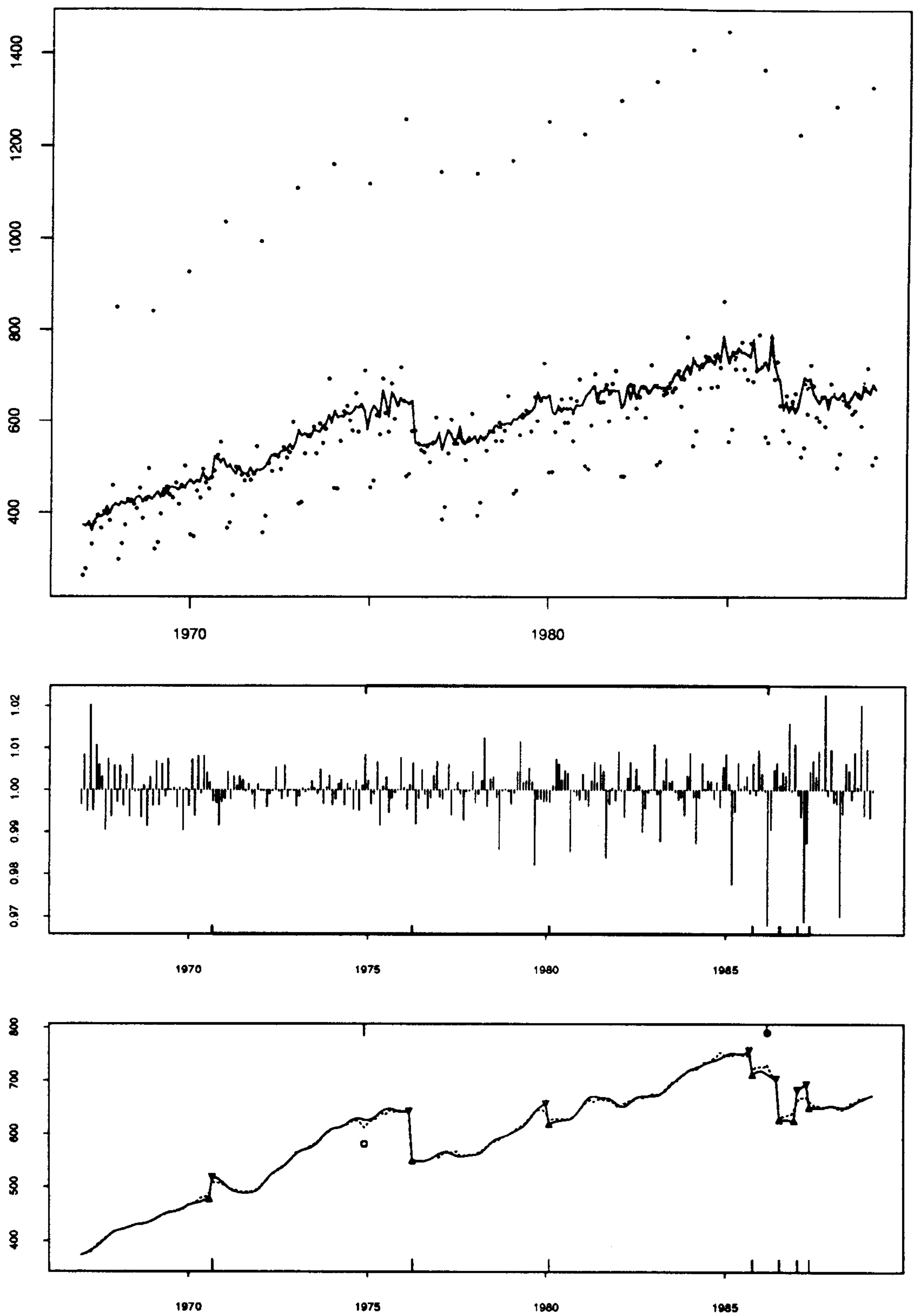


Figure 30: The top plot shows the data for BVARRS (points), the TRIG-6 seasonally adjusted data (dashed line), and X-12-ARIMA seasonally adjusted data (solid line). The middle plot shows the ratio of the seasonally adjusted data and the bottom plot compares the trends.

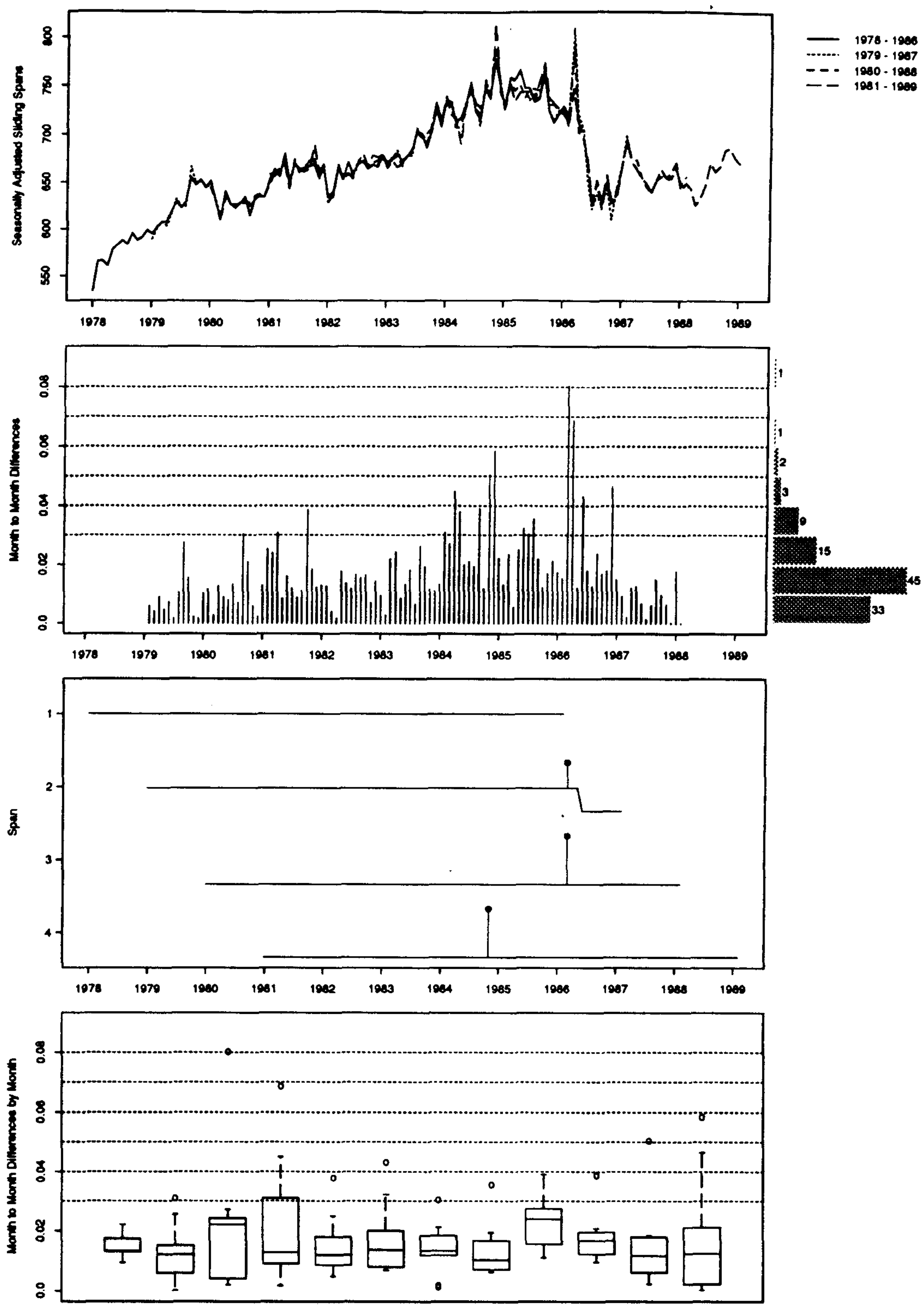


Figure 31: X-12-ARIMA sliding spans for BVARRS. The top plot shows the seasonally adjusted data for the four spans The second and fourth plots display the statistic $MM_t^{\max}$. The third plot compares the outlier treatments over the spans.

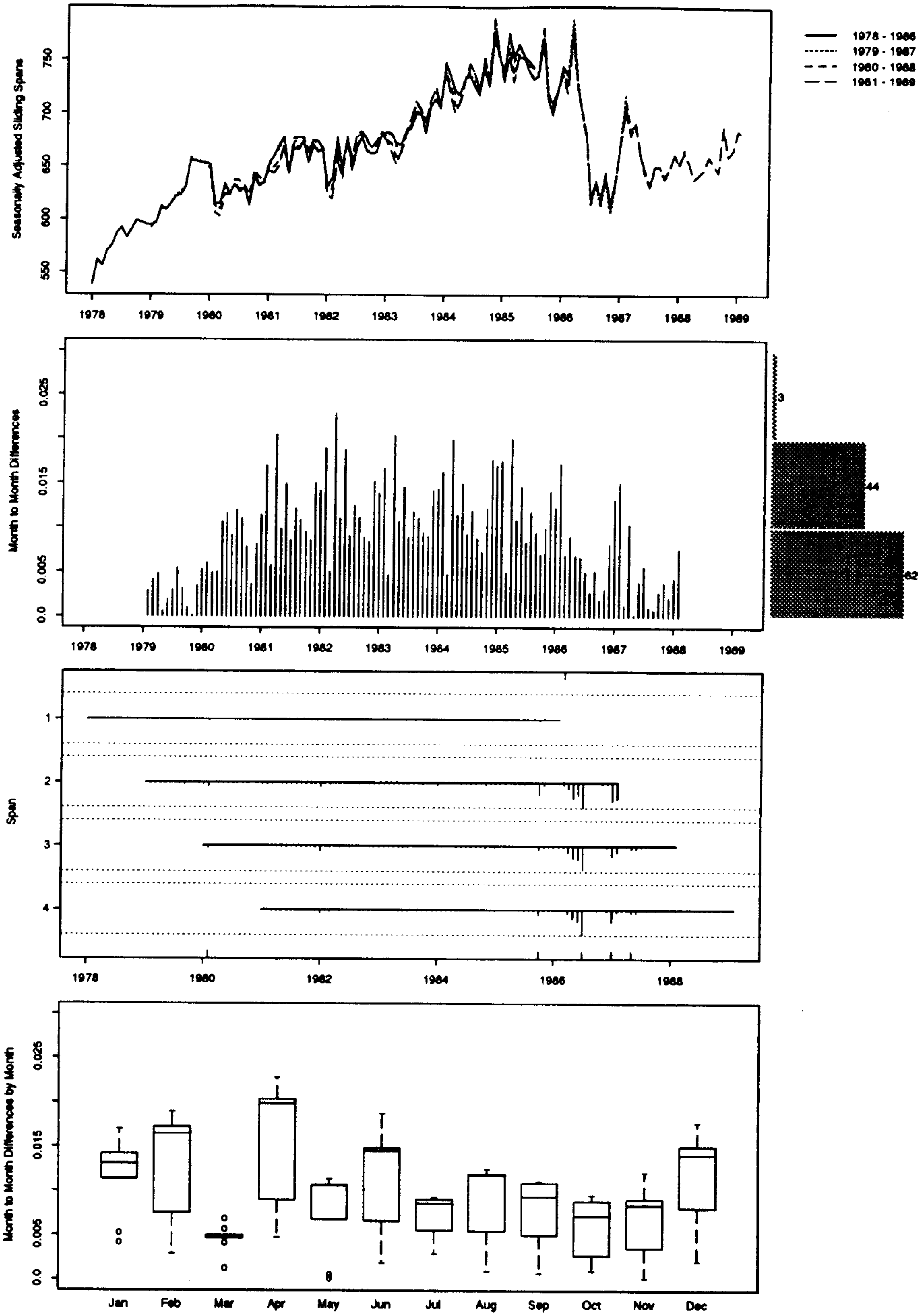


Figure 32: BSM sliding spans for BVARRS. The top plot shows the seasonally adjusted data for the four spans The second and fourth plots display the statistic $MM_t^{\max}$. The third plot compares the outlier treatments over the spans.

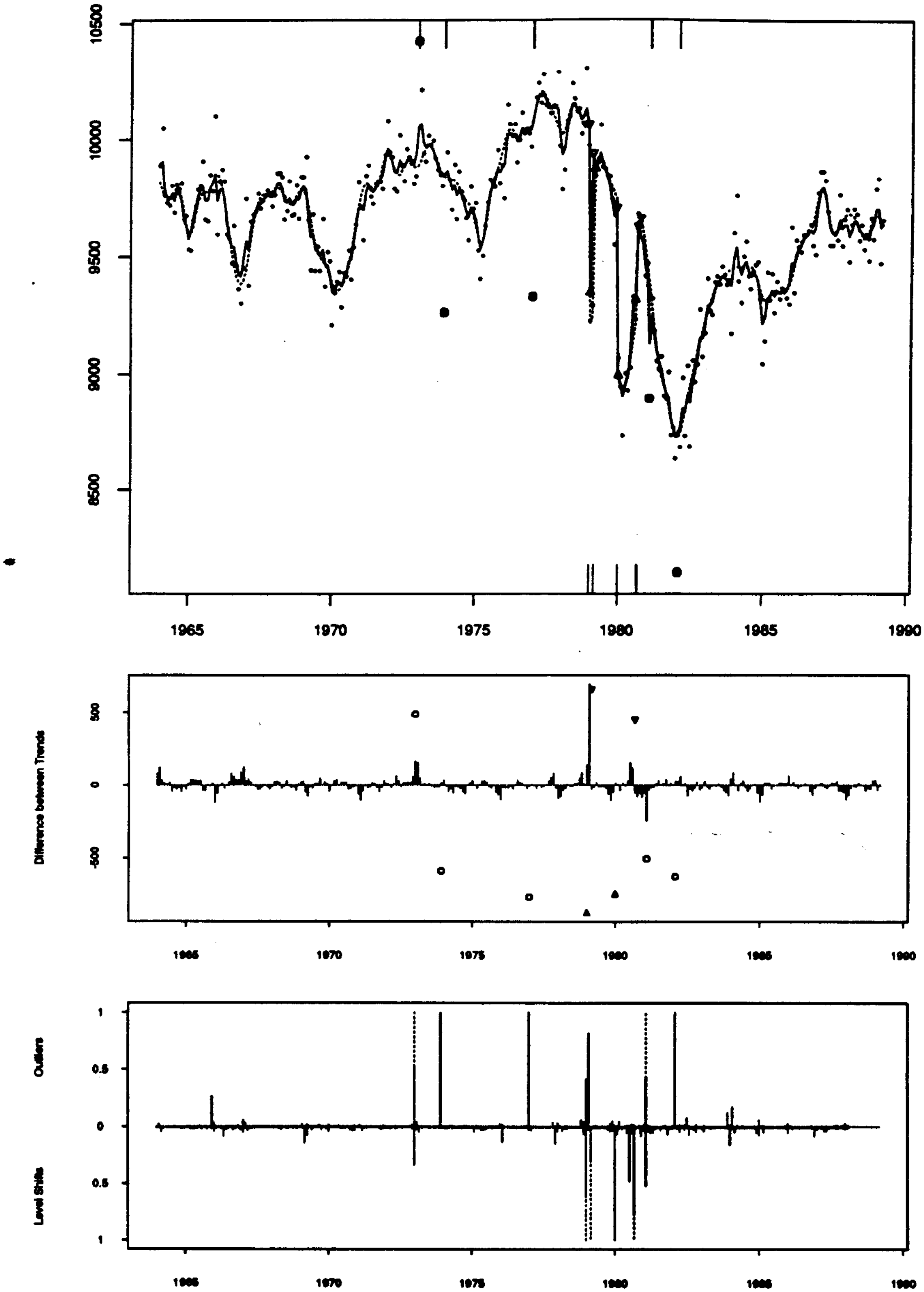


Figure 33: The top plot compares the trends for TRIG-6 and X-12-ARIMA for CMW1HS. The middle plot displays the ratio of the trends. The bottom plot gives the posterior probability of AO's and LS's for TRIG-6 (solid lines) the X-12-ARIMA (dashed lines - either 0 or 1).

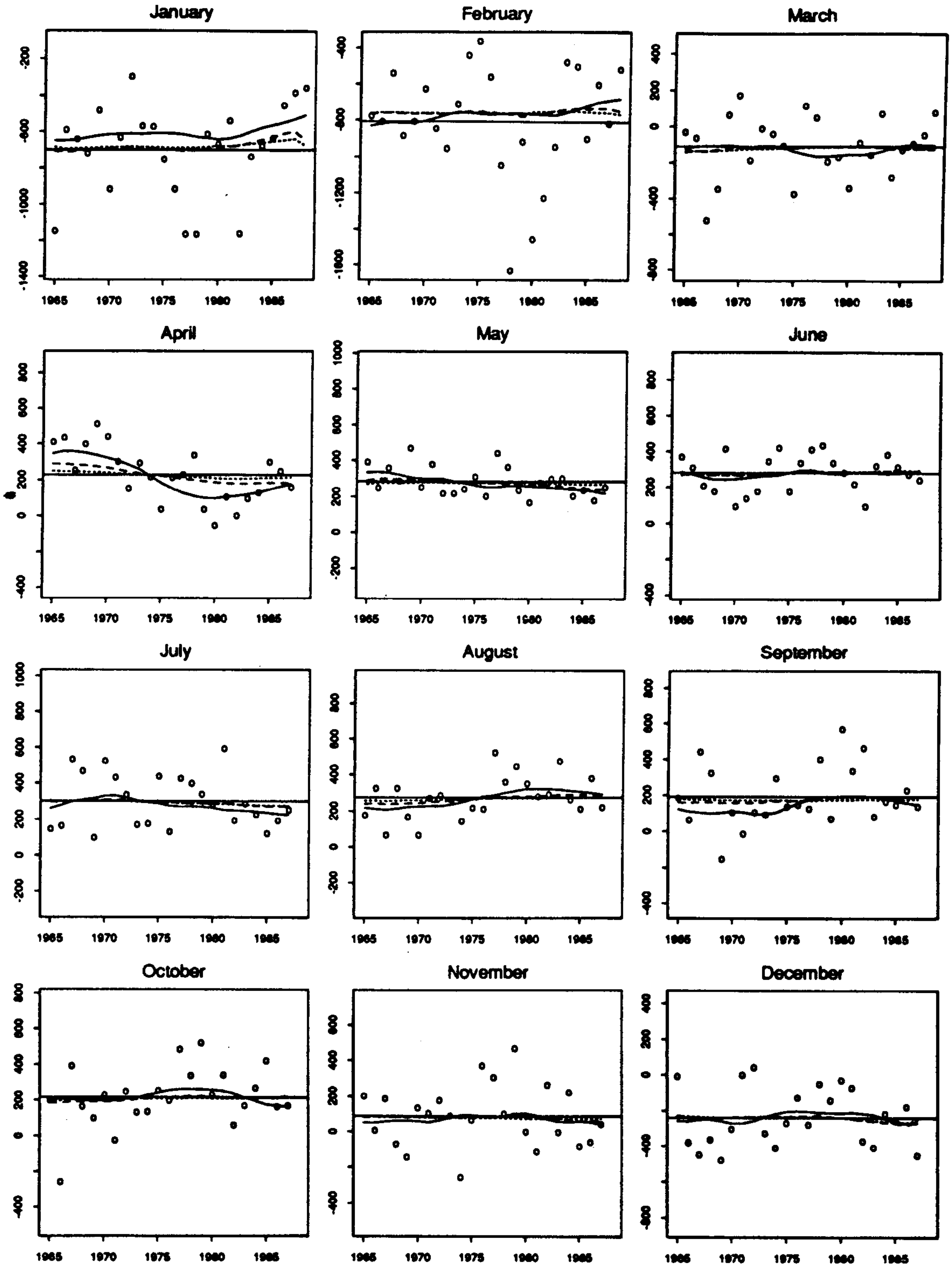


Figure 34: SSI plots for CNETHS. The procedures compared include the BSM (short dashed line), TRIG-1 (medium dashed line), TRIG-6 (long dashed line) and X-12-ARIMA (solid line). The horizontal solid line corresponds to the mean.

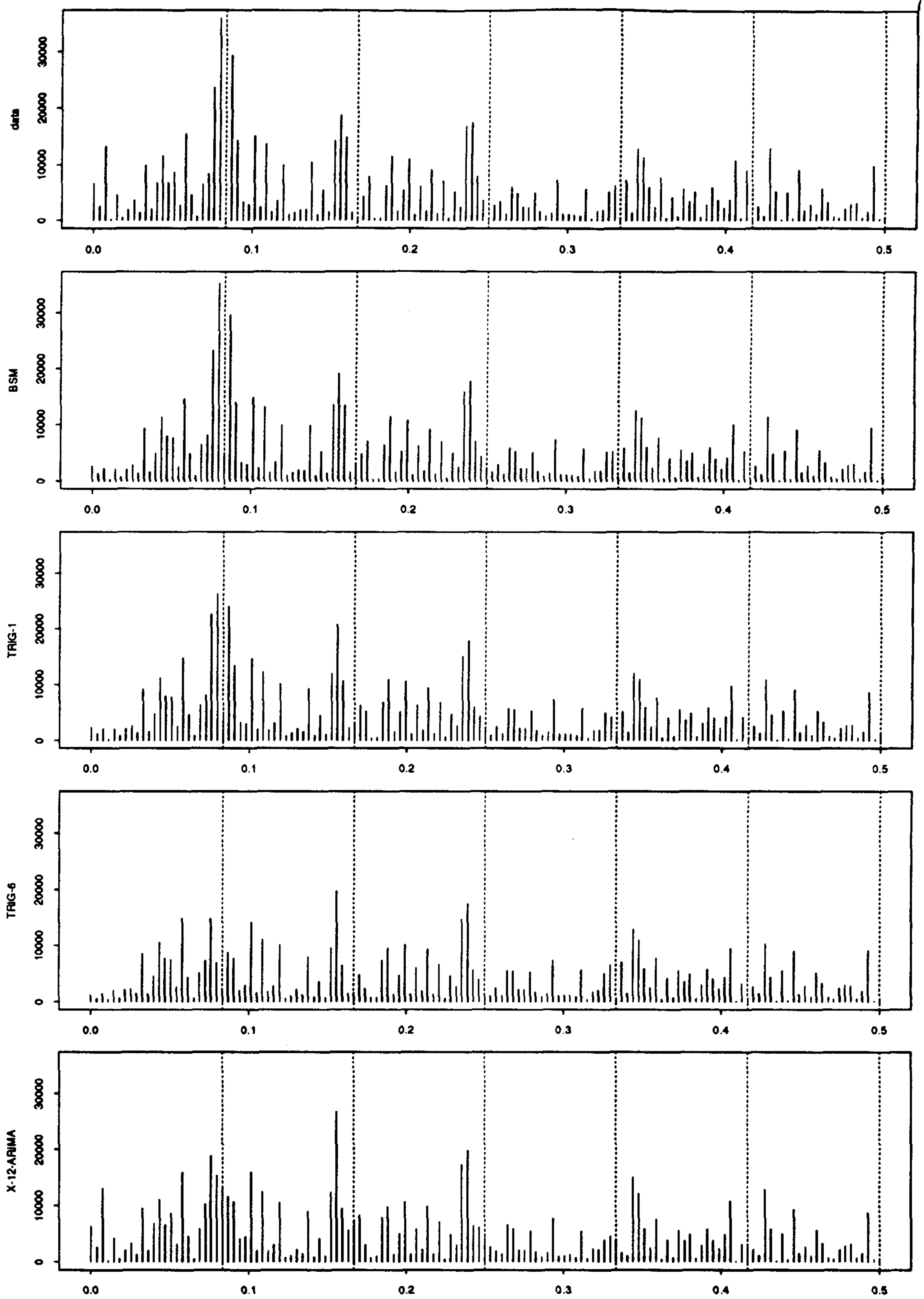


Figure 35: The top plot is the periodogram of the detrended data for CNETHS (with the fundamental and its harmonics suppressed). The other plots give the periodograms for the detrended seasonally adjusted data for BSM, TRIG-1, TRIG-6, and X-12-ARIMA.

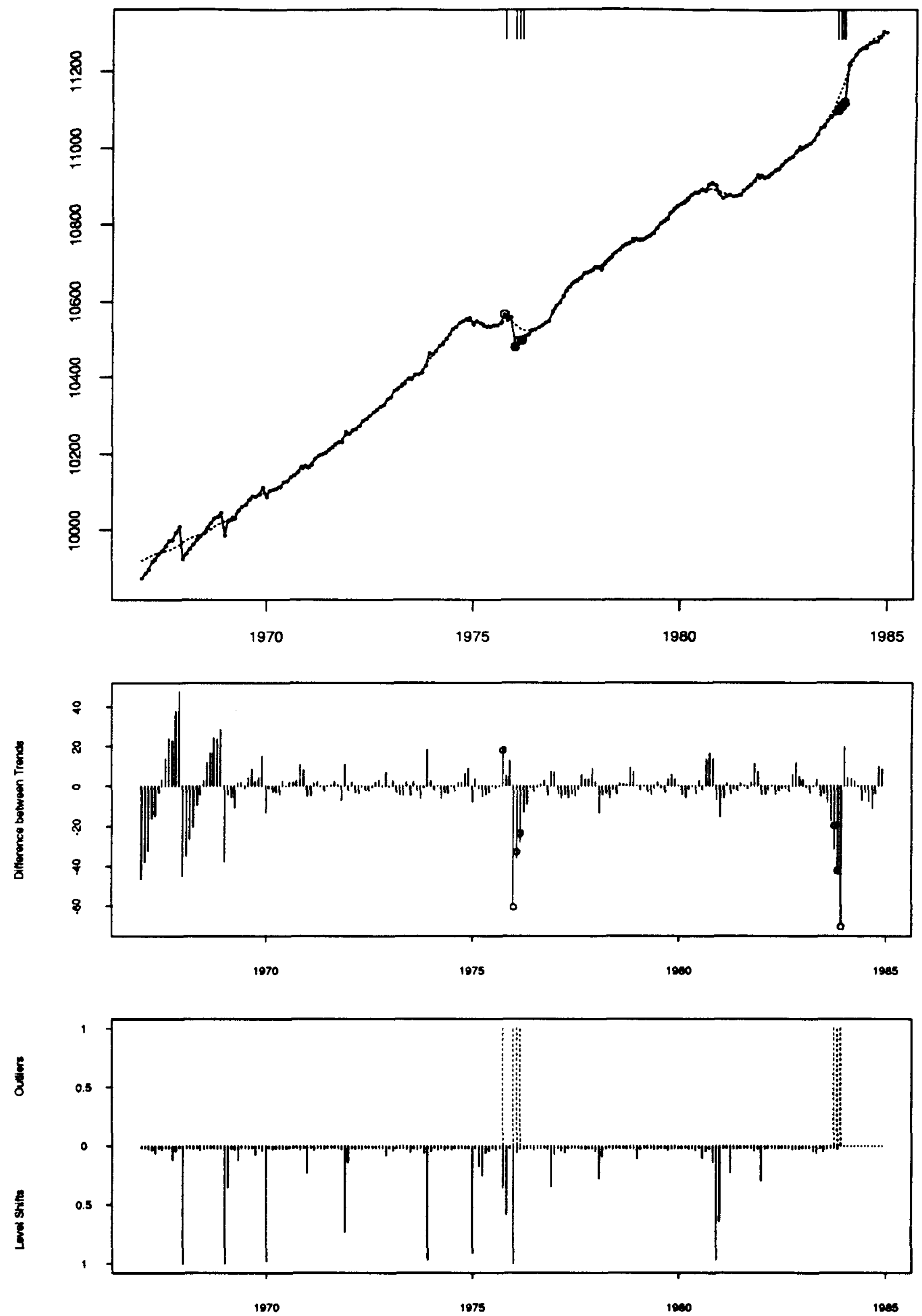


Figure 36: The top plot compares the trends for TRIG-6 and X-12-ARIMA for BTNDRI. The middle plot displays the ratio of the trends. The bottom plot gives the posterior probability of AO's and LS's for TRIG-6 (solid lines) the X-12-ARIMA (dashed lines - either 0 or 1).

References

Alspach, D. L. and Sorenson, H. W. (1972). Nonlinear Bayesian estimation using Gaussian sum approximations. *IEEE Transactions on Automatic Control*, AC-17:439–448.

Ansley, C. F. and Kohn, R. (1985). Estimation, filtering, and smoothing in state space models with incompletely specified initial conditions. *The Annals of Statistics*, 13(4):1286–1316.

Becker, R. A., Chambers, J. C., and Wilks, A. R. (1988). *The New S Language: An Programming Environment for Data Analysis and Graphics.* Wadsworth.

Bell, W. (1986). A computer program for detecting outliers in time series. In *American Statistical Association Annual Conference.*

Bell, W. and Hillmer, S. (1983). Modeling time series with calendar variations. *Journal of the American Statistical Association*, 78:526–534.

Bell, W. and Hillmer, S. (1987). Initializing the Kalman filter in the nonstationary case. Research Report CENSUS/SRD/RR-87/33, Statistical Research Division, Bureau of the Census, Washington, D.C. 20233.

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

Box, G. E. P., Hillmer, S. C., and Tiao, G. C. (1978). Analysis and Modeling of Seasonal Time Series. In Zellner, A., editor, *Seasonal Analysis of Economic Time Series*, pages 309–334. U. S. Bureau of the Census, Washington, D. C.

Bruce, A. G. (1992). Notes on Two Filter Smoother Formulae. Technical report, Victoria University, Wellington, New Zealand. In preperation.

Bruce, A. G. and Jurke, S. (1992). Empirical comparison of two methods for non-Gaussian seasonal adjustment: a book of plots. Technical report, Victoria University, Wellington, New Zealand.

Bruce, A. G. and Martin, R. D. (1992). Tree based robust Bayesian estimation of time series structural models. Technical Report 107, Deptartment of Statistics, University of Washington, Seattle, WA, 98195.

Burman, P. J. (1980). Seasonal Adjustment by signal extraction. *Journal of the Royal Statistical Society, Series A*, 143:321–337.

Chang, I. and Tiao, G. C. (1983). Estimation of time series parameters in the presence of outliers. Technical Report 8, University of Chicago, Graduate School of Business.

Chang, I., Tiao, G. C., and Chen, C. (1988). Estimation of time series parameters in the presence of outliers. *Technometrics*, 30:193–204.

Cleveland, W. S. and Devlin, S. J. (1980). Calendar effects in monthly time series: detection by spectrum analysis and graphical methods. *Journal of the American Statistical Association*, 75:487–496.

Cleveland, W. S. and Terpenning, I. J. (1982). Graphical methods for seasonal adjustment. *Journal of the American Statistical Association*, 77(377):52–62.

Dagum, E. B., , Huot, G., and Morry, M. (1988). Seasonal adjustment in the eighties: some problems and solutions. *Canadian Journal of Statistics*, 16:109–126.

Dagum, E. B. (1978). Modelling, forecasting and seasonally adjusting economic time series with the X-11 ARIMA metohd. *The Statistician*, 27(3,4):203–216.

Dagum, E. B. (1980). The X-11-ARIMA seasonal adjustment method. Research paper, Seasonal Adjustment and Time Series Staff, Statistics Canada, 25-A, R.H. Coats Bldg., Ottawa K1A 0T6.

Dagum, E. B. and Morry, M. (1984). Basic issues on the seasonal adjustment of the Canadian consumer price index. *Journal of Economic and Business Statistics*, 2(3):250–259.

De Jong, P. (1988). The likelihood for a state space model. *Biometrika*, 75(1):165–169.

De Jong, P. (1991). The diffuse Kalman filter. *The Annals of Statistics*, 19(2):1073–1083.

den Butter, F. A. G., , Coenen, R. L., and van de Gevel, F. J. J. S. (1985). The use of ARIMA models in seasonal adjustment. *Emprical Economics*, 10:209–230.

den Butter, F. A. G. and Mourik, T. J. (1990). Seasonal adjustment using structural time series models: an application and a comparison with the Census X-11 method. *Journal of Economic and Business Statistics*, 8(4):385–394.

Dennis, J. E., Gay, D. M., and Welsch, R. E. (1981). An adaptive nonlinear least-squares algorithm. *ACM Transactions on Mathematical Software*, 7(3):348–383.

Findley, D. F., Monsell, B. C., Otto, M. C., Bell, W. R., and Pugh, M. G. (1988). Toward X-12-ARIMA. In *Proceedings of the Fourth Annual Research Conference, U.S. Bureau of the Census*, pages 591–622.

Findley, D. F., Monsell, B. C., Shulman, H. B., and Pugh, M. G. (1990). Sliding-spans diagnostics for seasonal and related adjustments. *Journal of the American Statistical Association*, 85(410):345–355.

Fox, A. J. (1972). Outliers in time series. *Journal of the Royal Statistical Society, Series B*, 34:350–363.

Gay, D. M. (1979). SUMSOL: A general quasi-Newton nonlinear function optimizer.

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

Gordon, K. and Smith, A. F. M. (1990). Modeling and monitoring biomedical time series. *Journal of the American Statistical Association*, 85(410):328–337.

Granger, C. W. J. (1978). Seasonality: causation, interpretation, and implications (with discussion). In Zellner, A., editor, *Seasonal analysis of economic time series*, pages 33–55, U.S. Bureau of the Census, Washington, D.C.

Hannan, E. J. (1964). The estimation of a changing seasonal pattern. *Journal of the American Statistical Association*, 59:1063–1077.

Hannon, E. J., Terrell, R. D., and Tuckwell, N. E. (1970). The seasonal adjusmtnet of economic time series. *International Economic Review*, 11(1):24–52.

Harrison, P. J. and Stevens, C. F. (1971). A Bayesian approach to short-term forecasting. *Operational Research Quarterly*, 22:341–362.

Harrison, P. J. and Stevens, C. F. (1976). Bayesian Forecasting. *Journal of the Royal Statistical Society, Series B*, 38:205–247.

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

Harvey, A. C. (1984). A unified view of statistical forecasting procedures. *J. Forecast.*, 3:245–275.

Harvey, A. C. and Todd, P. J. J. (1983). Forecasting economic time series with structural and Box-Jenkins models: a case study. *Journal of Economic and Business Statistics*, 1(4):299–315.

Harvey, A. C. and Valls Pereira, P. L. (1989). Trend, seasonality and seasonal adjustment. In *Proceedings of the Seminar held in Mar del Plata*, pages 163–199, Argentina.

Hillmer, S. C., Bell, W. R., and Tiao, G. C. (1983). Modeling considerations in the seasonal adjustment of economic time series. In Zellner, A., editor, *Applied time series analysis of economic data*, pages 74–100. U. S. Bureau of the Census, Washington, D. C.

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.

Jain, R. K. (1989). The seasonal adjustment procedures for the consumer price indexes: some empirical results. *Journal of Economic and Business Statistics*, 7(4):461–469.

Kitagawa, G. (1987). Non-Gaussian state space modeling of nonstationary time series. *Journal of the American Statistical Association*, 82(400):1032–1063.

Kitagawa, G. (1988). Numerical Approach to Non-Gaussian smoothing and its applications. In *20th Symposium on the Interface: Computing Science and Statistics*, pages 82–91.

Kitagawa, G. (1990). The two-filter formula for smoothing and an implementation of the Gaussian-sum smoother. Technical report, The Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-ku Tokyo, Japan 106.

Kitagawa, G. (1991). Unpublished FORTRAN implementation of the Gaussian sum two filter smoother.

Kitagawa, G. and Gersch, W. (1984). A smoothness priors-state space approach modeling of time series with trend and seasonality. *Journal of the American Statistical Association*, 79:378–389.

Maravall, A. (1985). On structural time series models and the characterization of components. *Journal of Economic and Business Statistics*, 3:350–355.

Maravall, A. and Pierce, D. A. (1987). A prototypcial seasonal adjustment model. *Journal of Time Series Analysis*, 8:177–193.

Monsell, B. C. (1983). Using the Kalman smoother to adjust for moving trading day. Research Report SEC/RR-83/04, Statistical Research Division, U.S. Bureau of the Census, Room 3000-4, Washington, D.C. 20233.

Monsell, B. C. (1990). Documentation for the current prototype of the Census X-12-ARIMA seasonal adjustment program: Version 1.0. Technical report, Statistical Research Division, U.S. Bureau of the Census, Room 3000-4, Washington, D.C. 20233.

Monsell, B. C. and Otto, M. C. (1991). Documentation for the REGARIMA program for IBM compatible 386 PC's. Technical report, Statistical Research Division, U.S. Bureau of the Census, Room 3000-4, Washington, D.C. 20233.

Ozaki, T. and Thomson, P. J. (1992). Transformation and seasonal adjustment. (In preparation.) Institute of Statistics and Operations Researc h, Victoria University of Wellington.

Shiskin, J., Young, A. H., and Musgrave, J. C. (1967). The X-11 variant of the Census method II seasonal adjustment program. Technical Paper 15, U.S. Bureau of the Census, U.S. Government Printing Office, Washington, D.C.

Smith, A. and West, M. (1983). Monitoring renal transplants: an application of the Multiprocess Kalman filter. *Biometrics*, 39:867–878.

Tsay, R. S. (1988). Outliers, level shifts, and variance changes in time series. *J. Forecast.*, 7:1–20.