Discussion Paper 8

Institute for Empirical Macroeconomics
Federal Reserve Bank of Minneapolis
250 Marquette Avenue
Minneapolis, Minnesota 55480

January 1989

A SYSTEMS APPROACH TO RECURSIVE ECONOMIC FORECASTING AND SEASONAL ADJUSTMENT

Peter Young, Cho Ng, and Peter Armitage

Centre for Research in Environmental Systems,
Institute of Environmental and Biological Sciences
University of Lancaster

This paper will appear in a special issue of the International Journal on Computers and Mathematics with Applications on "System Theoretic Methods in Economic Modeling," edited by Stefan Mittnik, to appear in 1989.

ABSTRACT

The paper discusses a new, fully recursive approach to the adaptive modeling, forecasting and seasonal adjustment of nonstationary economic time-series. The procedure is based around a time variable parameter (TVP) version of the well known "component" or "structural" model. It employs a novel method of sequential spectral decomposition (SSD), based on recursive state-space smoothing, to decompose the series into a number of quasi-orthogonal components. This SSD procedure can be considered as a complete approach to the problem of model identification and estimation, or it can be used as a first step in maximum likelihood estimation. Finally, the paper illustrates the overall adaptive approach by considering a practical example of a UK unemployment series which exhibits marked nonstationarity caused by various economic factors.

A SYSTEMS APPROACH TO RECURSIVE ECONOMIC FORECASTING AND SEASONAL ADJUSTMENT

by

Peter Young, Cho Ng and Peter Armitage[1]
Centre for Research in Environmental Systems
Institute of Environmental and Biological Sciences
University of Lancaster

1. INTRODUCTION

Recursive estimation has a long history. Karl Freidrich Gauss, first derived the recursive least squares algorithm over one hundred and fifty years ago (see Appendix 2 in Young ,1984, where Gauss's "long hand" derivation is compared with the modern matrix approach). But the recent popularity of recursive estimation was undoubtedly stimulated by the appearance, in 1960, of the now famous paper on state-variable filtering and prediction by the system's theorist Rudolph Kalman. Since the 1960's and 70's, the significance of such recursive estimation and forecasting procedures to economic modelling and econometrics has became ever more apparent, and it is now quite common to see detailed references to the Kalman filter in standard econometric and statistical text books (see e.g. Harvey, 1981; Priestley, 1981). During this same time, systems research workers have been actively concerned with the development of recursive methods for the identification and estimation of parameters in the more common, linear representations of discrete time-series, such as the AR, ARMA, ARMAX and Box-Jenkins models (see e.g. Young, 1984; Ljung and Soderstrom, 1983).

In the present paper, we exploit the excellent spectral properties of certain special recursive estimation and smoothing algorithms to develop a practical and unified approach to **adaptive** economic forecasting and seasonal adjustment. The approach is based around the well known "structural" or "component" time-series model[2] and, like previous, similar, state-space solutions (e.g. Harrison and Stevens, 1976; Kitagawa, 1981; Harvey, 1984), it employs the standard Kalman filter-type recursive algorithms. Except in the final forecasting and smoothing stages of the analysis, however, the justification for using these algorithms is not the traditional one based on "optimality" in a prediction error or maximum likelihood (ML) sense. Rather, the algorithms are utilised in a manner which allows for straightforward and effective **sequential spectral decomposition** of the time series into the **quasi-orthogonal** components of the model. A unifying element in this analysis is the modelling of **nonstationary state variables** and **time variable parameters** by a class of **second order random walk models**. As we shall see, this simple device not only facilitates the development of the spectral decomposition algorithms but it also injects an inherent adaptive capability which can be exploited in both forecasting and seasonal adjustment.

2. THE COMPONENT TIME SERIES MODEL

Although the analytical procedures proposed in this paper can be applied to multivariable (vector) processes (see Ng et et al,1988), we will restrict the discussion, for simplicity of exposition, to the following two component models of a univariate (scalar) time-series y(k),

$$y(k) = t(k) + p(k) + e(k) \quad (1)$$

$$y(k) = t(k) + n(k) + e(k) \quad (2)$$

1. The Civil Service College, 11 Belgrave Rd., London SWIV 1RB.
2. The term "structural" has been used in other connections in both the statistical and economics literatures and so we will employ the former term.

where, t(k) is a low frequency or **trend** component; p(k) is a **periodic or seasonal component**; n(k) is a general **stochastic perturbation** component; and e(k) is a zero mean, serially uncorrelated **white noise** component, with variance σ^2. The model (1) is appropriate for economic data exhibiting pronounced trend and periodicity and is the main vehicle utilised in the present paper for the development of adaptive seasonal adjustment procedures. The second model (2) can also be used to represent such heavily periodic time-series but it has much wider applicability to quasi-periodic and non-periodic phenomena. It is the utilised here mainly for the development of recursive state-space forecasting algorithms. Both models, however, are special cases of the general component model discussed in detail by Young (1988) and Ng and Young (1988).

Component models such as (1) and (2) have been popular in the literature on econometrics and forecasting (e.g. Nerlove et al, 1979; Bell and Hillmer, 1984) but it is only in the last few years that they have been utilised within the context of state-space estimation. Probably the first work of this kind was by Harrison and Stevens (1971,1976) who exploited state-space methods by using a Bayesian interpretation applied to their "Dynamic Linear Model" (effectively a regression model with time variable parameters). More recent papers which exemplify this state-space approach and which are particularly pertinent to the present paper, are those of Jakeman and Young (1979,1984), Kitagawa (1981), Kitagawa and Gersch (1984), and Harvey (1984).

In the state-space approach, each of the components t(k), p(k) and n(k) is modelled in a manner which allows the observed time series y(k) to be represented in terms of a set of discrete-time state equations. And these state equations then form the basis for recursive state estimation, forecasting and smoothing. Before we investigate the use of these analytical techniques, therefore, it is appropriate to consider the specific form of the models for t(k), p(k) and n(k).

2.1 The Trend Model

It is assumed that the low frequency or trend behaviour t(k) can be represented by one of the family of stochastic, **Generalised Random Walk** (GRW) models. In practice, the most important of these is the second order GRW which can be written in the following form,

$$\mathbf{x}_t(k) = \mathbf{F}_t\, \mathbf{x}_t(k-1) + \mathbf{G}_t\, \eta_t(k-1) \qquad (3)$$

where,

$$\mathbf{x}_t(k) = [\ t(k)\ d(k)\ ]^T \ \text{ and } \ \eta_t(k) = [\ \eta_{t1}(k)\ \eta_{t2}(k)\ ]^T$$

and,

$$\mathbf{F}_t = \begin{bmatrix} \alpha & \beta \\ 0 & \gamma \end{bmatrix} \quad ; \quad \mathbf{G}_t = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Here, $\eta_{t1}(k)$ and $\eta_{t2}(k)$ represent zero mean, serially uncorrelated, discrete white noise inputs, with the vector $\eta_t(k)$ normally characterised by a covariance matrix $\mathbf{Q}_t$, i.e.,

$$E\{\ \eta_\tau(k)\, \eta_t(k)^T\} = \mathbf{Q}_t\, \delta_{k,j} \quad ; \quad \delta_{k,j} = \begin{cases} 1 \text{ for } k=j \\ 0 \text{ for } k\neq j \end{cases}$$

where, $\delta_{k,j}$ is the Kronecker delta function. Unless there is evidence to the contrary, $\mathbf{Q}_t$ is assumed to be diagonal in form with unknown elements q_{t11} and q_{t22}, respectively.

This GRW model subsumes, as special cases (see e.g. Young,1984): the very well known and used **Random Walk** (RW: α=1; β=γ=0; $\eta_{t2}(k)$=0); the **Smoothed Random Walk** (SRW: β=γ=1; 0< α<1.0; $\eta_{t1}(k)$=0); and, most importantly in the present paper, the **Integrated Random Walk** (IRW: α=β=γ=1; $\eta_{t1}(k)$=0). In the case of the IRW, we see that t(k) and d(k) can be interpreted as level and slope variables associated with the variations of the trend, with the random disturbance entering only through the d(k) equation. If $\eta_{t1}(k)$ is non-zero, however, then both the level and slope equations can have random fluctuations defined by $\eta_{t1}(k)$ and $\eta_{t2}(k)$, respectively. This variant has been termed the "Linear Growth Model" by Harrison (1967), Harrison and Stevens (1971,1976).

The advantage of these random walk models is that they allow, in a very simple manner, for the introduction of nonstationarity into the time series models. By introducing a trend model of this type, we are assuming that the time-series can be characterised by a stochastically variable mean value. The nature of this variability will depend upon the specific form of the GRW chosen: for instance, the IRW model is particularly useful for describing large smooth changes in the trend; while the RW model provides for smaller scale, less smooth variations (Young,1984). As we shall see later in Section 2.5, these models can also be utilised to allow for similar behaviour in the *parameters* (coefficients) of the component models for n(k) and p(k). And, by defining the stochastic inputs $\eta_{t1}(k)$ and $\eta_{t2}(k)$ in a specific manner, we shall also see how the same models can be used to handle large, abrupt changes in the level and slope of either the trend or the model coefficients.

2.2 The Periodic or Seasonal Model

It is assumed that the periodic component in model (1) can be defined by the following **Dynamic Harmonic Regression** (DHR) relationship,

$$p(k) = \sum_{i=1}^{i=F} \theta_{1i}(k)\ \cos(2\pi f_i k) + \theta_{2i}(k)\ \sin(2\pi f_i k) \tag{4}$$

where the regression coefficients $\theta_{ji}(k)$, j=1,2 and i=1,2,...,F, may be constant (i.e. $\theta_{ji}(k)=\theta_{ji}$ for all k), when the model is simply the conventional harmonic regression in F different but constant frequencies; or time-variable, in which case the model is able to handle nonstationary seasonality, as discussed later in Section 6.1. This latter version, in which the parameter variations are modelled as GRW processes, is extremely useful for time-series which exhibit amplitude modulated periodic behaviour, such as the growing amplitude seasonality of the airline passenger data in Fig.1 and the heavily modulated seasonality in the unemployment series shown in Fig.3(a) (see later). Since there are two parameters associated with each frequency component, the changes in the amplitude A(k) of each component, as defined by,

$$A_i(k) = \sqrt{\theta_{1i}(k)^2 + \theta_{2i}(k)^2}$$

provides a useful indication of the estimated amplitude modulation.

2.3 The Stochastic Perturbation Model

In order to allow for general stochastic perturbations with decaying or growing amplitude, n(k) and e(k) are combined and represented by a stochastic **General Transfer Function** (GTF) model: this is similar to the well known ARMA model employed in Box-Jenkins forecasting (Box and Jenkins,1970), but no stationarity restrictions are applied. The GTF model is best identified, and its parameters estimated, within a traditional transfer function framework. In order to consider the model in state-space form, however, it is most convenient to assume that the sum of the stochastic perturbation and the white noise component constitutes an ARMA process with the same white noise input e(k), i.e.,

$$n(k) + e(k) = \frac{D(z^{-1})}{C(z^{-1})} e(k) \tag{5}$$

where,

$$C(z^{-1}) = 1 + c_1 z^{-1} + c_2 z^{-2} + ,......, + c_p z^{-p}$$

$$D(z^{-1}) = 1 + d_1 z^{-1} + d_2 z^{-2} + ,......, + d_p z^{-p}$$

Note that, for convenience, the order p is assumed the same for both polynomials: different orders can, however, be introduced simply by assuming that appropriate trailing coefficients are zero. Similarly, "subset" models (see e.g Whittle,1952; Priestley,1981) can be specified by constraining the selected intermediate coefficients to zero value.

It is now straightforward to transform the model (5) into the following "innovations" state-space form (e.g. Åstrom,1970), defined completely by the estimated parameters of the GTF model, i.e.,

$$\left.\begin{aligned} \mathbf{x}_n(k) &= \mathbf{F}_n \mathbf{x}_n(k-1) + \mathbf{G}_n e(k-1) \\ n(k) &= [1\ 0\ 0\ .....\ 0\ ]\ \mathbf{x}_n(k) + e(k) \end{aligned}\right\} \tag{6}$$

where,

$$\mathbf{x}_n(k) = [\ n(k)\ n_2(k)\ .....\ n_p\ ]^T$$

and $\mathbf{F}_n$ and $\mathbf{G}_n$ have the canonical form,

$$\mathbf{F}_n = \begin{bmatrix} -c_1 & 1 & 0 & . & . & . & . & . & 0 \\ -c_2 & 0 & 1 & . & . & . & . & . & 0 \\ . & . & . & . & 1 & & . & . & 0 \\ . & . & . & & & & 1 & & 0 \\ . & . & . & . & . & . & . & . & 1 \\ -c_p & 0 & 0 & . & . & . & . & . & 0 \end{bmatrix} \quad ; \quad \mathbf{G}_n = \begin{bmatrix} g_1 \\ g_2 \\ . \\ . \\ . \\ g_p \end{bmatrix}$$

where,

$$g_i = d_i - c_i \ ; \quad i=1,2,...,p$$

with the order p defined by some form of statistical identification criterion (e.g. Akaike, 1974). This is the general state-space form for a GTF or ARMA model; if an AR or subset AR model is identified for the perturbations, then the g_i parameters are identically equal to the negative of the AR coefficients, i.e $-c_i$.

2.4 The Complete State-Space Model

Having defined state-space model structures for all of the components of the model, it is straight forward to assemble these into the following aggregate state-space form,

$$\left.\begin{aligned} x(k) &= \mathbf{F}\, x(k-1) + \mathbf{G}\, \eta(k-1) && \text{(i)} \\ y(k) &= \mathbf{H} x(k) + e(k) && \text{(ii)} \end{aligned}\right\} \qquad (7)$$

where the state vector x(k) is composed of all the state variables from the component sub-models; and the observation vector **H** is chosen to extract from the state vector x(k) the appropriate structural components t(k), p(k), or t(k) and n(k) in equations (1) or (2), respectively. In other words, depending on which model is being considered, either equation (1) or equation (2) will appear as the observation equation (7)(ii). The disturbance vector η(k) is defined by the disturbance inputs of the constituent sub-models. In the case of equation (2), for example, the state space model (7) can be represented in the following partitioned form,

$$x(k) = \begin{bmatrix} F_t & 0 \\ 0 & F_n \end{bmatrix} x(k-1) + \begin{bmatrix} G_t & 0 \\ 0 & G_n \end{bmatrix} \eta(k-1)$$

$$y(k) = \mathbf{H}\, x(k) + e(k)$$

where,

$$x(k) = [\, x_t^T(k)\; x_n^T(k)]^T$$

$$\eta(k) = [\, \eta_t^T(k)\;\; e(k)\,]^T$$

$$\mathbf{H} = [1\ 0\ 1\ 0\ \ 0 \ldots 0]$$

This particular form of the model with IRW trend and GTF or AR stochastic disturbance components is quite useful for general univariate economic and business forecasting applications (see Ng and Young,1988). As we shall see later, the alternative model (1) is more appropriate in seasonal adjustment applications when there is sustained, nonstationary periodicity.

2.5 Parametric NonStationarity and Variance Intervention

In the present context, the GRW model (3) is important not only as a convenient representation of the trend component, but also because we exploit it in the development of **Time Variable Parameter** (TVP) models. Here, it is assumed that any model parameters, such as the harmonic regression coefficients θ_{1i} and θ_{2i}, i=1,...,F in equation (4) or the coefficients c_i, d_i, i=1,...,p, in the GTF model (5), are potentially time-variable, with stochastic variations that can be represented by the GRW[3]. In other words, the time-series y(k), in either of the models (1) and (2), may possess a wide variety of nonstationary characteristics.

In general, we might assume that economic and business time series are particularly appropriate for the TVP approach to modelling. In the long term, the socioeconomic system is clearly nonlinear and subject to many changes caused by factors such as: variations in social behaviour and attitudes; modifications in government policies; and changes in the methods of acquiring, measuring and interpreting social statistics. It seems reasonable, therefore, to assume that even the small perturbational dynamic behaviour of such a system (i.e the fluctuations about the long term trends) will only be described adequately by linear models if we allow for the possibility of changes in the model parameters over the passage of time.

But the nature of such parametric variation is difficult to predict: while modifications in the socioeconomic system are often relatively slow and smooth, more rapid and violent changes do occur from time-to-time and lead to similarly rapid changes, or even discontinuities, in the related time series. Typical examples are shown in Figs. 2 and 3(a): Fig.2 is a plot of monthly car driver casualties in the

---

3. in the case of Regression relationships, they are termed "dynamic" models; i.e **Dynamic Linear, Harmonic** or **Auto-Regression** (see e.g. Young,1988, Young and Benner, 1988).

U.K over the period 1970 to 1984 (Harvey and Durbin,1986) in which changes of level, due to both the oil crisis of the 1970's and recent changes of U.K Government legislation on seat belts, are clearly apparent; Fig.3(a) shows the monthly variations in the unemployment figures for school leavers in the U.K. over the same period. These have been drastically affected by the oil crisis, changes in Government and several fairly major modifications in the method of measurement after 1979.

The GRW model is well able to characterise changes such as those shown in Figs.1 to 3. If the variances q_{tii}, i=1,2, are assumed constant, then the model, in its various RW, IRW and SRW forms, can describe a relatively wide range of variation in the associated trend or model parameters. Moreover, if we allow these variances to change over time, then an even wider range of behaviour can be accommodated. In particular, large, but otherwise arbitrary, instantaneous changes in q_{t11} and q_{t22} (e.g. increases to values > 10^2) introduced at selected "intervention" points, can signal to the associated estimation algorithm the possibility of significant changes in the level or slope, respectively, of the modelled variable at these same points. The sample number associated with such intervention points can be identified either objectively, using statistical detection methods (e.g. Jun,1988; Tsay,1988); or more subjectively by the analyst (see Young and Ng, 1988).

It is interesting to note that this same device, which we term **variance intervention** (Young and Ng,1988) can be applied to any state-space or TVP model: Young (1969,1970,1971,1981), for example, has used a similar approach to track the significant and rapid changes in the level of the model parameters of an aerospace vehicle during a rocket boost phase. It is straightforward to develop similar TVP versions of the instrumental variable (IV) and approximate maximum likelihood (AML; or extended least squares, ELS) algorithms for transfer function model estimation (see Section **6.2** and Kaldor,1978; Norton,1975).

3. THE RECURSIVE FORECASTING AND SMOOTHING ALGORITHMS

In this paper, recursive forecasting and smoothing is achieved using the state-space (Kalman) filtering and fixed-interval smoothing algorithms. The Kalman filtering algorithm (Kalman,1960) is, of course, well known (see e.g. Young,1984) and can be written most conveniently in the following general "prediction-correction" form,

Prediction :

$$\hat{x}(k/k-1) = \mathbf{F}\,\hat{x}(k-1)$$
$$\mathbf{P}(k/k-1) = \mathbf{F}\,\mathbf{P}(k-1)\,\mathbf{F}^T + \mathbf{G}\,[\mathbf{Q}_r]\,\mathbf{G}^T \qquad (8)$$

Correction :

$$\hat{x}(k) = \hat{x}(k/k-1) + \mathbf{P}(k/k-1)\mathbf{H}^T[\,1 + \mathbf{H}\,\mathbf{P}(k/k-1)\,\mathbf{H}^T\,]^{-1}\,\{\,y(k) - \mathbf{H}\,\hat{x}(k/k-1)\,\}$$
$$\mathbf{P}(k) = \mathbf{P}(k/k-1) - \mathbf{P}(k/k-1)\mathbf{H}^T[1 + \mathbf{H}\mathbf{P}(k/k-1)\mathbf{H}^T]^{-1}\mathbf{H}\,\mathbf{P}(k/k-1) \qquad (9)$$

In these equations, we use $\hat{x}(k)$ to denote the estimate of either one of the state vectors associated with the structural components (i.e $\mathbf{x}_t$ and $\mathbf{x}_n$ as defined in equations (3) and (6)) *or* the composite state vector x of the complete state-space model (7)). The other matrices are defined accordingly, with $\mathbf{Q}_r$ denoting the covariance matrix of the input white noise disturbances, i.e.,

$$E\{\eta(k)\eta(j)^T\} = \mathbf{Q}_r\,\delta_{k,j}$$

where $\delta_{k,j}$ is the Kronecker delta function.

Given the nature of the structural models (1) and (2), it is clear that this assumption of white observational errors will not apply *unless we consider all the components simultaneously*. It is this assumption that makes the analytical procedures presented here sub-optimal in a strict maximum likelihood and Bayesian sense. This in no way negates the utility of the proposed approach, however, since we do not view the algorithms from such a theoretical standpoint. Rather, we justify their use on the basis of their spectral properties which, as we shall see, are particularly attractive for achieving spectral decomposition.

It will be noted that, since the random walk class of models are all characterised by a scalar observation equation, the filtering algorithm has been manipulated into the well known form (see e.g. Young,1984) where the "noise variance ratio" (NVR) matrix $\mathbf{Q}_r$ and the P(k) matrix are both defined in relation to the white measurement noise variance σ^2, i.e.,

$$\mathbf{Q}_r = \mathbf{Q}/\sigma^2 ; \quad \mathbf{P}(k) = \mathbf{P}^*(k)/\sigma^2 \tag{10}$$

where $\mathbf{P}^*(k)$ is the error covariance matrix associated with the state estimates. In the RW and IRW, models, moreover, there is only a single white noise input term, so that only a scalar NVR value has to be specified by the analyst.

There are a variety of algorithms for off-line, fixed interval smoothing but the one we will consider here utilises the following backwards recursive algorithm for the smoothed estimate $\hat{\mathbf{x}}(k/N)$, subsequent to application of the above Kalman filtering forwards recursion (see e.g. Norton,1975; Young,1984),

$$\hat{\mathbf{x}}(k/N) = \mathbf{F}^{-1}[\, \hat{\mathbf{x}}(k+1/N) + \mathbf{G}\, \mathbf{Q}_r \mathbf{G}^T\, \mathbf{L}(k)\,] \tag{11}$$

where,

$$\mathbf{L}(N) = 0 \, ;$$

N is the total number of observations (the "fixed interval"); and

$$\mathbf{L}(k) = [\, \mathbf{I} - \mathbf{P}(k+1)\, \mathbf{H}^T\mathbf{H}\,]^T \{\, \mathbf{F}^T\mathbf{L}(k+1) - \mathbf{H}^T[\, y(k) - \mathbf{H}\, \mathbf{F}\, \hat{\mathbf{x}}(k)\,]\, \}$$

is an associated backwards recursion for the "Lagrange Multiplier" vector $\mathbf{L}(k)$ required in the solution of this two point boundary value problem. Finally, the covariance matrix $\mathbf{P}^*(k/N) = \sigma^2\mathbf{P}(k/N)$ for the smoothed estimate is obtained by reference to $\mathbf{P}(k/N)$ generated by the following matrix recursion,

$$\mathbf{P}(k/N) = \mathbf{P}(k) + \mathbf{P}(k)\mathbf{F}^T[\mathbf{P}(k+1/k)]^{-1}\{\mathbf{P}(k+1/N)-\mathbf{P}(k+1/k)\}\,[\mathbf{P}(k+1/k)]^{-1}\,\mathbf{F}\mathbf{P}(k) \tag{12}$$

while the smoothed estimate of original series y(k) is given simply by,

$$\hat{y}(k/N) = \mathbf{H}\, \hat{\mathbf{x}}(k/N) \tag{13}$$

i.e., the appropriate linear combination of the smoothed state variables.

As we shall see in the next Section, these recursive smoothing equations for the various component models are exploited to decompose the signal y(k) into its various quasi-orthogonal elements. In this manner, the component models are identified and estimated and it is possible to formulate the complete discrete-time, state-space model (7). The procedures for smoothing and forecasting of y(k) then follow straightforwardly, once again by the application of the state-space filtering/smoothing algorithm but this time applied, in their more general form, to the complete state equations (7). This then allows for the following operations:-

(1) **Forecasting** The f step ahead forecasts of the composite state vector x(k) in equation (7) are obtained at any point in the time series by repeated application of the prediction equations (8) which, for the complete model, yields the equation,

$$\hat{\mathbf{x}}(k+f/k) = \mathbf{F}^{f}\,\hat{\mathbf{x}}(k) \tag{14}$$

where f denotes the forecasting period. The associated forecast of y(k) is provided by,

$$\hat{y}(k+f/k) = \mathbf{H}\,\hat{\mathbf{x}}(k+f/k) \tag{15}$$

with the variance of this forecast computed from,

$$\mathrm{var}\,\{\tilde{y}(k+f/k)\} = \sigma^2[\,1+\mathbf{H}\;\mathbf{P}(k+f/k)\mathbf{H}^T] \tag{16}$$

where $\tilde{y}(k+f/k)$ is the f step ahead prediction error, i.e.,

$$\tilde{y}(k+f/k) = y(k+f) - \hat{y}(k+f/k)$$

In relation to more conventional alternatives to forecasting, such as those of Box and Jenkins, the present state-space approach, with its inherent component decomposition, has the advantage that the estimates and forecasts of individual component state variables can be obtained simply as by-products of the analysis. For example, it is easy to recover the estimate and forecast of the trend component, which can be considered as a simple, on-line estimate of the "seasonally adjusted" series and provides a measure of the underlying "local" trend at the forecasting origin.

(2) **Forward Interpolation** Within this discrete-time setting, the process of forward interpolation, in the sense of estimating the series y(k) over a section of missing data, based on the data up to that point, follows straightforwardly: the missing data points are accommodated in the usual manner by replacing the observation y(k) by the predicted value $\hat{y}(k/k-1)$ and omitting the correction equations (9). Such a procedure can be used for the complete model (7) or for the component sub-models discussed in Section 2.

(3) **Smoothing** Finally, the smoothed estimate $\hat{y}(k/N)$ of y(k) for all values of k is obtained directly from equation (13); and associated smoothed estimates of all the component states are available from equation (11). Smoothing can, of course, provide a superior interpolation over gaps in the data, in which the interpolated points are now based on all of the N samples. As in the case of forward interpolation, suboptimal smoothed estimates of the structural model components can be obtained by applying the same two-pass smoothing algorithm separately and sequentially to the component sub-models; indeed this is precisely the procedure utilised in the spectral decomposition technique discussed in the next Section of the paper.

4. IDENTIFICATION AND ESTIMATION OF THE STRUCTURAL MODELS

The problems of structure identification and subsequent parameter estimation for the complete state space model (7) are clearly non-trivial. From a theoretical standpoint, the most obvious approach is to formulate the problem in Maximum Likelihood (ML) terms. If the stochastic disturbances in the state-space model are normally distributed, the likelihood function for the observations may then be obtained from the Kalman Filter via "prediction error decomposition" (Schweppe,1965)). For a suitably identified model, therefore, it is possible, in theory, to maximise the likelihood with respect to any or all of the unknown parameters in the state-space model, using some form of numerical optimisation.

This kind of maximum likelihood approach has been tried by a number of research workers but their results (e.g. Harvey and Peters,1984) suggest that it can be quite complex, even if particularly simple structural models are utilised (e.g. those containing merely trend and seasonal models, in which the only unknown parameters are the variances of the stochastic disturbances, and where no stochastic perturbation component n(k) is included). In addition it is not easy to solve the ML problem in practically useful and **completely** recursive terms; i.e. with the parameters being estimated recursively as well as the states.

The alternative approach suggested here can be considered from two standpoints:-

First, it can be interpreted as a first step in ML estimation; one which allows for the identification of an appropriate model structure and provides the initial, sub-optimum (in the ML sense) estimates of the unknown parameters that characterise this model structure. In this manner, the initial estimates required for numerical optimisation should be close to their final optimum values. This seems particularly important in the present context, since the the likelihood function is not always well defined in the region of the optimum (see Ng, 1987).

Second, the proposed method can be considered simply in spectral or filtering terms as *a method of decomposing the original time-series y(k) into a number of quasi-orthogonal components*; namely, the components of the models (1) and (2). These components are then modelled separately, prior to their use in the formulation of the aggregate state-space model (7). Here, the nominal sub-optimality of the solution in the strict ML sense is counteracted by its simple "filtering" interpretation, which should appeal to the practical user. Moreover, while it will normally be suboptimal in the ML sense (i.e. minimisation of the overall one step ahead prediction errors), the proposed method appears to function very well in a wider sense and quite often seems to out-perform the ML solution in longer period forecasting terms.

5. THE SPECTRAL PROPERTIES OF THE SMOOTHING ALGORITHMS

The process of **Sequential Spectral Decomposition** (SSD) proposed here is based on the application of the state-space, "fixed interval" smoothing algorithms discussed in the previous Section **3**, as applied to the various component models discussed in Section **2**. In particular, it exploits the excellent spectral properties of the smoothing algorithms derived in this manner. These spectral properties are illustrated in Figs.4 and 5, which show how the amplitude spectra, for the most important IRW and DHR model algorithms, are controlled by the selected NVR value. It is clear that, in all cases, the scalar NVR defines the "bandwidth" of the smoothing algorithm. The phase characteristics are not shown, since the algorithms are all of the "two-pass" smoothing type and so exhibit *zero phase lag* for all frequencies.

Fig.4 shows that the IRW trend algorithm (termed IRWSMOOTH in the microCAPTAIN program) is a very effective "low-pass" filter, with particularly sharp "cut-off" properties for low values of the NVR. The relationship between the $\log_{10}(F_{50})$, where F_{50} is the 50% cut-off frequency, and $\log_{10}(NVR)$ is approximately linear over the useful range of NVR values, so that the NVR which provides a specified cut-off frequency can be obtained from the following approximate relationship (T.J Young, 1987),

$$NVR = 1605[\, F_{50}\,]^4 \tag{17}$$

In this manner, the NVR which provides specified low-pass filtering characteristics can be defined quite easily by the analyst. The band-pass nature of the DHR recursive smoothing algorithm (DHRSMOOTH) is clear from Fig.5 and a similar simple relationship once again exists between the bandwidth and the NVR value. These convenient bandwidth-NVR relationships for IRWSMOOTH and DHRSMOOTH are useful in the proposed procedure for spectral decomposition discussed below.

Clearly, smoothing algorithms based on other simple random walk and TVP models can be developed: for instance the double integrated random walk (DIRW, see Young,1984) smoothing algorithm has even sharper cut-off characteristics than the IRW, but its filtering characteristics exhibit much higher levels of distortion at the ends of the data set (Ng,1987).

5.1 Sequential Spectral Decomposition

The procedure recommended for spectral decomposition of y(k) is as follows:-

1. Plot the y(k) series and review its major statistical characteristics by reference to both the AR (maximum entropy) spectrum, with order identified via the Akaike AIC criterion, and the periodogram (e.g. Priestley,1981). The periodogram provides a very good and detailed spectral description of y(k) which is a useful reference against which to judge the spectral decomposition; while the relative smoothness and superior resolution of peaks provided by the AR spectrum is often more useful in identifying the principal modes of dynamic behaviour.

2. If a very low frequency component or trend is identified in step **1.**, then this should be estimated and removed by the IRWSMOOTH algorithm, if necessary using variance intervention to allow for any sharp changes in the level or slope. The NVR for this operation can be based initially on the F_{50}-NVR relationship, or an equivalent relationship for another bandwidth criterion, e.g. F_{95}. An NVR=0.0001 (F_{50} approx. 0.016 cycles/sample) provides a useful starting value, which appears appropriate to many of the economic time series we have evaluated (see example, Section **7**).

3. Checks on the adequacy of the low pass filtration resulting from the choice of NVR in **2** can now include: (a) comparison of the periodograms for the trend t(k) and detrended, $\delta(k) = y(k) - \hat{t}(k)$, components with the original periodogram for y(k) computed in **1**; reference to the cross correlation function between t(k) and δ(k) to verify reasonable independence; and evaluation of the trend derivative estimate d(k) produced by IRWSMOOTH to ensure that "leakage" of higher frequency components is at a minimum (see Young,1987; Ng,1987).

4. If necessary return to step **2** and choose a revised value of the NVR.

5. Investigate the periodogram and/or the AR spectrum of the detrended δ(k) series and identify the major characteristics and peaks in the spectrum.

6. Either: (i) utilise the DHRSMOOTH algorithm to sequentially estimate the frequency components associated with the spectral peaks identified in **5** (see Section **6.1** below); or (ii) identify and estimate an AR or ARMA model for δ(k), with the order defined by the Akaike Information Criterion (AIC) or some alternative order identification criterion (see Section **6.2** below).

7. Finally, filtering and smoothing should be repeated based on the complete state-space model, as defined by the component models identified in steps **1.** to **6.** Also, if required, this model can provide the initial conditions for ML estimation.

Note that it is possible to automate the above "manual" selection of the trend NVR (see Young,T.J.,1987) in the loop of steps **2** to **4** but, since the subsequent analysis is not particularly sensitive to the NVR value, manual selection is normally quite adequate and helps to expose the nature of the trend component. For instance, since the estimate of the trend derivative is provided by IRWSMOOTH, it can be used to assess the long term variations in the trend and will often reveal features such as trade cycle effects. The selection of the NVR values in the DHRSMOOTH algorithm is discussed below in Section **6.1**.

6. IDENTIFICATION AND ESTIMATION OF THE DETRENDED DATA

Having estimated and removed any low frequency trend component on the data in steps 1 to 4 of the above SSD procedure, it is necessary to identify and estimate an appropriate model for the detrended data δ(k). The most appropriate model will tend to depend upon the application and the requirements of the analysis: as pointed out in Section **2**, the DHR model is clearly most useful for the estimation, adaptive forecasting and smoothing, including the seasonal adjustment, of heavily periodic data; while the GTF model has wider applicability to quasi-periodic and non-seasonal time-series.

6.1 The DHR Model and Seasonal Adjustment

Recursive identification and estimation of the DHR model is straightforward. In the case where the regression parameters are assumed constant, the normal **Recursive Least Squares** (RLS) algorithm (see Appendix 1) can be used. When the parameters are assumed time-variable, then it is simply necessary to represent the variations by the GRW model, with or without variance intervention as appropriate, and use the recursive least squares filtering and fixed interval smoothing algorithms outlined in Section 3. (see Young,1984).

In the stationary parameter case, the conventional, constant parameter harmonic regression model, estimated in the above fashion, can be combined with the GRW trend model to construct the complete state-space model (7) for the component model (1). Recursive state-space forecasting then follows straightforwardly by application of the estimation and forecasting equations (8), (9), (14) and (15). If a variable parameter DHR model is found to be necessary, then the same basic approach can be utilised but the variable parameters will also be estimated on the basis of appropriate GRW models. This automatically yields the self-adaptive version of the recursive forecasting equations.

If off-line analysis of the nonstationary time-series is required, then the recursive fixed interval smoothing equations (11) to (13) can be used to provide smoothed estimates of the structural model components and any associated time variable parameters. The main effect of allowing the parameters and, therefore, the amplitude and phase of the identified seasonal components to vary over time in this manner, is to include in the estimated seasonal component other frequencies components with periods close to the principal period. As pointed out above, the chosen NVR then controls the band of frequencies that are taken into account by the DHRSMOOTH algorithm (see T.J.Young, 1987). If it is felt that the amplitude variations in the seasonal component are related to some known longer period fluctuations (e.g. an economic cycle) then such prior knowledge could be used to influence the choice of the NVR.

If the DHR model is identified and estimated for all the major periodic components identified in the data (i.e., those components which are associated with the peaks in the periodogram or AR spectrum) then the DHRSMOOTH algorithm can be used to construct and remove these "seasonal" components in order to yield a "seasonally adjusted" data set (see T.J.Young et al., 1988). When carrying out seasonal adjustment in this manner, it is advisable to estimate the GRW trend model and the DHR model simultaneously, since this tends to reduce "end effects" introduced by the smoothing algorithm. This is equivalent to the situation in ordinary HR model estimation where a constant term is introduced to allow for a non-zero mean series (see Appendix 1). In the present context, of course, the mean value is allowed to vary over the observation interval and represents the trend behaviour.

This kind of "adaptive" seasonal adjustment (SA) procedure is, of course, most important in the evaluation of business and economic data, where existing SA methods, such as the Census X-11 procedure (Shiskin et al, 1967; which uses a procedure based on centralised moving average (CMA) filters) are well established. We cannot review such methods here, except to point out that much debate has gone on about the validity of seasonal adjustment and its vulnerability to abuse. In this connection, we feel that the proposed DHR-based approach may offer various advantages over techniques such as X-11: as we shall see in the later example (Section 7.), it can handle large and sudden changes in the dynamic characteristics of the series, including amplitude and phase changes; it is not limited to the normally specified seasonal periods (i.e. annual periods of 12 months or 4 quarters); and it is more objective and simpler to apply in practice (see Young,T.J.,1987; Ng,1987).

6.2 Recursive Identification and Estimation of the AR or ARMA model

Since the appearance of Box and Jenkins book (1970) on time-series analysis, forecasting and control, the ARMA and ARIMA models have become the accepted, standard representations of stochastic time-series. However, the success of these general models has tended to overshadow the many attractive features (see e.g. Priestley,1981) of the simpler AR and subset AR models. As in the case of the DHR model, it is well known that the RLS algorithm, coupled with the AIC (or similar) identification criterion, yields asymptotically unbiased recursive estimates of the parameters in the AR model. And if statistically insignificant parameters are identified during this initial RLS analysis, then the equivalent subset model can be estimated using the same RLS algorithm with the insignificant parameters constrained to zero in the normal manner (see Young,1984).

The resulting AR or subset models may be less parametrically efficient than the equivalent ARMA models, but they still provide powerful representations of general stochastic behaviour. For example, *unstable* AR(14) or subset AR(14) models of the detrended airline passenger data (i.e models whose eigenvalues lie outside the unit circle of the complex plane) are both easy to estimate and, when combined with the IRW trend model, yield forecasting performance that appears, in this case, to be superior to that obtained originally by Box and Jenkins using their ARIMA model (Ng, 1987). We have found that this ability to easily identify and characterise mildly unstable behaviour is an attractive feature of the AR model. Moreover, as we have pointed out above, it is well known that the spectrum of the high order AR model normally provides a very good spectral description of time-series data, being equivalent to the maximum entropy spectrum.

Recursive estimation of the ARMA model is not so straightforward but it can be achieved using various algorithms, as discussed by Young (1984). These include: the **Prediction Error Recursion** (PER) method (see e.g. Ljung and Soderstrom,1983); the **Approximate Maximum Likelihood** (AML) method (see Young, 1968,1976,1984); and the two step procedure based on initial AR identification, followed by recursive refined (i.e. optimal) instrumental variable estimation, as proposed by Young (1985). For off-line analysis, we have found the latter approach to be the most useful since the first stage AR model is often satisfactory for most practical purposes.

7. A PRACTICAL EXAMPLE: ANALYSIS OF THE LEAVERS DATA

The school leavers unemployment data set shown in Fig.3 is an excellent example of a nonstationary economic time-series. As we have pointed out, this 184 sample data set is clearly influenced heavily by the socio-economic and political changes over the period 1970 to 1984 and it seems reasonable to conjecture, therefore, that the obvious nonstationarity of both the mean and the seasonal components are functions of these factors. With this in mind, we can assume that the sensitivity of the recursive estimation algorithms should be chosen to account for factors such as: the possible presence of an "economic" or "business" cycle; the occurence of major world events, like the oil crisis of the 1970's; and the effects of the change in U.K. government after 1979. In the latter case, the new Conservative administration's modification of the unemployment registration regulations affecting young people clearly led to dramatic changes in the amplitude *and* phase of the series after 1979. Such factors are taken into account in the analysis reported in the next two sub-sections, most of which was carried out using the **microCAPTAIN** micro-computer program (Young and Benner, 1988)

7.1 Sequential Spectral Decomposition and Adaptive Seasonal Adjustment

The IRWSMOOTH estimate of the low frequency trend is shown in Fig.3(a) for an NVR=0.0001and variance intervention introduced at sample 55 to account for the significant change in the level of the series at this point. We see from equation (17) that this NVR value yields $F_{50} = 0.158$ which, for monthly data, corresponds to a 50% attenuation of components with periods less than about 5 years in length. In this manner, we allow the estimated trend to account for any "economic cycle" behaviour but ensure that the detrended data, shown in Fig.6(a), contains all of the information on the important 12 monthly annual cycle (note that, for NVR=0.0001, F_{95}=0.0316, so ensuring 95% attenuation of all components with periods less than about 2.5 years). The variance intervention is introduced simply by boosting the NVR associated with the second equation of the IRW model from zero to 100 *only* at the 55th sample point where the significant change in amplitude is seen to occur.

Fig. 6(b) shows the amplitude periodogram of the detrended data. We see, by reference to Fig.3(b), that the IRWSMOOTH detrending has adequately removed the longer period behaviour and left the seasonal pattern, which is composed of the 12 monthly component and its associated harmonic components at 6, 4, 3, 2.4, and 2 months, respectively. Each of these harmonic components is now estimated *separately and sequentially* using the DHRSMOOTH algorithm based on an RW model with NVR=0.01 in each case, and with variance intervention introduced at samples 55 and 124 to reflect the dramatic changes in amplitude at thes sample points. The RW model was chosen to model the parameter variations since, as we shall see, the amplitude changes occur mainly at the intervention points, with much smaller changes elsewhere. The NVR=0.01 was selected for the RW model since it produces a 50% cut-off frequency about the same as that for the IRW model with an NVR=0.0001, so ensuring that the estimator responds to any medium term amplitude modulation of the seasonal components associated with factors such as economic cycle behavior.

The total seasonal component, as obtained by summing all of the separately estimated harmonic components, and its amplitude periodogram are presented in Figs. 7(a) and (b), respectively. Figs. 8(a) and (b) show the estimate of the first seasonal component (12 months) together with the A(k) amplitude variations. Clearly, the seasonality in the series has been modelled satisfactorily, with the large changes in amplitude at the selected intervention points and the much smaller variations between these points both captured well. This demonstrates the utility of the proposed methodology in allowing for such a mix of widely different types of nonstationarity. Notice also that the procedure has accounted for the change in phase at the second intervention point caused by the changed registration procedures.

The non-seasonal or "irregular" component, as obtained by subtracting the the trend and seasonal components from the original data, is given in Fig.9(a), with the associated periodogram in Fig.9(b): we see that it contains some low frequency serial correlation and could be modelled as an AR or ARMA process, if required. It might also be considered as an estimated "anomaly" series, revealing medium term, non-seasonal perturbations about the smooth, long term trend (we comment further on this series in Section 7.3). When this non-seasonal component is added to the estimated trend, we obtain the seasonally adjusted series shown in Fig.10.

7.2 Adaptive Forecasting and Smoothing

In order to obtain multi-step-ahead forecasts, a complete state-space model of the form shown in equation (7) is formed based on the IRW and DHR models with interventions, as discussed in **7.1**. For the sake of presentation, however, we only model the 12 and 6 month period seasonal components. Fig.11(a) shows the one-step-ahead predictions up to the user-specified forecasting origin, and up to 24 step-ahead forecasts beyond this origin. Two interventions are introduced at samples 55 and 124 but, this time, applied both to the trend and seasonal processes. Fig.11(b) is a magnified plot of the forecasting errors and the associated 2 S.E. boundaries: it is noteworthy that these standard error bounds widen sufficiently in the region of the interventions to allow for the sudden but short-lived increase in the prediction errors over these regions. The forecast of the complete seasonal component, as plotted in Fig.12(a), is obtained simply by summing the estimated 12 and 6 month harmonics. The associated estimates of the A(k) amplitude variations for the two harmonics is given in Fig.12(b). These results show how successfully the estimates adapt to the sudden changes in both the amplitude and phase of the seasonal component, even though each prediction in this filtering (in contrast to smoothing) process is only based on the information up to the current data point.

Of course, it is possible to refine the filtering estimates obtained up to the forecasting origin by resort to fixed interval smoothing. In the present context, this is achieved most conveniently by applying the smoothing equations (11)-(13) directly to the *complete* state-space model. Figs.13 and 14 show the results obtained in this manner: it is interesting to note how the strong perturbations in the estimates and forecasts around 1980 during the forecasting run, as induced by the phase change in the data (see particularly Fig.12(a)), are removed during the smoothing run. Also, although we subjectively introduced a second intervention into the trend process, the estimator indicates that this is not really significant by introducing only a minimal change in level at this point.

7.3 Other Possibilities.

It must be stressed that the above analysis of the school leavers data is utilised here merely to exemplify the recursive procedures discussed in this paper and should not be considered, in any sense, as a final evaluation of the data. Indeed, the nature of the SA residuals in Fig.9(a), particularly around 1978-81, might suggest that some modification of the SA is required in this region: in particular, the the associated SA series in Fig.10 indicates possible under-adjustment before the intervention at sample 124 and over-adjustment following it. In a modified analysis of these data (T.J.Young et al,1988), therefore, we have softened the variance intervention effect by replacing the abrupt intervention at sample 124 by a more distributed and milder intervention , with the NVR on the seasonal components increased to only 0.1, but applied over a longer period between samples 118 and 124. This has the effect of transfering the nonstationarity more to the seasonal component and results in a seasonally adjusted series that is smoother in this region. Nevertheless, we do not expect that this will constitute the last word on this interesting data set.

8. CONCLUSIONS

In this paper, we have considered what we believe to be the first, fully recursive approach to the modelling, forecasting and seasonal adjustment of nonstationary time-series; an approach which seems particularly relevant to the analysis of socio-economic and business data. We have concentrated on the application of these techniques to univariate time-series, but they can be extended quite easily to multivariable (vector) processes. Ng et al(1988), for example, have used these multivariable procedures to model and forecast the monthly sales of competitive group of products from two organisations. This kind of multivariable analysis, which is based on recursively estimated vector AR or ARMA models of the multiple time-series (see Wang and Young,1988), is clearly of potential importance in a socio-economic context, where the variables are highly interactive and there is the strong possibility of ill-defined feedback connections with uncertainty about the direction of causation. The additional dimension of smoothed TVP estimation should be particularly helpful in allowing the analyst to examine the assumption that the model parameters may change over time in response to nonstationarity in the system's dynamic characteristics. Recent multivariable analysis of quarterly US macro-economic data has yielded promising results in this regard for low dimensional models, but the difficulty of extending the procedures to high dimensional model forms should not be underestimated.

Finally it should be noted that the adaptive seasonal adjustment procedure proposed in the paper is still at its first stages of development and requires various enhancements to allow for factors such as sampling period variations, holidays and festival effects, and the minimisation of end effects in SA revisions. Until such enhancements are introduced, it cannot properly be compared with existing well tried and tested procedures such as the Census X-11.

ACKNOWLEDGEMENTS

This paper was completed by the first author while he was a visitor in the Institute for Empirical Macroeconomics at the Federal Reserve Bank of Minneapolis. He is grateful to the Institute for its support during this period and to the members of the Institute for the making the visit so intellectually stimulating.

REFERENCES

Akaike,H. (1974). A new look at statistical model identification. *I.E.E.E. Trans. on Aut. Control*, AC-**19**, 716-722.

Astrom,K.J. (1970). *Introduction to Stochastic Control Theory.* Academic Press: New York.

Bell, W.R., and Hillmer, S.C. (1984). Issues involved with the seasonal adjustment of economic time series. *Jnl. Bus. and Econ. Stat.*, **2**, 291-320.

Box G.E.P., and Jenkins,G.M.(1970). *Time Series Analysis, Forecasting and Control.* Holden-Day: San Francisco.

Harrison, P.J. (1967). Exponential smoothing and short term sales forecasting. *Management Science*, **13**, 821-842.

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

Harrison,P.J., and Stevens,C.F. (1976). Bayesian forecasting. *Jnl. Royal Stat. Soc., Series B.*, **38**, 205-247.

Harvey,A.C. (1981). *The Econometric Analysis of Time Series.* John Wiley: London.

Harvey,A.C. (1984). A unified view of statistical forecasting procedures. *Jnl. of Forecasting*, **3**, 245-275.

Harvey,A.C., and Durbin,J. (1986). The effects of seat belt legislation on British road casualties: a case study in structural time-series modelling. *Jnl. Royal Stat. Soc., Series A*, **149**, 187-210.

Harvey,A.C., and Peters,S. (1984). Estimation procedures for structural time-series models. *London School of Economics, Discussion Paper No. A28*.

Jakeman,A.J., and Young,P.C. (1979,1984). Recursive filtering and the inversion of ill-posed causal problems. *Utilitas Mathematica*, **35**, 351-376, 1984 (appeared originally as *Report No. AS/R28/1979, Centre for Resource and Environmental Studies, Australian National University, 1979*)

Jun,D.B. (1988). On detecting and estimating major level or slope change in general exponential smoothing. *Jnl. of Forecasting*, to appear 1988.

Kaldor,J. (1978). *The Estimation of Parametric Change in Time-Series Models.* M.A. Thesis, Australian National University.
Kalman,R.E.(1960). A new approach to linear filtering and prediction problems. *ASME Trans.,Jnl. Basic Eng.*, **83-D**, 95-108.
Kitagawa,G. (1981). A non-stationary time-series model and its fitting by a recursive filter. *Jnl. of Time Series*, **2**, 103-116.
Kitagawa,G., and Gersch,W. (1984). A smoothness priors state-space modelling of time series with trend and seasonality. *Jnl. American Stat. Ass.*, **79**, 378-389.
Ljung,L., and Soderstrom,T. (1983). *Theory and Practice of Recursive Estimation.* MIT Press: Cambridge, Mass.
Nerlove, M., Grether, D.M., and Carvalho, J.L. (1979). *Analysis of Economic Time Series : A Synthesis*. Academic Press: New York.
Ng,C.N. (1987). *Recursive Identification, Estimation and Forecasting of Non-Stationary Time-Series.* Ph.D Thesis, Centre for Research on Environmental Systems, Univ. of Lancaster, England.
Ng,C.N., and Young,P.C. (1988). Recursive estimation and forecasting of nonstationary time-series, to appear *Jnl. of Forecasting*, 1989.
Ng.C.N., Young,P.C., and Wang,C.L. (1988). Recursive identification, estimation and forecasting of multivariate time-series, to appear in *Proc. IFAC Symposium on Identification and System Parameter Estimation, Beijing, China, 1988*
Norton,J.P.(1975). Optimal smoothing in the identification of linear time-varying systems. *Proc. I.E.E.*, **122**, 663-668.
Priestley,M.B. (1981). *Spectral Analysis and Time Series.* Academic Press: London.
Schweppe,F. (1965). Evaluation of likelihood function for Gaussian signals. *I.E.E.E. Trans. on Inf. Theory*, **11**, 61-70.
Shiskin,J., Young,A.H., and Musgrave,J.C. (1967). The X-11 variant of the Census Method 11 seasonal adjustment program. *Tech. Paper No. 15, U.S Dept. of Commerce, Bureau of Economic Analysis.*
Tsay,R.S. (1988). Outliers, level shifts and variance changes in time series. *Jnl of Forecasting*, **7**, 1-20.
Whittle,P. (1952). Tests of fit in time series. *Biometrica*, **39**, 309-318.
Young,P.C. (1968). The use of linear regression and related procedures for the identification of dynamic processes. *Proc. 7th I.E.E.E. Symposium on Adaptive Processes.*
Young,P.C. (1969). *The Differential Equation Error Method of Process Parameter Estimation.* Ph.D. Thesis, University of Cambridge, England.
Young,P.C. (1970). An instrumental variable method for real time identification of a noisy process. *Automatica*, **6**, 271-287.
Young,P.C.(1971). A second generation adaptive pitch autostabilisation system for a missile or aircraft. *Tech. Note 404-109,Naval Weapons Centre, China Lake, California (100 pages).*
Young,P.C. (1976). Some observations on instrumental variable methods of time-series analysis. *Int. Jnl. of Control*, **23**, 593-612.
Young,P.C. (1981). A second generation adaptive autostabilisation system for airborne vehicles. *Automatica*, **17**, 459-469.
Young,P.C. (1984). *Recursive Estimation and Time-Series Analysis.* Springer-Verlag: Berlin.
Young,P.C. (1985). The instrumental variable method: a practical approach to identification and system parameter estimation, appears in H.A.Barker and P.C.Young (eds.) *Identification and System Parameter Estimation 1985, Vols 1 and 2*, Pergamon press: Oxford.
Young,P.C. (1987). Recursive estimation, forecasting and adaptive control, chapter to appear in C.T.Leondes (ed) *Control and Dynamic Systems, Vol. XXXii,* Academic Press: Florida.
Young,P.C. (1988). Recursive extrapolation, interpolation and smoothing of nonstationary time-series. *Proc. IFAC Symposium on Identification and System Parameter Estimation, Beijing, China, 1988*
Young,P.C., and Benner,S. (1988). *microCAPTAIN Handbook : Version 2.0.*
Young,P.C, and Ng,C.N. (1988). Variance intervention, *Jnl. of Forecasting*, to appear, 1989.
Young,P.C., and Wang,C.L. (1988). Identification and estimation of multivariable dynamic systems, chapter in J. O'reilly (ed.) *Multivariable Control for Industrial Applications*,. Peter Perigrinus: London.
Young,T.J. (1987). *Recursive Methods in the Analysis of Long Time Series in Meteorology and Climatology*. Ph.D. Thesis, Centre for Research on Environmental Systems, Univ. of Lancaster, England.
Young,T.J., Ng.C.N., and Young,P.C.(1988). Seasonal adjustment by optimally smoothed time-variable parameter estimation. *Proc. IFAC Symposium on Identification and System Parameter Estimation, Beijing, China, 1988 .*

APPENDIX 1 : THE RECURSIVE LEAST SQUARES (RLS) and INSTRUMENTAL VARIABLE (RIV) ALGORITHMS

The RIV algorithm can be written in the following form:

$$\hat{\mathbf{a}}(k) = \hat{\mathbf{a}}(k-1) + \mathbf{P}(k-1)\,\hat{\mathbf{z}}(k)\,[\, r + \mathbf{z}(k)^T\mathbf{P}(k-1)\,\hat{\mathbf{z}}(k)\,]^{-1}\,\{\, y(k) - \mathbf{z}^T\,\hat{\mathbf{a}}(k-1)\,\}$$

$$\mathbf{P}(k) = \mathbf{P}(k-1) + \mathbf{P}(k-1)\,\hat{\mathbf{z}}(k)\,[\, r + \mathbf{z}(k)^T\mathbf{P}(k-1)\,\hat{\mathbf{z}}(k)\,]^{-1}\,\mathbf{z}^T\,\mathbf{P}(k-1)$$

where,

$$\hat{\mathbf{a}}(k) = [\,\hat{a}_1\ \hat{a}_2\ .....\ \hat{a}_n\,]^T \ ;\ \mathbf{z}(k) = [\, z_1\ z_2\ .....\ z_n\,]^T \ ;\ \hat{\mathbf{z}}(k) = [\,\hat{z}_1\ \hat{z}_2\ .....\ \hat{z}_n\,]^T$$

In this algorithm, y(k) is the observation of the "dependent variable; r is a scalar constant; **P**(k) is a nxn matrix, $\hat{\mathbf{a}}(k)$ is the estimate at the kth recursion of the parameter vector **a**(k), as defined for the model under consideration; **z**(k) is the data vector associated with this model; and $\hat{\mathbf{z}}(k)$ is the **instrumental variable vector** associated with the data vector **z**(k). In the case where $\hat{\mathbf{z}}(k)$ is set equal to **z**(k) and r=1.0 , the algorithm becomes the RLS algorithm and **P**(k) is then symmetric.

Special cases of this algorithm (see e.g. Young,1984) are:-

The Recursive Regression Algorithm In this case, $\hat{\mathbf{z}}(k) = \mathbf{z}(k)$ is defined in terms of the n regression (or "independent") variables.

The Recursive Autoregression Algorithm In this case, $\hat{\mathbf{z}}(k) = \mathbf{z}(k)$ is defined as,

$$\mathbf{z}(k) = [\,-y(k-1)\ -y(k-2)\ .....\ -y(k-n)\,]^T.$$

The Recursive Harmonic Regression Algorithm In this case, $\hat{\mathbf{z}}(k) = \mathbf{z}(k)$ is defined in terms of the sine and cosine variables on the right hand side of equation (4), possibly with a constant term to allow for non-zero mean series.

The Recursive IV Algorithm for a Bivariate (Input-Output) Model Here,

$$\mathbf{z}(k) = [\,-y(k-1)\ -y(k-2)\ .....\ -y(k-n)\ u(k)\ u(k-1)\ .....\ u(k-n)\,]^T$$

$$\hat{\mathbf{z}}(k) = [\,-\hat{x}(k-1)\ -\hat{x}(k-2)\ .....\ -\hat{x}(k-n)\ u(k)\ u(k-1)\ .....\ u(k-n)\,]^T$$

where u(k) is the input variable, y(k) the *noisy* output variable, and $\hat{x}(k)$ an instrumental variable which is generated as an adaptive estimate of the *noise free* output x(k) of the system and is statistically independent of the observational errors on y(k). In the case of "Refined" or optimal IV algorithms, all the variables in the algorithm are adaptively prefiltered, as discussed in Young (1984).

In all these algorithms, if the observational errors are NID$(0,\sigma^2)$ and $r=\sigma^2$, then **P**(k) is an estimate of the covariance matrix of $\hat{\mathbf{a}}(k)$; while the "deterministic" versions of the algorithms are obtained with r=1. Dynamic versions of the algorithms are obtained by assuming that the elements of a(k) are all described by independent GRW models and modifying the algorithms accordingly (see Young,1984)

FIGURE CAPTIONS

Fig.1 The airline passenger series (Box and Jenkins, 1970)

Fig.2 The U.K. road casualty series (Harvey and Durbin, 1986). The Figure shows the series and estimated trend, plus the standard errors, as obtained from the IRWSMOOTH filter.

Fig.3 The school leaver's unemployment series: (a) series and estimated trend, as obtained from the IRWSMOOTH filter; (b) amplitude periodogram of the series.

Fig.4 Frequency response characteristics of the IRWSMOOTH filter for different values of the variance ratio (NVR).

Fig.5 Frequency response characteristics of the DHRSMOOTH filter for different values of NVR.

Fig.6 Detrended school leaver's unemployment series: (a) detrended series; (b) amplitude periodogram.

Fig.7 Total seasonal component of the school leaver's unemployment series: (a) total seasonal component; (b) amplitude periodogram.

Fig.8 First seasonal component (12 month period) of the school leaver's unemployment series: (a) first seasonal component; (b) amplitude A(k) for the first seasonal component.

Fig.9 Non-seasonal (residual) component of the school leaver's unemployment series: (a) non-seasonal component; (b) amplitude periodogram.

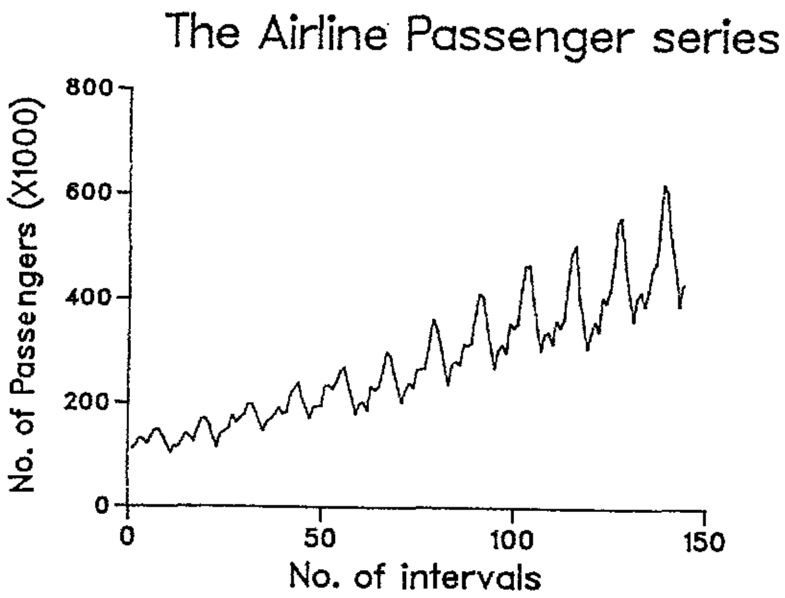
Fig.10 Seasonally adjusted school leaver's unemployment series compared with the original series.

Fig.11 Adaptive (TVP) forecasting of the school leaver's unemployment series: (a) forecasts of the series and trend component, showing one-step-ahead forecasts up to the *forecasting origin* and a two year ahead forecast from this origin; (b) forecasting residuals and associated standard errors.

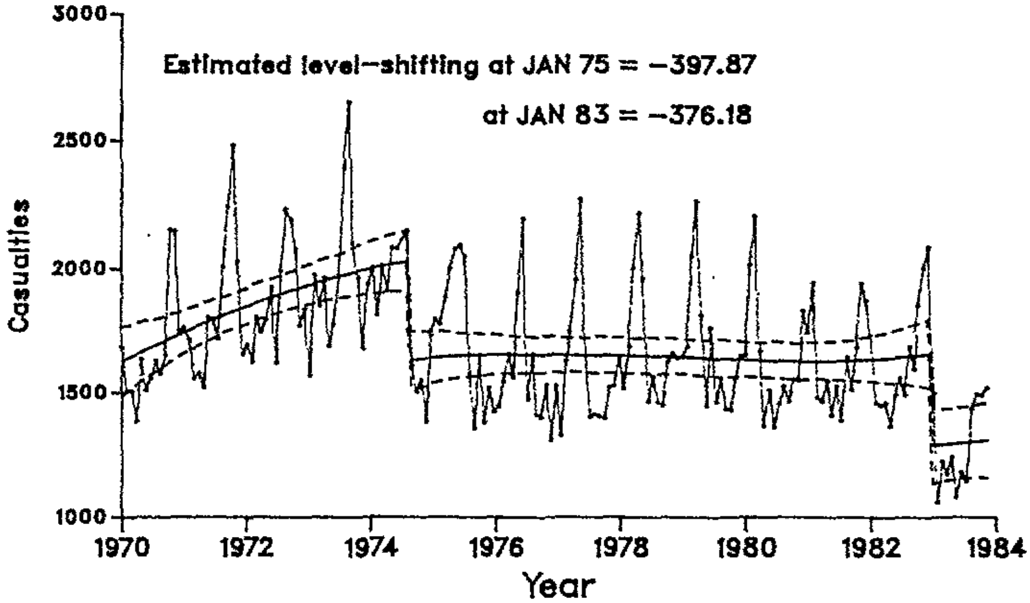
Fig.12 Adaptive (TVP) forecasting of the school leaver's unemployment series: (a) forecasts of the total seasonal component associated with Fig.11; (b) estimates and forecasts of the amplitude A(k) for the 12 and 6 month period components in (a).

Fig.13 Adaptive (TVP) smoothing of the school leaver's unemployment series: (a) smoothed estimate of the series and the trend component compared with the original data; (b) smoothing residuals.

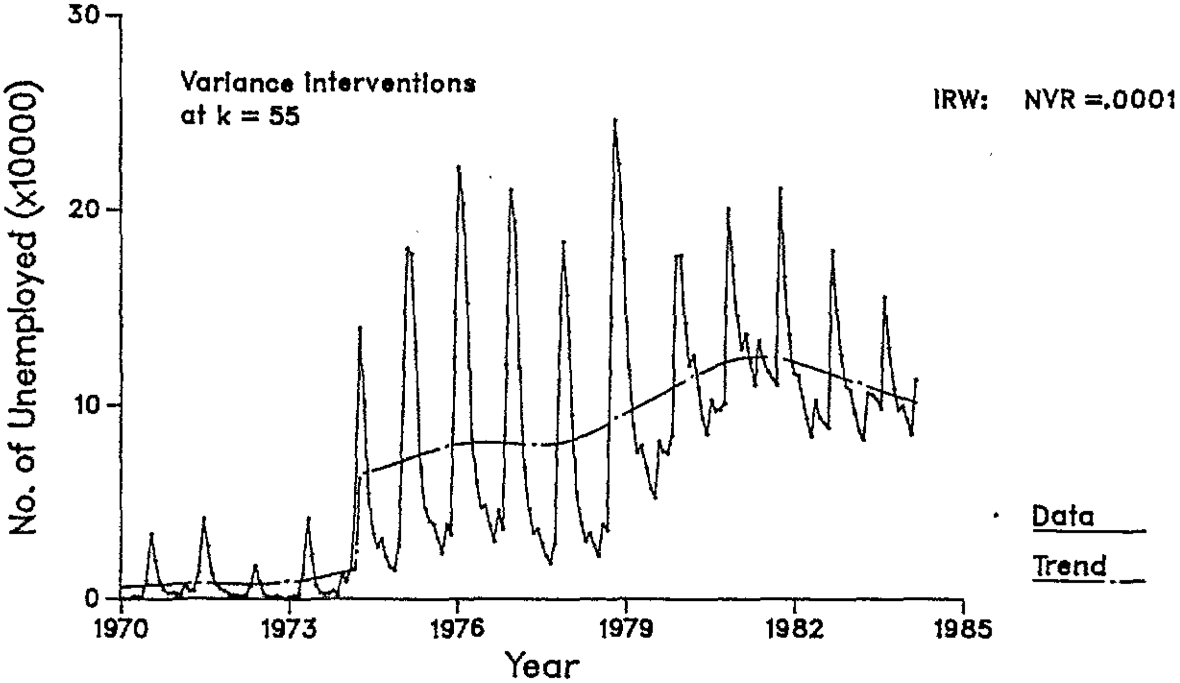
Fig.14 Adaptive (TVP) smoothing of the school leaver's unemployment series: (a) smoothed estimate of the total seasonal component associated with Fig.13; (b) smoothed estimates of the amplitude A(k) for the 12 and 6 month period components in (a).


The Airline Passenger series
No. of Passengers (X1000)
800
600
400
200
0
0
50
100
150
No. of intervals

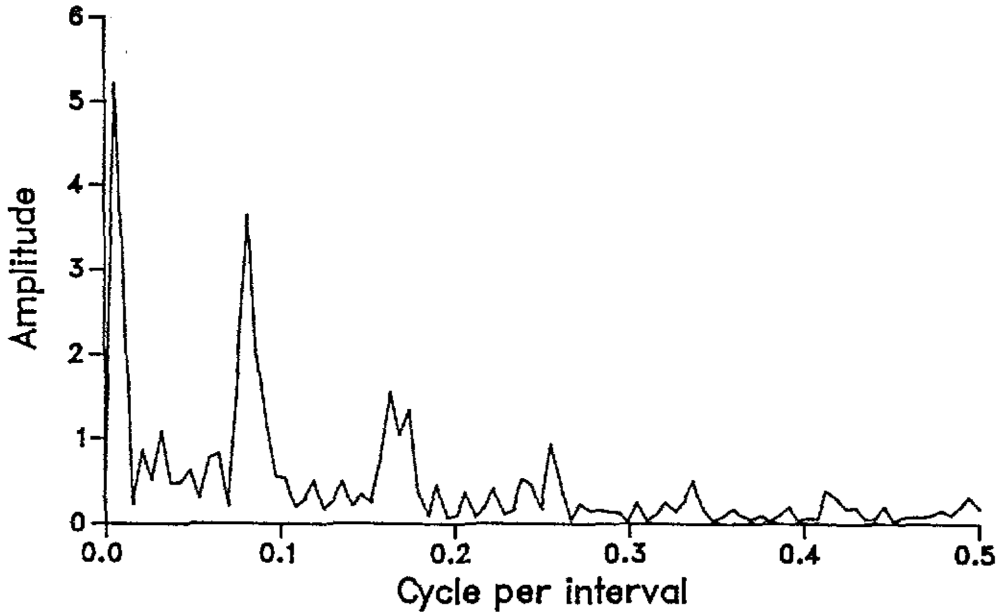
GRWSMOOTH estimation of the British road casualties series with variance interventions



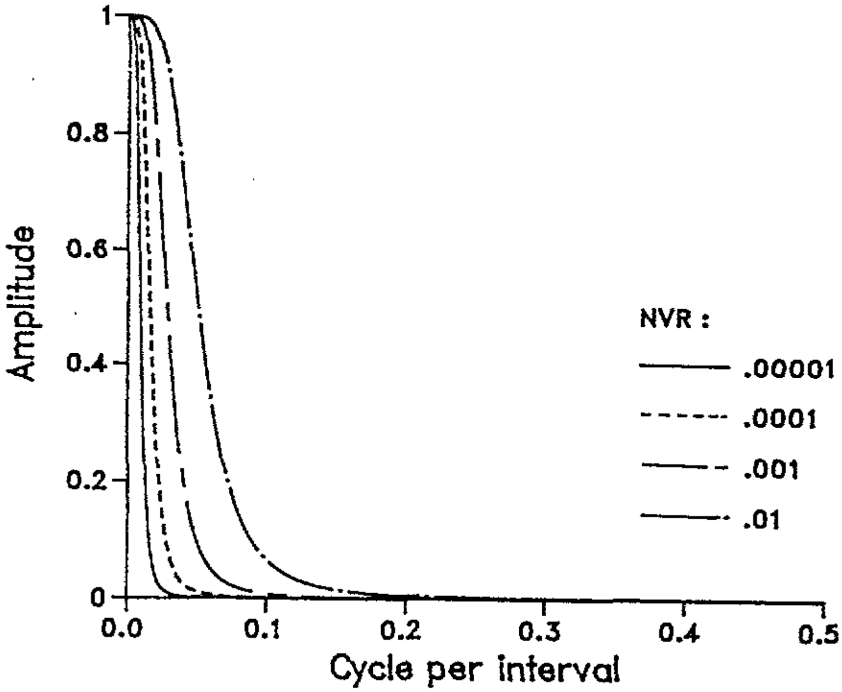
IRWSMOOTH trend of the School Leavers Unemployment series with intervention

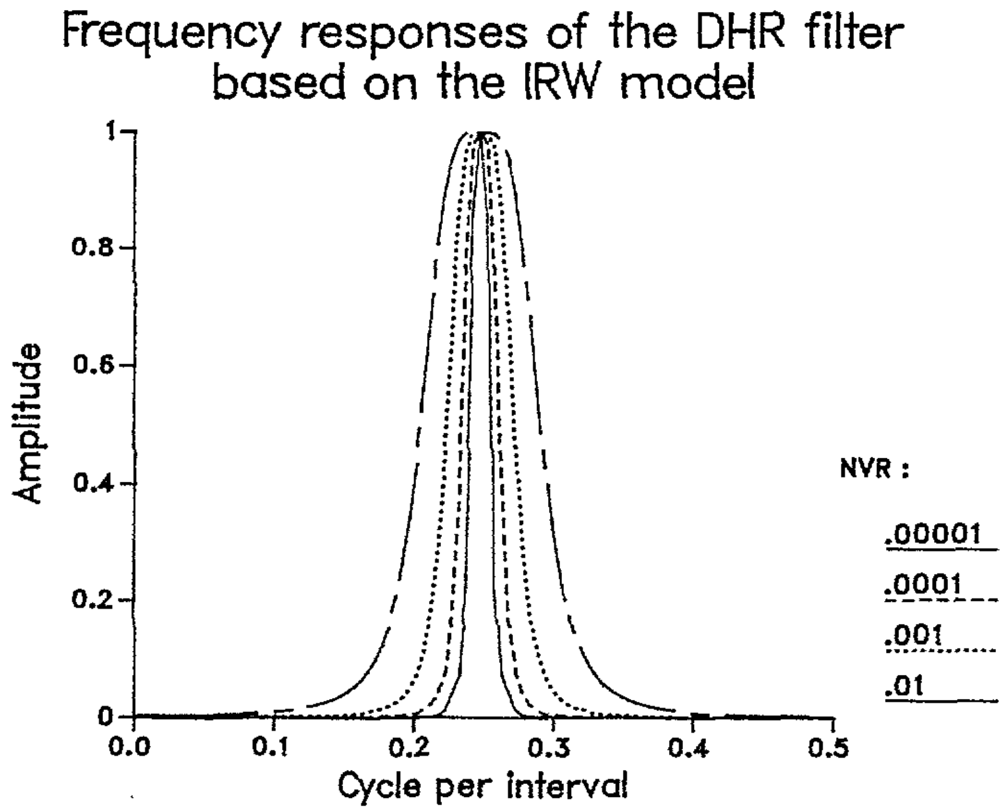


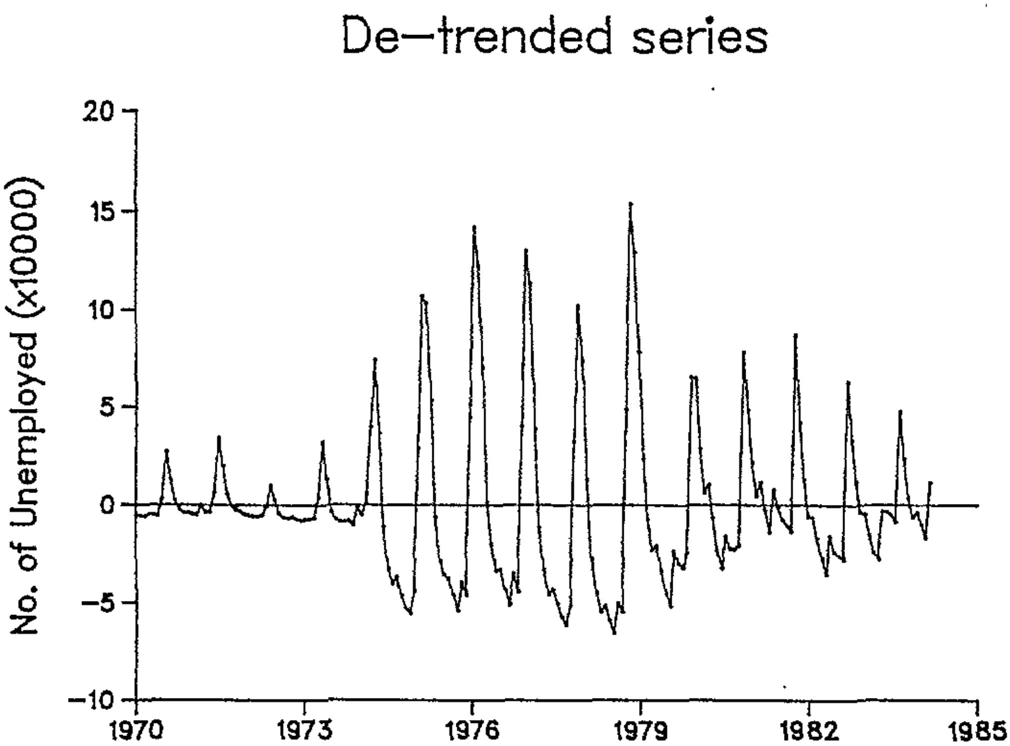
Amplitude spectrum of the original series

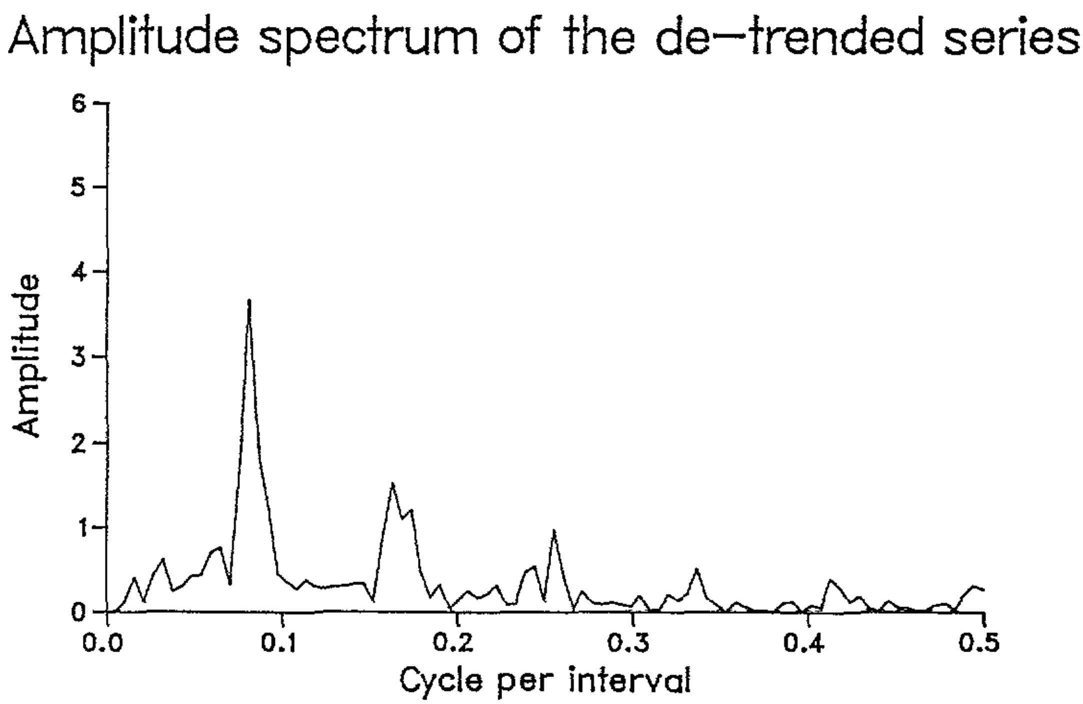


Frequency responses of the IRWSMOOTH filter

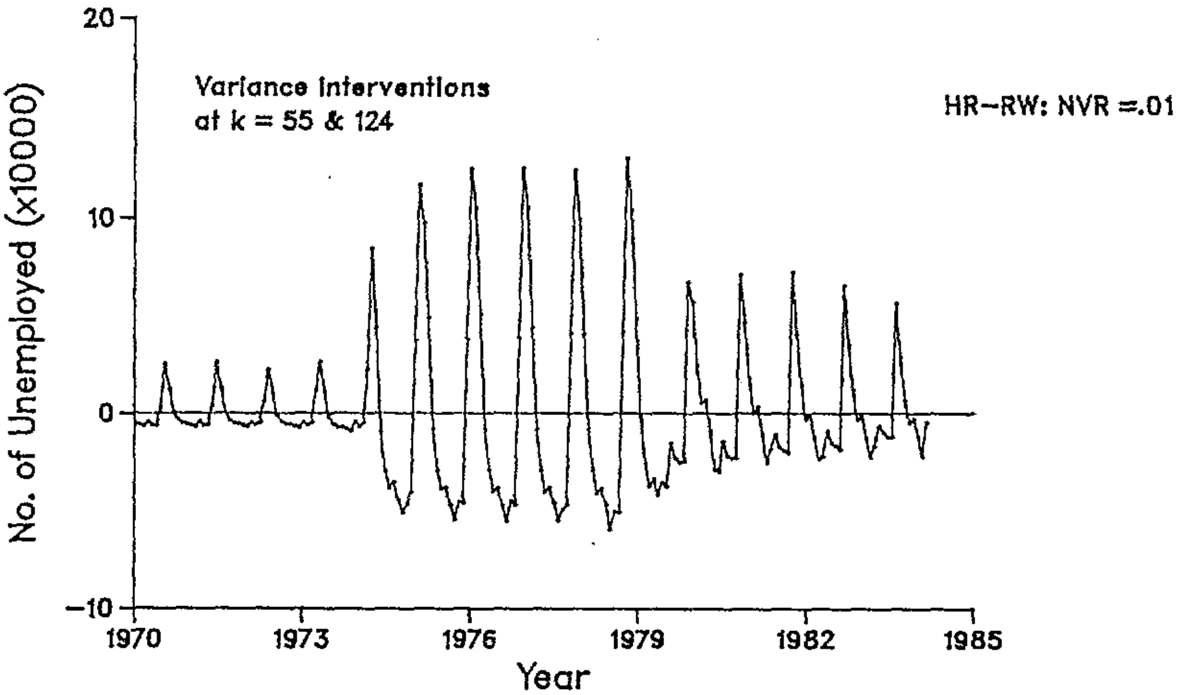



Frequency responses of the DHR filter
based on the IRW model
Amplitude
1
0.8
0.6
0.4
0.2
0
0.0
0.1
0.2
0.3
0.4
0.5
Cycle per interval
NVR :
.00001
.0001
.001
.01

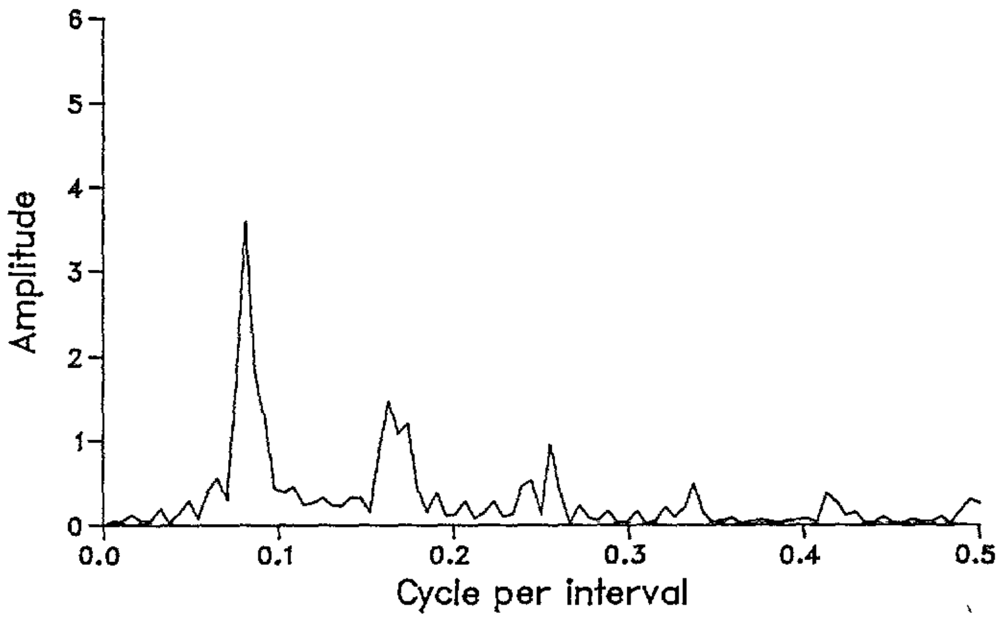

De–trended series
No. of Unemployed (x10000)
20
15
10
5
0
–5
–10
1970
1973
1976
1979
1982
1985


Amplitude spectrum of the de–trended series
Amplitude
6
5
4
3
2
1
0
0.0
0.1
0.2
0.3
0.4
0.5
Cycle per interval

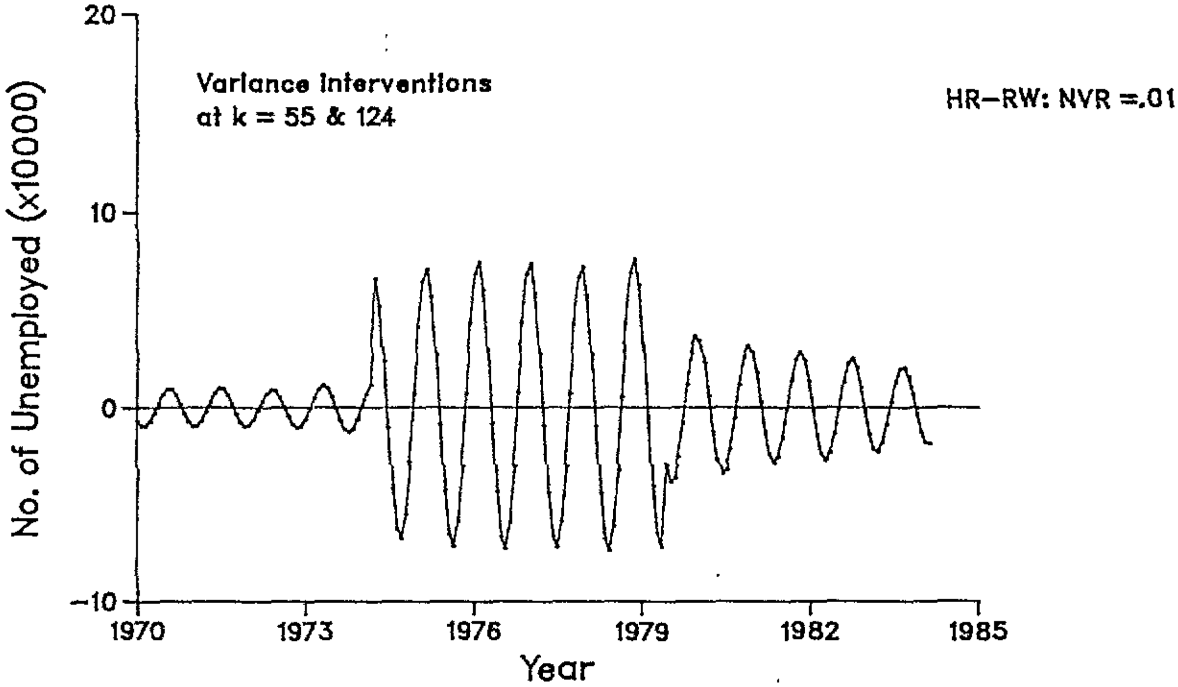
Seasonal component of the School Leavers Unemployment series with interventions



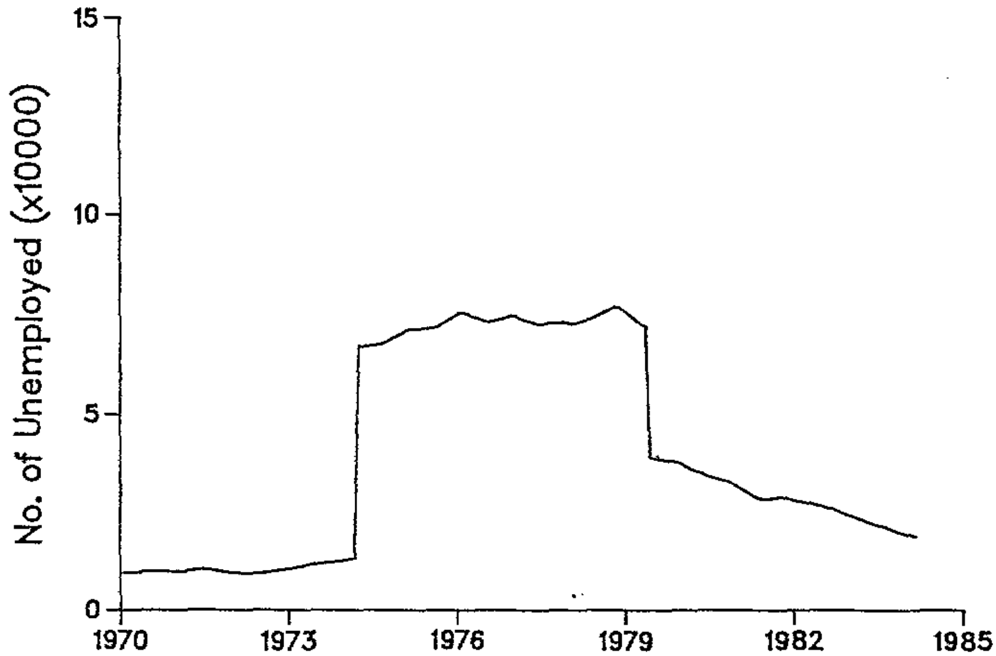
Amplitude spectrum of the seasonal component



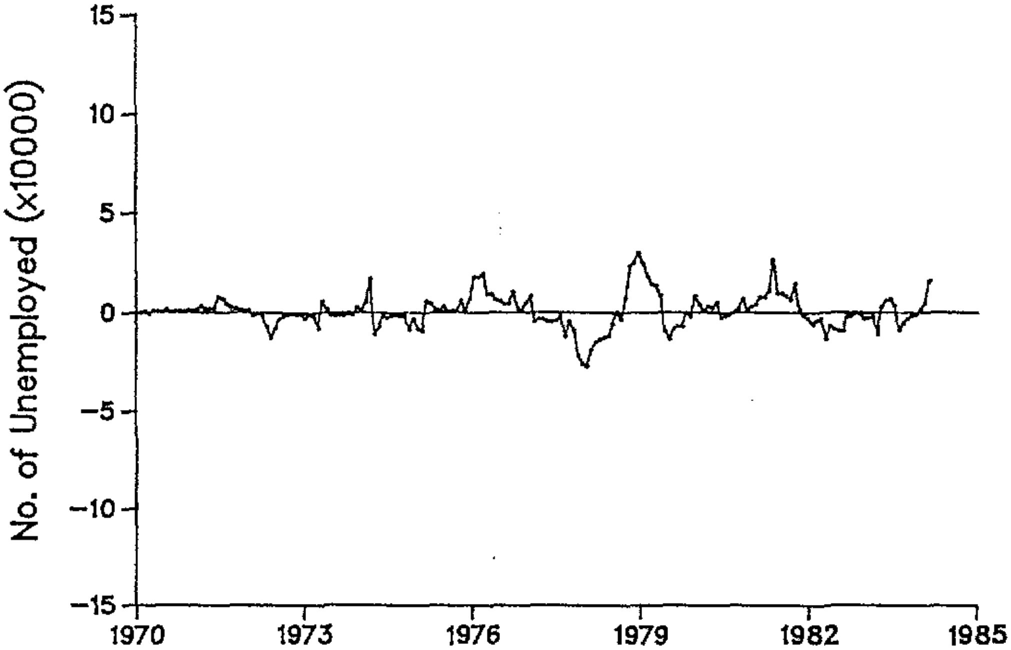
The 1st seasonal component of the School Leavers Unemployment series with interventions



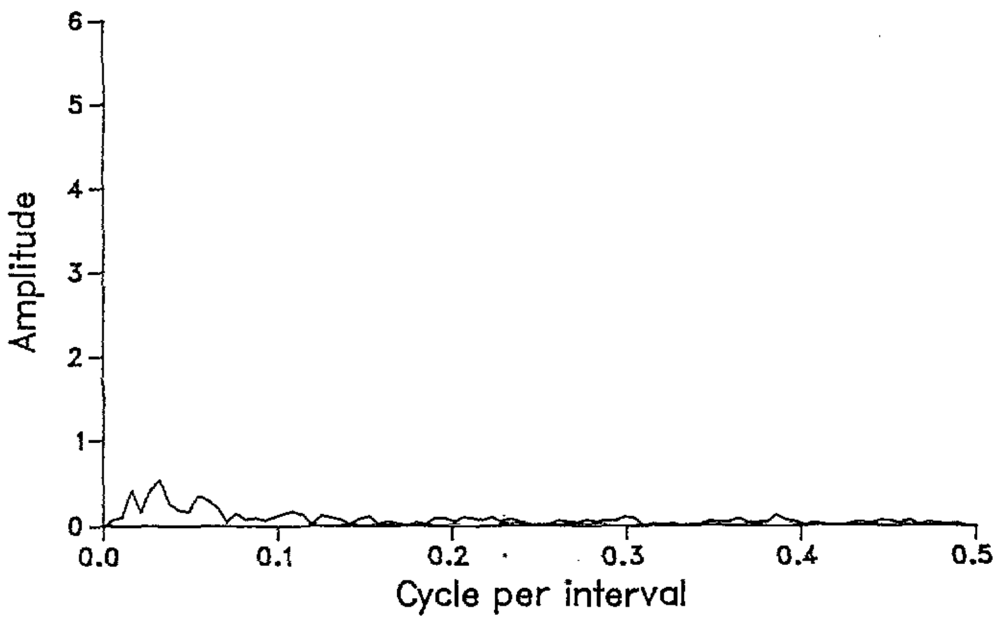
Amplitude of the 1st seasonal components

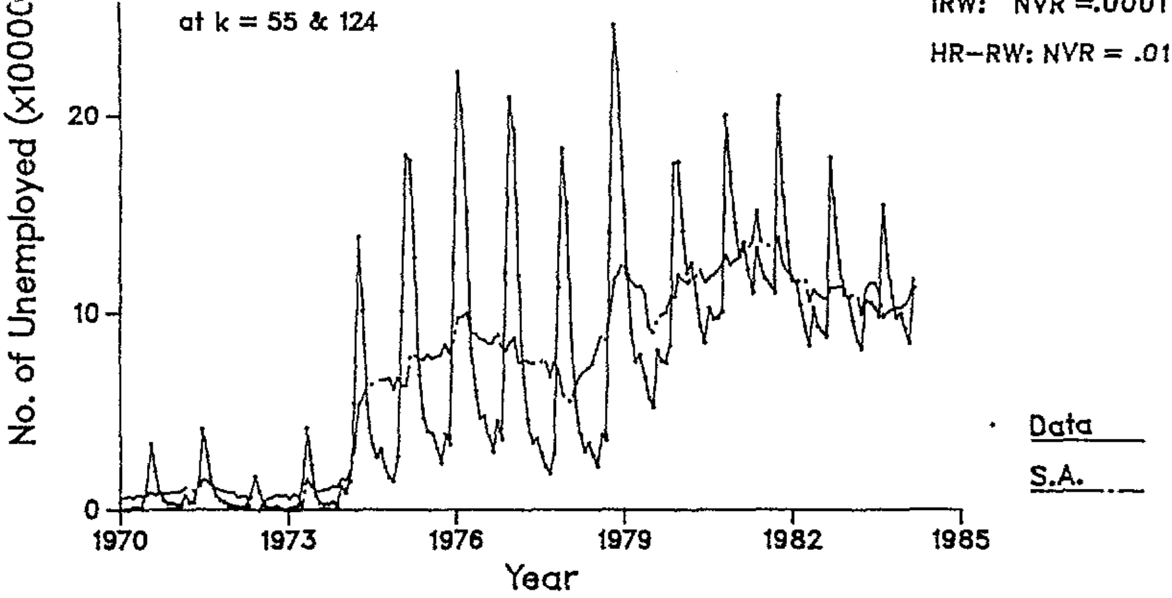


Non–seasonal component

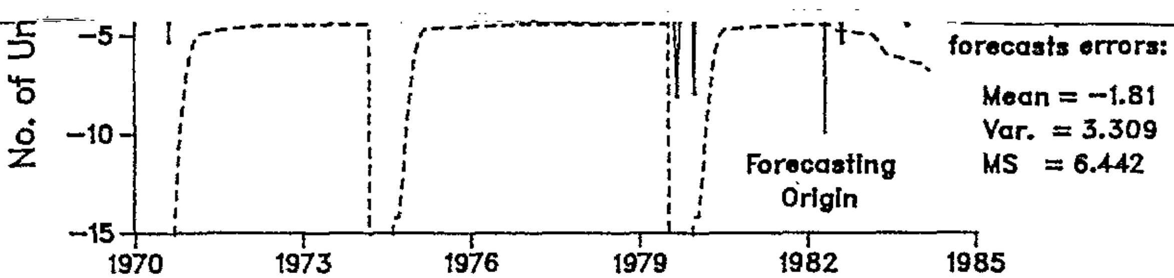


Amplitude spectrum of the non–seasonal component



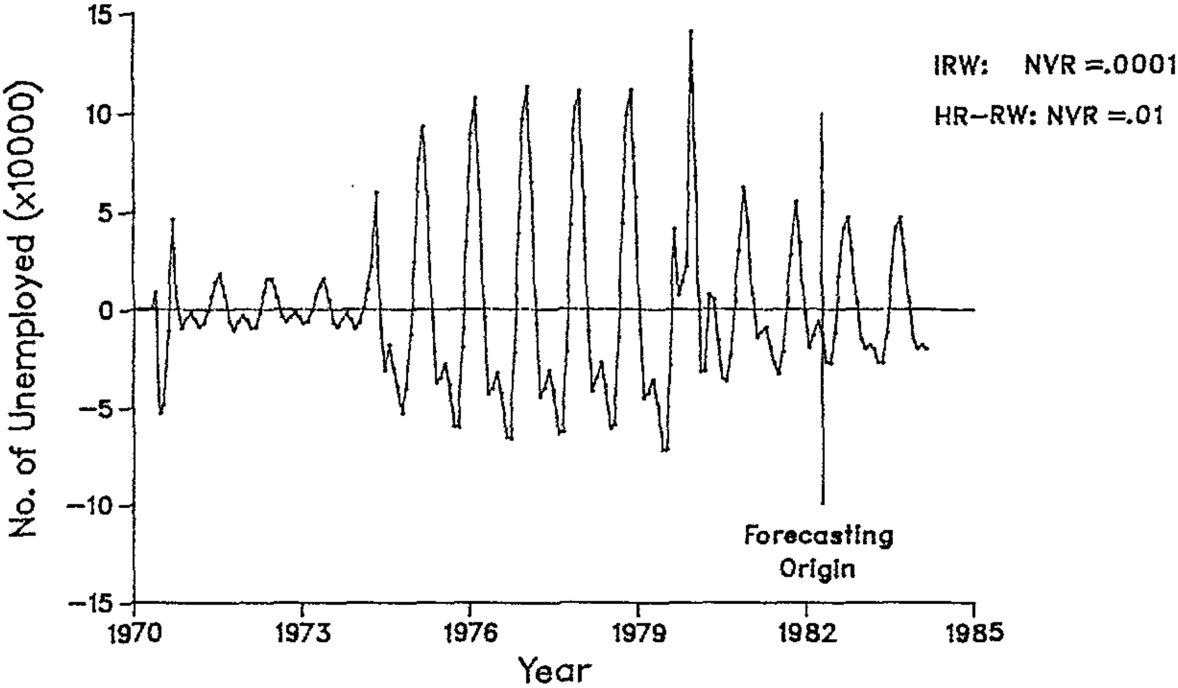


at k = 55 & 124
IRW: NVR =.0001
HR–RW: NVR = .01
No. of Unemployed (x1000C
0
10
20
1970
1973
1976
1979
1982
1985
Year
Data
S.A.

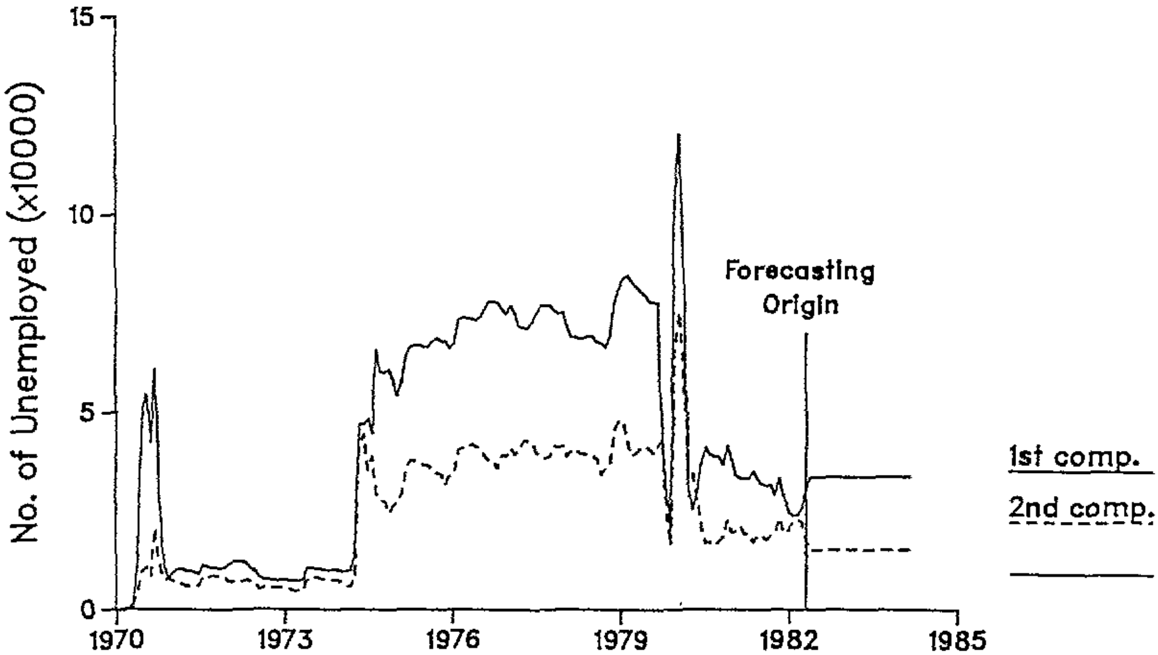


No. of Un
-5
-10
-15
1970
1973
1976
1979
1982
1985
Forecasting Origin
forecasts errors:
Mean = -1.81
Var. = 3.309
MS = 6.442

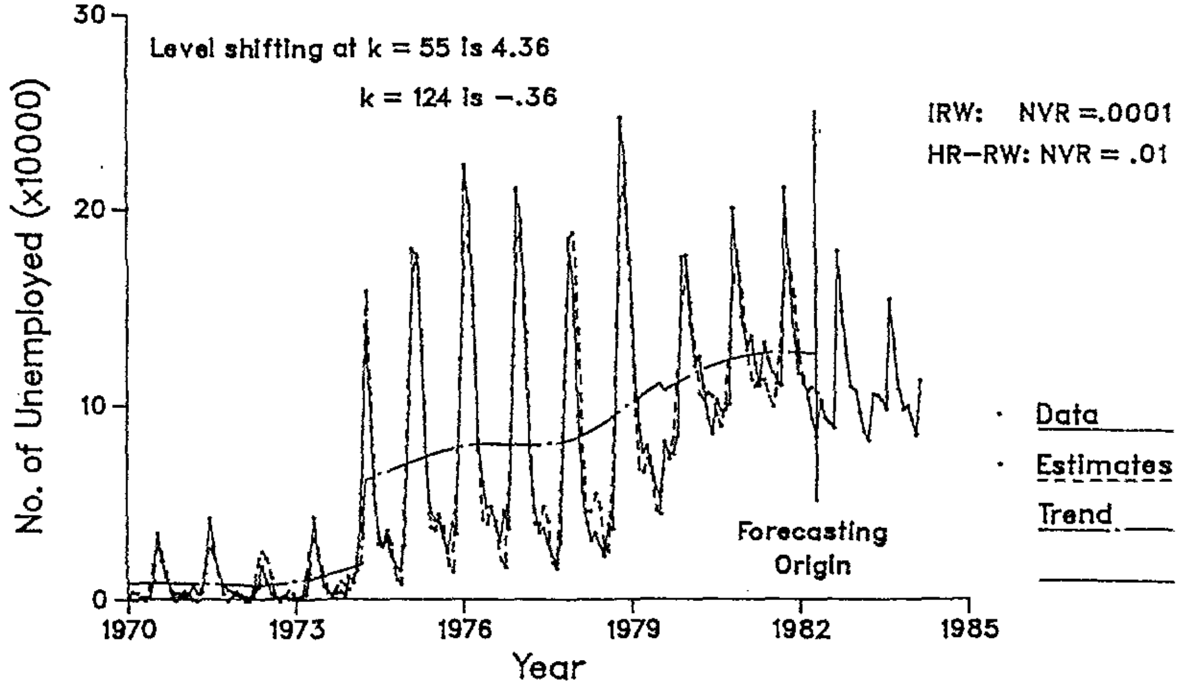
Seasonal components of the School Leavers Unemployment series based on the IRW + HR–RW model



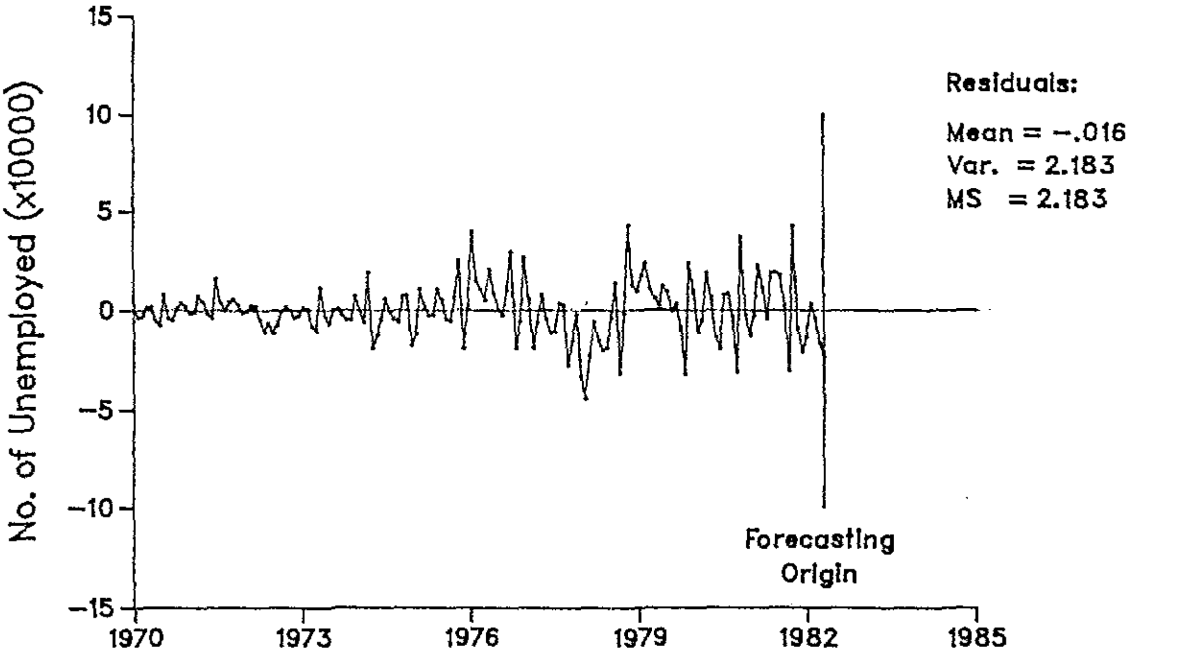
Amplitude of the seasonal components



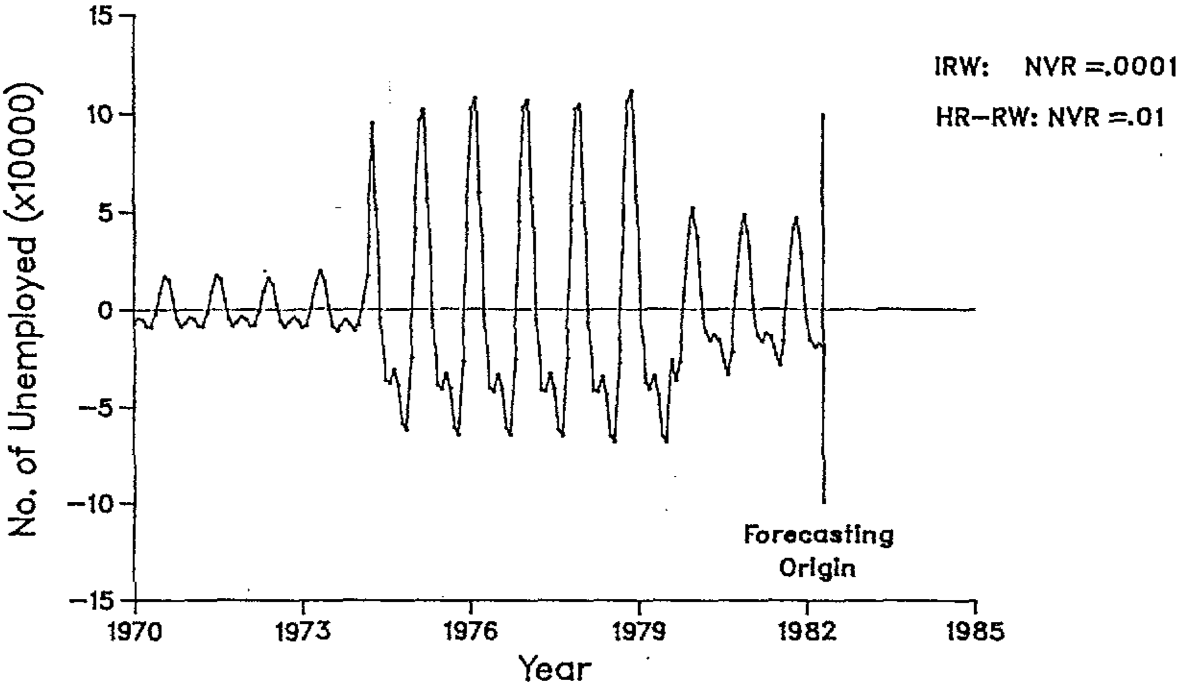
KF smoothing of the School Leavers Unemployment series based on the IRW + HR–RW model



Residuals of smoothing



KF smoothed estimates of the seasonal components of the School Leavers Unemployment series based on the IRW + HR–RW model



Amplitude of the seasonal components