Survey Methodology, June 1989
Vol. 15, No. 1, pp. 29–45
Statistics Canada

Modelling and Estimation for Repeated Surveys

D.A. BINDER and J.P. DICK[1]

ABSTRACT

Estimation of the means of a characteristic for a population at different points in time, based on a series of repeated surveys, is briefly reviewed. By imposing a stochastic parametric model on these means, it is possible to estimate the parameters of the model and to obtain alternative estimators of the means themselves. We describe the case where the population means follow an autoregressive-moving average (ARMA) process and the survey errors can also be formulated as an ARMA process. An example using data from the Canadian Travel Survey is presented.

KEY WORDS: Kalman filter; Overlapping surveys; State-space models; Time series modelling; Small area estimates.

1. INTRODUCTION

When surveys with similar data items are conducted on repeated occasions, certain estimation and data analysis methods are available which are not possible with single occasion surveys. For example, efficient estimation methods for the current occasion can depend on data from previous occasions. This occurs when there are overlapping sampling units between occasions and, hence, the survey errors can be correlated over time. As well, the series of estimates from a repeated survey are often modelled by the data users. A common example of this is to assume an autoregressive-moving average (ARMA) model. However, most existing procedures for estimating the unknown parameters of this model assume that the input data are not subject to survey error.

In this paper we develop procedures for estimating these model parameters when the data contain survey errors. The covariance structure of the survey errors we consider include some cases where the survey errors are correlated over time.

When such a model for the behaviour of the population characteristics is assumed, the minimum mean squared error (MMSE) linear estimator can be derived. This estimator incorporates the model structure which the classical minimum variance linear unbiased estimator (MVLUE) ignores. The MVLUE is discussed in Section 2.

Blight and Scott (1973), Scott and Smith (1974), Scott, Smith and Jones (1977), R.G. Jones (1980) and others considered the implications of such stochastic models for the population means over time. These results and a more general formulation using state-space models and Kalman filters are discussed in Section 3, for the case where the stochastic model for the population characteristics is completely specified. These methods can be developed in a setting which is equivalent to a Bayes formulation, where the prior distribution is completely specified.

When the assumed model is an ARMA process in the presence of survey errors, the state-space formulation can be used to derive the maximum likelihood estimates of the unknown

[1] D.A.Binder and J.P. Dick, Social Survey Methods Division, Statistics Canada, 4th Floor, Jean Talon Building, Tunney's Pasture, Ottawa, Ontario, K1A 0T6.

parameters. We note that this approach can be viewed as empirical Bayes. We assume that the survey errors can be described through an ARMA process up to a multiplicative factor. This is discussed in Section 4.

An example of this model is described in Section 5 using data from the Canadian Travel Survey. This example shows the implications on the estimates of the model parameters when the survey errors are taken into account. We also derive a smoothed estimate of the underlying process under the model assumptions. In this example, the survey errors are independent, so that the full machinery of the general formulation in this paper is not required. However, the example demonstrates that the impact of ignoring the survey errors even in this case can be appreciable.

Section 6 contains some concluding remarks.

2. MINIMUM VARIANCE LINEAR UNBIASED ESTIMATION IN OVERLAPPING REPEATED SURVEYS

In this section we briefly review the literature for the case where the population values of a characteristic such as a mean or total are taken as fixed unknown constants. In Section 3, we study the case where a stochastic model is assumed for the population characteristic.

In overlapping surveys, where the same individual provides responses on repeated occasions, the sampling errors between occasions are usually correlated. Correlations can also occur in a multi-stage survey where some of the first stage sampling units overlap, even though the ultimate respondents differ.

Estimators which ignore these correlations and use only the data collected in the single reference period are in general inefficient relative to the minimum variance linear unbiased estimator (MVLUE). The relative efficiency depends on the size of the correlation of the sampling errors between occasions. When the correlations are zero, as in our example in Section 5, the MVLUE is simply the estimator based on data from a single reference period.

Jessen (1942) was the first to incorporate the overlapping information from the same individual on two successive occasions. Patterson (1950) provided a general theory for repeated surveys with overlapping units. He considered in detail the special case of simple random sampling from an infinite population, where the correlation for individuals is exponentially declining in time lag. On each occasions, a sample of individuals is removed from the sample of the previous occasion and a sample of individuals is added. All data are collected with reference to the current occasion only. Patterson derived the MVLUE for this setup.

Extensions have been made to the basic assumptions of Patterson (1950). Eckler (1955) called Patterson's design one-level rotation sampling. Eckler derived the MVLUE when individuals report for two successive time periods, which he termed two-level rotation sampling. He also derived the MVLUE for surveys with higher order rotation sampling designs.

Rao and Graham (1964) relaxed the infinite population assumption by incorporating the finite population correction factor into the variances of the survey error. Singh (1968) was the first to consider multi-stage designs. He examined two-stage sampling with the assumption that the correlation between responses on different occasions can be considered in two parts: (i) the correlation between second stage units (SSU's) within primary sampling units (PSU's) and (ii) the correlation between PSU means on successive occasions. If both of these correlation patterns are assumed to be that of a first order autoregressive process, then the form of the MVLUE follows the general form given by Patterson (1950).

Tikkiwal (1979) and others considered the implications of relaxing the assumption of a first order autoregressive correlation pattern. Tikkiwal concluded that if a completely general correlation structure is assumed, the simple form of the MVLUE is lost and approximations must be used in practice. Rao and Graham (1964) and Gurney and Daly (1965) proposed the use of composite estimators which are approximations to the optimal estimators. These estimators are easily implemented and have high relative efficiency. For a discussion on the use of these estimators, see Binder and Hidiroglou (1988).

Gurney and Daly (1965) also generalized the results of Patterson (1950) to a linear model framework. They introduced the concept of an "elementary estimate". This is an estimate which uses data from a specific time period, based on individuals which all join and leave the survey at the same time. The expected value of these elementary estimates can be expressed as a linear combination of the population parameters, $\{\theta_t\}$. When the correlation structure is known, standard general linear model theory can be used to derive the MVLUE.

To formalize this discussion, let y_{tj} be the j-th elementary estimate from the t-th time period, where $E(y_{tj}) = \theta_t$. If $\boldsymbol{Y}$ and $\boldsymbol{\Theta}$ are vectors with components y_{tj} and θ_t respectively, we can write:

$$\boldsymbol{Y} = \boldsymbol{X}'\boldsymbol{\Theta} + \boldsymbol{e}, \tag{2.1}$$

where $\boldsymbol{X}$ is a fixed $(n \times T)$ matrix of 0's and 1's, $E(\boldsymbol{e}) = 0$ and $E(\boldsymbol{ee}') = \boldsymbol{U}$, which is the known variance-covariance matrix of the elementary estimates. Thus, the MVLUE is given by:

$$\tilde{\boldsymbol{\Theta}} = (\boldsymbol{X}'\boldsymbol{U}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{U}^{-1}\boldsymbol{Y}, \tag{2.2a}$$

with

$$\text{Var}(\tilde{\boldsymbol{\Theta}}) = (\boldsymbol{X}'\boldsymbol{U}^{-1}\boldsymbol{X})^{-1}. \tag{2.2b}$$

These results imply that every new survey would require the updating of all previous estimates. However, since estimates from the earlier occasions often have a much smaller effect than the recent occasions, composite estimates, such as proposed by Gurney and Daly (1965), are simpler to use and have a high relative efficiency. Binder and Hidiroglou (1988) discussed the appropriateness of these methods and their application in a number of surveys. In general, they found that good results can be achieved using composite estimators, providing the rotation group biases are not substantial.

3. SIGNAL-NOISE EXTRACTION

It is quite common for economists and sociologists to treat the underlying parameters, $\{\theta_t\}$, as random inputs for their stochastic models (Smith 1978). However, if the sampling errors associated with the input data are ignored, the estimates of the parameters of the stochastic model are biased.

In this section, we show how the stochastic model assumptions can also be used to obtain model-dependent, design-consistent estimators. In Section 4, we discuss maximum likelihood estimation of these parameters. Since misspecification of the model could lead to serious biases,

hypothesis testing methods should be used to check the consistency of the model with the data. The model should also reflect the subject matter knowledge of the underlying phenomenon.

First we consider the case where the survey errors are independent. (This would be approximately true for non-overlapping surveys with small sampling fractions.) In this case, the MVLUE for θ_t is $\tilde{\theta}_t = y_t$. However, by imposing a stochastic model for the sequence of parameters, $\{\theta_t\}$, an improvement in the mean squared error of the estimate can be achieved.

Scott and Smith (1974) proposed the following model for non-overlapping surveys. They wrote the model for the survey estimates at time t as:

$$y_t = \theta_t + e_t \tag{3.1}$$

where the e_t's are independent $N(0,S_t^2)$. They assumed that the sequence of parameters, $\{\theta_t\}$, can be modelled such that, conditional on $\boldsymbol{\Theta}'_{t-1} = (\theta_1, \ldots \theta_{t-1})$,

$$\theta_t = \underline{\alpha}_t' \boldsymbol{\Theta}_{t-1} + \epsilon_t, \tag{3.2}$$

where the ϵ_t's are independent $N(0,S_t^2)$ and independent of $\{e_t\}$, and $\underline{\alpha}_t$ is a $(t-1)$ dimensional vector of constants.

In general at time $t-1$, conditional on $\boldsymbol{Y}'_{t-1} = (y_1, \ldots, y_{t-1})$, we have $\boldsymbol{\Theta}_{t-1} \sim N(\tilde{\boldsymbol{\Theta}}_{t-1}, \tilde{V}_{t-1})$. Conditional arguments then yield

$$\mathrm{E}(\theta_t|y_t) = \tilde{\theta}_t = \pi_t(\underline{\alpha}_t'\tilde{\boldsymbol{\Theta}}_{t-1}) + (1-\pi_t)y_t \tag{3.3a}$$

and

$$\mathrm{Var}(\theta_t|y_t) = (1-\pi_t)S_t^2, \tag{3.3b}$$

where

$$\pi_t = \frac{\mathrm{Var}(y_t|\theta_t)}{\mathrm{Var}(y_t)} = \frac{S_t^2}{\underline{\alpha}_t'\tilde{V}_{t-1}\,\underline{\alpha}_t + \sigma_t^2 + S_t^2}. \tag{3.3c}$$

Note that the estimator in (3.3a) is a weighted average of two components. The first consists of the best linear forecast of θ_t given the previous value of $\tilde{\boldsymbol{\Theta}}_{t-1}$; the second consists of the best estimate of θ_t from the survey. The contribution of each term is controlled by π_t, the ratio of the survey variance to the total variance. As the survey error component becomes small, then the contribution from $\tilde{\boldsymbol{\Theta}}_{t-1}$ becomes small and the estimate of θ_t in (3.3a) is composed primarily of y_t, the estimate from the survey data. Therefore, the estimator of θ_t is design-consistent whenever y_t is design-consistent.

However, as the survey error component becomes large, the estimate of θ_t is due primarily from the linear forecast of $\boldsymbol{\Theta}_{t-1}$. The relative efficiency of the estimator, $\tilde{\theta}_t$, in (3.3a) is given by $1/(1-\pi_t)$, where π_t is defined in (3.3c). The greatest efficiency gains occur when the survey error is large relative to σ_t^2, the variance of the "shocks" of the model process.

Scott and Smith (1974) and R.G. Jones (1980) also considered the case of overlapping surveys. Jones' formulation for this case was as follows. Let $\boldsymbol{\Theta}_t$ be multivariate normal with mean zero and variance matrix V_t^*. Now the observations at time t may be generalized to a vector of elementary estimates, y_t. The conditional distribution of $Y_t = (y_1', \ldots, y_t')'$ given Θ_t is assumed to be of the form:

$$Y_t = X_t' \boldsymbol{\Theta}_t + e_t, \tag{3.4}$$

where X_t is a fixed matrix of 0's and 1's linking the parameters and the observations, and e_t is the survey error, assumed to be multivariate normal with mean zero and covariance matrix U_t.

Using conditional arguments, the best estimate of θ_t given Y_t is:

$$\mathrm{E}(\boldsymbol{\Theta}_t | Y_t) = \tilde{\boldsymbol{\Theta}}_t = (X_t' U_t^{-1} X_t + \tilde{V}_t^{*-1})^{-1} X_t' U_t^{-1} Y_t \tag{3.5a}$$

with a variance of

$$\mathrm{Var}(\boldsymbol{\Theta}_t | Y_t) = (X_t' U_t^{-1} X_t + V_t^{*-1})^{-1}. \tag{3.5b}$$

This result is very general. If we allow the underlying stochastic model for $\boldsymbol{\Theta}_t$ to be very diffuse, then the inverse of V_t^* is approximately zero, thus yielding the MVLUE given by (2.2a). R.G. Jones (1980) derived (3.5) by application of stochastic least squares, so that the estimator $\tilde{\boldsymbol{\Theta}}_t$ is the minimum mean squared error (MMSE) linear estimator, even when the normality assumptions are dropped.

Applying (3.5) directly would involve inverting matrices which have the same dimensionality as the vector of all the elementary estimates for all time periods. Computing such inverses can be numerically unstable. However, expression (3.5) can often be restructured using state-space models, which are useful for describing many time series models. See Harvey (1984) for a review of such models. As we demonstrate below, this would avoid the inversion of large matrices. Some structure for $\{\theta_t\}$ and $\{e_t\}$ would be required to take advantage of the reduction in dimensionality afforded by the state-space approach. An example of such a structure, which is often used in time series applications, is an autoregressive-moving average (ARMA) process, not necessarily homogeneous in time.

For applications such as small area estimation, where the sample size is not large, modelling the variances of the survey error, U_t, using such ARMA models can be useful. This is not usually done for repeated surveys. This would also alleviate the problem of applying the result in (3.5) directly when the dimensions of V_t^* and U_t are large and the inverses are numerically unstable.

In the state-space model, two processes occur simultaneously. The first process, the observation system, details how the observations depend on the current state of the process parameters. The second process, the transition system, details how the parameters evolve over time.

State-space models can be written as follows. The observation equation is written as:

$$y_t = H_t z_t + \underline{\omega}_t, \tag{3.6a}$$

and the transition equation is written as:

$$z_t = F_{t-1}\, z_t + G_t\, \underline{\epsilon}_t, \tag{3.6b}$$

where z_t is an $(r \times 1)$ state vector, H_t is a fixed $(n_t \times r)$ matrix, F_t is a fixed $(r \times r)$ transition matrix, G_t is a fixed $(r \times m)$ matrix and $\underline{\omega}_t$ and $\underline{\epsilon}_t$ are independent random disturbances with mean zero and covariances given by $\mathrm{E}(\underline{\omega}_t\, \underline{\omega}_t') = U_t$ and $\mathrm{E}(\underline{\epsilon}_t\, \underline{\epsilon}_t') = V_t$.

As an example of this formulation, we rewrite the model studied by Blight and Scott (1973) in terms of the state-space model. Blight and Scott considered data from Patterson's (1950) one-level rotation design. They let $\bar{y}_t''$ be the mean of the new units at time t, and $\bar{y}'$ and $\bar{x}_t'$ the means of the overlapping units at times t and $t-1$, respectively. They assumed that $\bar{y}_t''$ and $\bar{y}_t' - \rho\bar{x}_t'$ are independent observations at time t, where ρ is the between-occasion correlation of the responses from the same individual. They also assumed that the mean process $\{\theta_t\}$ is first order autoregressive.

We let the state vector be $z_t' = (\theta_t, \theta_{t-1})$. The observation equation can be written as:

$$\begin{bmatrix} \hat{y}_t'' \\ \bar{y}_t' - \rho\bar{x}_t' \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & -\rho \end{bmatrix} \begin{bmatrix} \theta_t \\ \theta_{t-1} \end{bmatrix} + \begin{bmatrix} \omega_{1t} \\ \omega_{2t} \end{bmatrix},$$

where $(\omega_{1t}, \omega_{2t})'$ has a diagonal covariance matrix.

The transition equation would be written as:

$$\begin{bmatrix} \theta_t \\ \theta_{t-1} \end{bmatrix} = \begin{bmatrix} \alpha & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \theta_{t-1} \\ \theta_{t-2} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \epsilon_t,$$

where ϵ_t is $N(0,\sigma^2)$. Thus, the Blight-Scott model can be written in state-space form.

Harvey and Phillips (1979) described a method to put the ARMA (p,q) model, defined by:

$$y_t - \alpha_1 y_{t-1} - \ldots - \alpha_p y_{t-p} = \epsilon_t - \beta_1 \epsilon_{t-1} - \ldots - \beta_q \epsilon_{t-q}, \tag{3.7}$$

where the ϵ_t's are independent $N(0,\sigma^2)$, into state-space form. The dimension of z_t is $r = \mathrm{MAX}(p,q+1)$. Where necessary, $\underline{\alpha} = (\alpha_1, \ldots, \alpha_p)$ or $\underline{\beta} = (\beta_1, \ldots, \beta_q)$ is augmented with zeroes to have dimension r. The matrix, U_t is set to zero. The ARMA (p,q) model is equivalent to (3.6) when $H_t' = (1, 0, \ldots, 0)$, $G_t' = (1, -\beta_1, \ldots, -\beta_{r-1})$ and

$$F_t = \left[\begin{array}{c|c} \alpha_1 & \\ \vdots & I_{r-1} \\ \alpha_{r-1} & \\ \hline \alpha_r & O' \end{array}\right],$$

where I_{r-1} the $(r-1) \times (r-1)$ identity matrix and O' is a row vector of zeroes.

In this formulation, the state vector $z_t = (z_{1t}, \ldots, z_{rt})'$ is defined as follows:

$$z_{it} = \alpha_i y_{t-1} + \alpha_{i+1} y_{t-2} + \ldots + \alpha_r y_{t-(r-i+1)}$$

$$- \beta_{i-1}\epsilon_t - \beta_i \epsilon_{t-1} - \ldots - \beta_{r-1}\epsilon_{t-(r-i)} ,$$

for $i = 2, 3, \ldots, r$ and $z_{1t} = y_t$ as in (3.7).

A necessary condition for stationarity is that $\mathrm{Var}(z_t) = \mathrm{Var}(z_{t-1})$ for all t. From expression (3.6b), we see that this implies that

$$\mathrm{Var}(z) = \boldsymbol{F}'\mathrm{Var}(z)\boldsymbol{F} + \boldsymbol{GVG}',$$

where $V_t \equiv V$ is constant for all t. Pearlman (1980) pointed out that this can be used to obtain the initial conditions for z_1.

Often the survey error process can be included in the state-space model, when some structure for the survey errors can be assumed. We have already demonstrated this for the Blight and Scott (1973) model. Scott and Smith (1974) and Miazaki (1985) considered a variety of models which were special cases of (θ_t) being ARMA (p,q), $\{e_t\}$ being ARMA (p^*,q^*) and the scalar observations satisfying $y_t = \theta_t + e_t$. State-space models for this process can be formulated analogously to the Harvey-Phillips representation above, where the state vector z_t is the vector formed by concatenating the state vectors from each of the individual ARMA processes.

For example, suppose $\{\theta_t\}$ is an ARMA (3,0) process with parameter $(\alpha_1, \alpha_2, \alpha_3)$ and model variance σ^2 and, $\{e_t\}$ is an ARMA (0,1) process with parameter β^* and model variance s^2. An ARMA (0,1) process for $\{e_t\}$ would be plausible for a survey which follows Eckler's two-level rotation sampling pattern, where the survey estimate for θ_t is given by $\bar{y}_t$, the mean of all individuals reporting for the t-th occasion.

This can be written in state-space form by letting

$$\boldsymbol{F}_t = \left[\begin{array}{ccc|cc} \alpha_1 & 1 & 0 & 0 & 0 \\ \alpha_2 & 0 & 1 & 0 & 0 \\ \alpha_3 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right], \quad \boldsymbol{G}_t = \left[\begin{array}{cc} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ \hline 0 & 0 \\ 0 & -\beta^* \end{array}\right], \quad V_t = \left[\begin{array}{cc} \sigma^2 & 0 \\ \hline 0 & s^2 \end{array}\right], \tag{3.8}$$

$\boldsymbol{U}_t = 0$ and $\boldsymbol{H}_t' = (1\,0\,0|1\,0)$. The first three components of the state vector correspond to the state-space formulation for the $\{\theta_t\}$ process and the last two components are for the $\{e_t\}$ process.

Note that the state-space approach allows for measurement error, given by $\underline{\omega}_t$ in (3.6a). However, unless the survey design has non-overlapping units with independent sampling errors, the measurement error terms cannot be used to model the survey error. Instead, we have absorbed the measurement (survey) error into the state vector.

From the general state-space framework, the Kalman filter equations can be derived. If, as in Meinhold and Singpurwalla (1983), we let the conditional distribution of z_{t-1} given Y_{t-1} be $N(\tilde{z}_{t-1|t-1}, \tilde{P}_{t-1|t-1})$, then recursive relationships for $\tilde{z}_{t|t}$ and $\tilde{P}_{t|t}$ can be constructed. Harvey (1984) shows these relationships are equivalent to the Kalman filter.

The Kalman filter, in general, consists of two parts. The first is a one-step ahead prediction of the state vector and its covariance; the second part provides an update of the mean and covariance matrix of the state-space vector after the new observations are available.

Following the notation used in (3.6), we let $Y_1 = y_1$ and $Y'_{t+1} = (Y'_t, y'_{t+1})'$, then the one-step ahead prediction has a mean and variance given by

$$\mathrm{E}(z_1) = \tilde{z}_{1|0} \tag{3.9a}$$

$$\mathrm{Var}(z_1) = \tilde{P}_{1|0} \tag{3.9b}$$

$$\mathrm{E}(z_t|Y_{t-1}) = \tilde{z}_{t|t-1} = F_t \tilde{z}_{t-1|t-1} \tag{3.9c}$$

$$\mathrm{Var}(z_t|Y_{t-1}) = \tilde{P}_{t|t-1} = F_t \tilde{P}_{t-1|t-1} F'_t + G_t V_t G'_t . \tag{3.9d}$$

The update of the mean and variance for the state vector at time t after the observation at time t becomes available is:

$$\begin{aligned} \mathrm{E}(z_t|Y_t) &= \tilde{z}_{t|t} \\ &= \tilde{z}_{t|t-1} + \tilde{P}_{t|t-1} H_t (H'_t \tilde{P}_{t|t-1} H_t + U_t)^{-1} (y_t - H'_t z_{t|t-1}) \end{aligned} \tag{3.10a}$$

$$\mathrm{Var}(z_t|Y_t) = \tilde{P}_{t|t} = \tilde{P}_{t|t-1} - \tilde{P}_{t|t-1} H_t (H'_t \tilde{P}_{t|t-1} H_t + U_t)^{-1} H'_t \tilde{P}_{t|t-1} \tag{3.10b}$$

The equations (3.9) and (3.10) are the well-known Kalman filter equations. The formulation followed here is essentially Bayesian; however, it is possible to derive equivalent results using orthogonal projections; see Young (1984).

The simplification in the computations due to the Kalman filter formulation in the sample survey setting can be seen by comparing equations (3.9) and (3.10) with R.G. Jones' (1980) result (3.5). Note that Jones' result required the inversion of a matrix with dimensionality given by the complete vector of survey estimates.

The Kalman filter can also be used to obtain smoothed estimates given by $\mathrm{E}(z_t|Y_t)$ for $T > t$. Details of this backcasting may be found in Harvey (1984).

Remarks

1. Although the Kalman filter assumes an infinite population model, when the sample survey is based on a large sample, the central limit theorem often allows the survey errors to be approximately normally distributed. As well, since the smoothed estimators for $\{\theta_t\}$ are the same as those obtained by R.G. Jones (1980) in (3.5a), these are the linear MMSE estimators even when the normality assumptions are dropped.

2. Missing time points can be incorporated in the state-space approach. If y_t is missing at time t, then the updating equations analagous to (3.9) become $\tilde{z}_{t|t} = \tilde{z}_{t|t-1}$ and $\tilde{P}_{t|t} = \tilde{P}_{t|t-1}$ as in R.H. Jones (1980). However, smoothed estimates for the missing time points will depend strongly on the model selected, since no survey estimate is available. Therefore, the risks of model misspecification here are high.

3. The likelihood function, which we discuss in Section 4 for obtaining the maximum likelihood estimates of the unknown parameters, can also be obtained when some data are missing, using the same approach given by R.H. Jones (1980). However, missing data will tend to increase the standard errors of the parameter estimates. In our example of Section 5, we encounter a case with missing time points.

4. ESTIMATION OF THE PARAMETERS IN A STATE-SPACE MODEL

When data are generated from the ARMA model (3.7) and the parameters $\underline{\alpha}$, $\underline{\beta}$, and σ^2 are unknown, the maximium likelihood estimates for the unknown parameters can be obtained using the likelihood function derived from the state-space model. This approach was suggested by Harvey and Phillips (1979), R.H. Jones (1980) and others.

The usual state-space models can also be used when the input data have independent measurement errors. This is the case for our example of Section 5, where we show the effect on the parameter estimates when the survey errors are taken into account.

Maximum likelihood estimation of these parameters when the data have correlated survey errors has not previously been studied in detail. For a model with univariate stationary observations $\{y_t\}$, Scott, Smith and Jones (1977) suggested using the estimated autocovariance function of the observations $\{y_t\}$ to estimate the parameters of the ARMA process. Here, the data model is $y_t = \theta_t + e_t$. The variances and covariances of the survey errors, $\{e_t\}$ can be estimated using design-based methods; see, for example, Wolter (1985).

Efficient estimation of the autocovariances of the survey errors, assuming stationarity of the series, is an area which has not received attention in the literature, so ad hoc methods would be used in practice. Future research in modelling these survey errors would be worthwhile. In our example in Section 5, we could assume independent survey errors, so this was not problematic.

Assuming the autocovariance of $\{e_t\}$ is available, the autocovariance of $\{\theta_t\}$ can be estimated by $\text{Cov}(\theta_t, \theta_{t-s}) = \text{Cov}(y_t, y_{t-s}) - \text{Cov}(e_t, e_{t-s})$. However, this method is not fully efficient (Smith; 1978). Moreover, this method would not incorporate non-stationary survey errors.

Miazaki(1985) considered the case where $\{\theta_t\}$ is an ARMA $(p,0)$ process. She also assumed $\{e_t\}$ to be an ARMA $(0,q)$ process which could be estimated directly from the survey. Miazaki then wrote the observations $\{y_t\}$ as an ARMA $(p,p+q)$ process which she estimated by restricted maximum likelihood methods.

Representing non-stationarity of survey errors in the state-space representation can sometimes be handled through nonhomogeneous matrices for V_t, the variance matrix of the random "shocks" from the transition equation (3.6b). For example, in (3.7) s^2 would be replaced by s_t^2 to allow for non-homogeneous survey errors. This approach is taken in the example in Section 5.

In general, for state-space models given by (3.5), Harvey and Phillips (1979) write the exact likelihood function as follows. Letting

$$\hat{y}_{t|t-1} = \mathrm{E}(y_t|Y_{t-1}) = H_t'\tilde{z}_{t|t-1}$$

and

$$R_t = \mathrm{Var}(y_t|Y_{t-1}) = H_t'\tilde{P}_{t|t-1}H_t + U_t,$$

the log-likelihood function for $Y_T' = (y_1', \ldots, y_T')$ is

$$\log f(Y_t) = (1/2)\sum_{t=1}^{T}\log|R_t| - (1/2)\sum_{t=1}^{T}(y_t - \hat{y}_{t|t-1})'R_t^{-1}(y_t - \hat{y}_{t|t-1}). \quad (4.1)$$

The unknown parameters in (4.1) are contained in $\hat{y}_{t|t-1}$ and in R_t. Depending on the algorithm used to maximize (4.1) with respect to the unknown parameters, it may be necessary to compute first and second derivatives of (4.1) with respect to the unknown parameters. This generally involves finding derivatives of $\tilde{z}_{t|t-1}$ and $\tilde{P}_{t|t-1}$. These can be computed numerically using the recursions given in (3.8) and (3.9). For example, (3.8c) yields $\partial\tilde{z}_{t|t-1} = (\partial F_t)\tilde{z}_{t-1|t-1} + F_t(\partial\tilde{z}_{t-1|t-1})$. The other expressions using (3.8) and (3.9) can be determined similarly.

The inclusion of regression parameters into (4.1) can be accomplished by replacing y_t by the deviation of y_t from the regression line. Tam (1987) generalized this concept even further by considering a model where the underlying stochastic process is determined by a state-space model for the regression coefficients which evolve over time.

To maximize the likelihood function (4.1) with respect to unknown parameters, an iterative procedure is needed. We omit details of the procedure used for the application in Section 5 since efficient procedures are still in the development stage.

Once having estimated the parameters, smoothed values for the state vector, $\tilde{z}_{t|T} = \mathrm{E}(z_t|Y_T)$ after time $T > t$, can be obtained using the backcasting formulae given by the Kalman filter; see Harvey (1984). Thus, for example, if $y_t = \theta + e_t$ as in (3.1), after backcasting we may formulate $y_t = \tilde{\theta}_{t|T} + \tilde{e}_{t|T}$, so that $\tilde{\theta}_{t|T}$ becomes the smoothed estimate of the mean at time t after observing Y_t.

To derive the standard error of the smoothed estimate it is necessary to account for the fact that the unknown parameters have been estimated from the data, particularly when the data series is short; see Jones (1979). Hamilton (1986) suggests doing this by Monte Carlo simulations. He generates a set of multivariate normal random variables with mean given by the maximum likelihood estimates for the parameters and variance given by the inverse of the estimated Fisher information matrix. He then estimates $\mathrm{E}(\tilde{P}_{t|T})$ and $\mathrm{Var}(\tilde{z}_{t|T})$, where the expectation and variance are taken over the generated parameter values. The sum of these two components is the estimated covariance matrix of the estimated state vector. This method assumes that the sample size is large, so that the normal approximation to the sampling distribution of the parameter estimates is valid.

In the examples of Section 5, we approximate the standard deviation of the sampling errors of the smoothed estimates, ignoring the variation due to estimating certain model parameters. We then compare these with the actual root mean squared errors of the sampling distribution obtained from simulated data.

5. DATA ANALYSIS

In this section we show the impact of the survey errors on estimates of the parameters of a first order autoregressive model with regression terms. In our example the survey errors are assumed to be independent between occasions. More complicated cases with correlated survey error and higher order ARMA models for the population characteristic could be handled within the framework we have described. We chose this example to demonstrate that the impact of accounting for the survey errors can be appreciable even for this relatively simple model.

We used data from Saskatchewan respondents to the Canadian Travel Survey (CTS). The CTS is conducted by Statistics Canada to collect descriptive statistics on the travelling habits and characteristics of Canadian residents. This survey is conducted as an "add-on" to the Labour Force Survey (LFS). The LFS is a monthly rotating panel survey with six rotation groups. However, the CTS is conducted at most four times a year, with at least one, but possibly as many as three rotation groups. The rotation groups used by the CTS for the quarters when the CTS is conducted are chosen so that there are no overlapping panels between occasions.

The survey errors are assumed to be independent. This is only approximately true. The LFS is a multi-stage survey and the primary sampling units (PSU's) do not rotate out as quickly as the individual rotating panels. The same PSU's are used on a number of occasions. Therefore, although the CTS sample is selected such that the panels do not overlap between occasions, the independence assumption is approximately true only when the correlation of the sampling errors between quarterly periods within the same PSU is small. This assumption was not verified.

The coefficients of variation (as a percentage) were calculated using the function:

$$\text{CV} = \alpha y^{-\beta}/\sqrt{\text{number of rotation groups}},$$

where y is the survey estimate in thousands. This is the function recommended to users of the CTS for data on Saskatchewan residents; see Statistics Canada (1985). In this report, the parameters α and β were estimated at 91.7528 and 0.353253, respectively, using a loglinear regression model applied to 1979 data. For the purposes of our example, these coefficients of variation were rounded to the nearest tenth of a percent.

The assumed model was:

$$y_t = \theta_t + e_t, \tag{5.1}$$

where the e_t's are independent survey errors, with $e_t \sim N(0,s_t^2)$ and

$$\theta_t = \gamma_0 + \gamma_1 t + \gamma_2 Q_{1t} + \gamma_3 Q_{2t} + \gamma_4 Q_{3t} + \epsilon_t, \tag{5.2}$$

where $\{\epsilon_t\}$ is ARMA (1,0) with parameters (α,σ^2). The regression terms in (5.2) are, respectively, the intercept, a term representing the quarter number with t taking values from -15.5 to 15.5 linearly in time and, finally, seasonal terms for the first three quarters of each year, where

$$\begin{aligned} Q_{it} &= 1 \text{ if the } t\text{-th observation is in the } i\text{-th quarter;} \\ &= -1 \text{ if the } t\text{-th observation is in the fourth quarter;} \\ &= 0 \text{ otherwise;} \end{aligned}$$

for $i = 1, 2, 3$.

Better models may be available for these data, although with such a small data set, tests of hypotheses against alternative models would not be very powerful.

To obtain the maximum likelihood estimates for the unknown parameters of this model, it is necessary to incorporate the assumptions made about the survey errors in the estimation procedure. Most users of official statistics ignore this survey error and implicitly assume that the input data are error-free. This does not seriously affect the results when the variance of the survey error is small relative to the variance of the model error.

The survey estimates and the coefficients of variation of the survey errors relative to these estimates are given in Tables 1 and 2. The results of the maximum likelihood estimation procedure are displayed in Tables 3 and 4. Two estimates are given for each model. The column labeled "Estimate: With Sampling Error" uses the method incorporating the assumed error structure; whereas the column labeled "Estimate: Ignoring Sampling Error" repeats the estimation under the assumption that the survey estimate is observed without error. In both cases model (5.2) is assumed.

Table 1

Overnight Person-Trips of Saskatchewan Residents to Destinations within Saskatchewan[1]

Year	Quarter	No. of Rotation Groups	Survey Estimate (000's)	Smoothed Estimate (000's)	Survey C.V. (%)	Smoothed C.V. (%)	Simulated RMSE (%)	Simulated Bias (%)
1979	Winter	1	598	611	9.6	5.9	6.9	0.1
	Spring	1	808	813	8.6	4.8	4.9	0.4
	Summer	3	1033	1103	4.6	3.0	3.1	0.0
	Fall	3	678	683	5.3	4.3	4.5	1.2
1980	Winter	1	578	608	9.7	5.5	5.8	0.1
	Spring	3	837	837	4.9	3.7	3.6	0.0
	Summer	1	1451	1169	7.0	3.3	3.5	0.3
	Fall	1	744	724	8.9	5.1	5.9	0.8
1981	Winter	3	631	632	5.4	4.3	5.0	-0.1
	Summer	3	1262	1172	4.2	2.9	3.3	0.1
1982	Winter	1	565	613	9.8	5.5	6.4	-0.4
	Spring	1	901	838	8.3	4.5	5.1	0.8
	Summer	3	1167	1147	4.4	2.9	3.1	0.1
	Fall	1	721	706	9.0	5.1	5.6	0.2
1984	Winter	1	585	598	9.6	5.8	6.7	-1.2
	Spring	1	788	804	8.7	4.6	5.2	-0.4
	Summer	3	1068	1107	4.5	2.9	3.6	-0.5
	Fall	1	711	686	9.0	5.3	6.7	0.7
1986	Winter	1	793	630	8.7	6.2	7.1	-1.3
	Spring	3	798	808	5.0	3.9	3.9	-0.4
	Summer	3	1053	1096	4.5	3.0	3.3	-0.3
	Fall	3	650	663	5.4	4.4	4.2	0.2

[1] The Canadian Travel Survey was not conducted in the Spring and Fall Quarters of 1981 and during 1983 and 1985.

Simulations in last two columns are based on a sample size of 100.

Table 2

Overnight Person-Trips of Saskatchewan Residents to Destinations in Manitoba[1]

Year	Quarter	No. of Rotation Groups	Survey Estimate (000's)	Smoothed Estimate (000's)	Survey C.V. (%)	Smoothed C.V. (%)	Simulated RMSE (%)	Simulated Bias (%)
1979	Winter	1	27	34	28.6	13.4	14.1	0.5
	Spring	1	33	48	26.7	11.0	10.2	0.9
	Summer	3	78	80	11.4	6.6	7.1	1.3
	Fall	3	55	48	12.9	10.1	10.8	0.6
1980	Winter	1	24	30	29.7	13.6	14.5	0.5
	Spring	3	63	50	12.3	9.5	9.4	0.7
	Summer	1	86	80	19.0	6.6	6.3	0.8
	Fall	1	75	46	19.9	11.0	12.2	0.5
1981	Winter	3	42	34	14.2	11.3	13.2	1.0
	Summer	3	79	82	11.3	5.9	5.7	0.1
1982	Winter	1	33	34	26.5	12.5	13.2	-2.8
	Spring	1	46	44	23.7	10.7	10.0	1.6
	Summer	3	78	82	11.4	5.7	5.4	0.1
	Fall	1	30	42	27.6	10.9	11.4	0.3
1984	Winter	1	36	34	25.7	13.8	16.8	-1.3
	Spring	1	48	43	23.4	11.4	11.5	0.1
	Summer	3	82	82	11.1	6.1	7.3	-0.2
	Fall	1	30	40	27.7	11.5	11.4	0.6
1986	Winter	1	33	33	26.7	16.3	19.9	-0.8
	Spring	3	38	41	14.6	10.9	11.7	-0.1
	Summer	3	90	81	10.8	7.1	8.8	-0.3
	Fall	3	42	40	14.1	11.2	10.5	1.7

[1] The Canadian Travel Survey was not conducted in the Spring and Fall Quarters of 1981 and during 1983 and 1985. Simulations in last two columns are based on a sample size of 100.

Table 3

Parameter Estimates for Saskatchewan to Saskatchewan Person-Trips[1]

Parameter	Ignoring Sampling Error	With Sampling Error				
	Estimate	Estimate	Standard Error	Simulated RMSE	Simulated Bias	t-value of Bias
REGRESSION						
Intercept (γ_0)	831.4	815.0	15.6	14.4	1.8	1.29
Linear (γ_1)	-0.84	-0.86	1.52	1.51	-0.10	-0.65
1st Quarter (γ_2)	-209.6	-203.8	21.8	24.6	-3.5	-1.41
2nd Quarter (γ_3)	-4.0	7.1	22.9	23.8	0.4	0.17
3rd Quarter (γ_4)	340.1	316.0	21.2	23.4	-0.4	-0.18
ARMA						
Autoregressive (α)	0.14	0.47	0.66	0.68	-0.39	-6.77
Model Variance (σ^2)	7930.5	879.3	1205.6	770.0	-488.2	-8.16

[1] Simulations and t-values are based on a sample size of 100.

Table 4

Parameter Estimates for Saskatchewan to Manitoba Person-Trips[1]

Parameter	Ignoring Sampling Error	With Sampling Error				
	Estimate	Estimate	Standard Error	Simulated RMSE	Simulated Bias	t-value of Bias
REGRESSION						
Intercept (γ_0)	51.2	50.5	1.9	2.0	0.4	1.57
Linear (γ_1)	-0.17	-0.13	0.18	0.17	-0.04	-2.01
1st Quarter (γ_2)	-20.1	-17.2	3.4	3.5	-0.6	-1.52
2nd Quarter (γ_3)	-5.9	-6.1	3.6	3.7	-0.1	-0.32
3rd Quarter (γ_4)	30.7	30.8	3.7	3.7	0.0	-0.07
ARMA						
Autoregressive (α)	0.14	-0.75	0.66	0.71	0.49	7.90
Model Variance (σ^2)	100.0	5.7	18.7	9.5	-0.3	-0.29

[1] Simulations and t-values are based on a sample size of 100.

The estimates of the regression parameters are essentially the same under either assumption. However, the autoregressive component estimates differ considerably under the two assumptions. In particular, the model variance increases substantially. This variance estimate increases because the variation due to survey error is missing from the model. The reason that the estimates of the regression coefficients are not affected is that the estimators for these coefficients remain unbiased, although they are somewhat inefficient.

Once the parameters of the model have been estimated, it is possible to use the assumed model to adjust the individual estimates of the number of overnight person-trips. The results discussed below demonstrate how the procedure reduces the coefficients of variation for these smoothed estimates when the model assumptions are correct. Such a procedure is analogous to model-dependent small area estimation methods.

The smoothed estimates and their coefficients of variation are given in Tables 1 and 2. These coefficients of variation are calculated, taking into account the sampling error of the regression coefficients, $\gamma_0, \ldots, \gamma_4$. This is possible since, given α and σ^2, the smoothed estimates are linear functions of the original survey estimates, so that the variances can be computed from this linear function and the assumed model variance of the regression residuals. However, the sampling errors for the estimated α and σ^2 were ignored at this point. The effect of ignoring these sampling errors is discussed below.

The smoothed estimates for travel within Saskatchewan are generally close to the original survey estimates, with possible exceptions for the Summer of 1980 and the Winter of 1986. Those for travel to Manitoba are also close, with a possible exception being the Fall of 1980. These exceptional cases could possibly be outliers or could be due to a special event that boosted tourism in those quarters. In general, such phenomena could be incorporated into the model by: (i) increasing the model variance in the state-space model for those periods or adding appropriate dummy variables for special events or (ii) increasing the sampling variance for outliers. A more in-depth knowledge of the circumstances would be required to decide whether such adjustments are appropriate. The analysis here can help pinpoint possible unusual cases.

Because the analysis so far has ignored the effect of the sampling error associated with estimating α and σ^2, we performed a simulation study to assess its seriousness. Jones (1979), Hamilton (1986) and Tam (1987) have suggested that these sampling errors should not be ignored, especially when the time series has few observations.

For the simulation, we generated sets of random data following the assumed model given by (5.1) and (5.2). We took as our parameter values the maximum likelihood estimates of the model. The same missing data pattern was used in the simulations as in the original data set. One hundred such data sets were generated for each model. In Tables 1 and 2, we report the percentage bias of the smoothed values and the percentage root mean squared error for the difference between the smoothed values and the true values based on these simulations.

To assess whether 100 was a sufficiently large number of simulations to estimate the root mean squared error (RMSE), we computed an estimate of the coefficient variation of the estimator of the RMSE. From the simulations we obtained an unbiased estimate of the variance of the estimator of the mean squared error. We then used Taylor linearization to estimate the variance of the estimator of the RMSE. The estimated coefficients of variation ranged from 6% to 11% for destinations within Saskatchewan and from 5% to 9% for destinations in Manitoba. Therefore, these estimates of the RMSE's do provide a reasonable assessment of the effect of ignoring the sampling error of the autoregressive parameters.

In Tables 1 and 2, the biases of the adjustment procedure are all small and, in fact, for the two sets of 22 observations only four were significant at the 5% level using a standard t-test.

We also note that the percentage root mean squared errors based on the imulations tend to be larger than those under the column entitled "Smoothed C.V.". This is to be expected since the simulations include sampling errors arising from the estimation of α and θ^2. However, the values of the "Smoothed C.V.'s" do give reasonable approximations to the simulated values, so the procedure which ignores the effect of the sampling error of α and θ^2 does not seriously affect the coefficients of variation.

In Table 3 and 4, we report some simulation results for the estimated parameters. For the regression coefficients, only one of the biases was significant at the 5% level. The standard errors are all consistent with the simulation results.

On the other hand, the simulations did point out a problem with the estimates for α and σ^2. The biases for the estimates of α were highly significant. As can be seen from Tables 3 and 4, one of the biases of σ^2 was also highly significant. The simulated root mean squared errors were not very close to the asymptotic approximation of the standard error obtained by inverting the Fisher information matrix. It seems that the sample size for our problem is not sufficiently large for the asymptotic approximations to be very accurate. This is a common problem for time series analyses of short series.

6. CONCLUSION

In cases where the variances of the survey errors are small relative to the variances of the model errors, the smoothed estimates would be close to the minimum variance linear unbiased estimates and there would be no appreciable reduction in the standard errors of the estimates, even when the assumed model is true. However, for cases such as small domain estimation where the sampling errors are not small, the standard errors for the smoothed estimates may be substantially smaller than those for the original survey estimates. For example, the smoothed estimates for the Saskatchewan-to-Manitoba data showed a greater improvement than the Saskatchewan-to-Saskatchewan data, since the sampling errors for the survey data were larger for the former data set.

One of the implications of assuming models for repeated surveys is that if the models are misspecified, the MMSE estimators may be seriously biased. It is important, therefore, to choose a model which is both consistent with the data and which reflects subject matter knowledge about the underlying phenomena. In our example the data set is small, so that a large number of statistical models would be consistent with the data.

Our simulation studies suggest that even for small data sets, the asymptotic approximations to the variances of the smoothed estimates are quite reasonable. However, as in the case of more traditional applications of time series analyses, the asymptotic approximations for the sampling errors of the parameter estimates may be poor.

ACKNOWLEDGEMENTS

The authors would like to thank an Associate Editor of this journal and the referees for helpful comments on earlier versions. In particular, we are grateful to one referee whose thorough and insightful comments have led to many improvements of the paper. We would also like to thank Pierre Hubert, Chief of the Tourism, Travel and Recreation Section, Education Culture and Tourism Division for making available the data from the Canadian Travel Survey. Some of the material presented here appeared in the second author's Master thesis at the University of Guelph.

REFERENCES

BINDER, D.A., and HIDIROGLOU, M.A. (1988). Sampling in Time. In *Handbook of Statistics, Vol. 6*, (Eds, P.R. Krishnaiah and C.R. Rao), Amsterdam: Elsevier Science, 187-211.

BLIGHT, B.J.N., and SCOTT, A.J. (1973). A stochastic model for repeated surveys. *Journal of the Royal Statistical Society*, Ser. B, 35, 61-68.

ECKLER, A.R. (1955). Rotation sampling. *Annals of Mathematical Statistics*. 26, 664-685.

GURNEY, M., and DALY, J.F. (1965). A multivariate approach to estimation in periodic sample surveys. *Proceedings of the Social Statistics Section, American Statistical Association*, 247-257.

HAMILTON, J.D. (1986). A standard error for the estimated state vector of a state-space model. *Journal of Econometrics*, 33, 387-397.

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

HARVEY, A.C., and PHILLIPS, G.D.A. (1979). Maximum likelihood estimation of regression models with autoregressive-moving average disturbances. *Biometrika*, 66, 49-58.

JESSEN, R.J. (1942). Statistical investigation of a farm survey for obtaining farm facts. *Iowa Agricultural Experimental Station Research Bulletin*, 304, 54-59.

JONES, R.G. (1979). The efficiency of time series estimators for repeated surveys. *Australian Journal of Statistics*, 21, 45-56.

JONES, R.G. (1980). Best linear unbiased estimators for repeated surveys. *Journal of the Royal Statistical Society*, Ser. B, 42, 221-226.

JONES, R.H. (1980). Maximum likelihood fitting of ARMA models to time series with missing observations. *Technometrics*, 22, 389-395.

KALMAN, R.E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82, 35-45.

MEINHOLD, R.J., and SINGPURWALLA, N.D. (1983). Understanding the Kalman Filter. *The American Statistician*, 37, 123-127.

MIAZAKI, E.S. (1985). Estimation for time series subject to the error of rotation sampling. Ph.D. Thesis, Iowa State University, Ames, Iowa.

PATTERSON, H.D. (1950). Sampling on successive occasions with partial replacement of units. *Journal of the Royal Statistical Society*, Ser. B, 12, 241-255.

PEARLMAN, J.G. (1980). An algorithm for the exact likelihood of a high-order autoregressive-moving average process. *Biometrika*, 67, 232-233.

RAO, J.N.K., and GRAHAM, J.E. (1964). Rotation designs for sampling on repeated occasions. *Journal of the American Statistical Association*, 59, 492-509.

SCOTT, A.J., and SMITH, T.M.F. (1974). Analysis of repeated surveys using time series methods. *Journal of the American Statistical Association*, 69, 674-678.

SCOTT, A.J., SMITH, T.M.F., and JONES, R.G. (1977). The application of time series methods to the analysis of repeated surveys. *International Statistics Review*, 45, 13-28.

SINGH, D. (1968). Estimates in successive sampling using multi-stage design. *Journal of the American Statistical Association*, 63, 99-112.

SMITH, T.M.F. (1978). Principles and problems in the analysis of repeated surveys. In *Survey Sampling and Measurement*, (Ed. N.K. Namboodini), New York: Academic Press, 201-216.

STATISTICS CANADA (1985). Canadian Travel Survey: Estimation and variance estimation procedures. Technical Report, Statistics Canada.

TAM, S.M. (1987). Analysis of repeated surveys using a dynamic linear model. *International Statistical Review*, 55, 63-73.

TIKKIWAL, B.D. (1979). Successive sampling — a review. *Bulletin of the International Statistics Institute*, 48, 367-384.

WOLTER, K.M. (1985). *Introduction to Variance Estimation*. New York: Springer-Verlag.

YOUNG, P. (1984). *Recursive Estimation and Time Series Analysis: An Introduction*. New York: Springer-Verlag.