Research Report
Department of Statistics
Göteborg University
Sweden

Some statistical aspects on methods for detection of turning points in business cycles

Eva Andersson
David Bock
Marianne Frisén

Research Report 2002:7
ISSN 0349-8034

Mailing address:	Fax	Phone	Home Page:
Dept of Statistics	Nat: 031-773 12 74	Nat: 031-773 10 00	http://www.stat.gu.se/stat
P.O. Box 660	Int: +46 31 773 12 74	Int: +46 31 773 10 00	
SE 405 30 Göteborg			
Sweden			

Some Statistical Aspects on Methods for Detection of Turning Points in Business Cycles

Eva Andersson, postdoc research associate.
Department of Statistics, P.O.Box 660, SE-40530, Sweden, phone +46 317731264, fax +46 317731274, Eva.Andersson@Statistics.GU.se

David Bock, graduate student.
Department of Statistics, P.O.Box 660, SE-40530, Sweden, phone +46 317731262, fax +46 317731274, David.Bock@Statistics.GU.se

Marianne Frisén, professor.
Department of Statistics, P.O.Box 660, SE-40530, Sweden, phone +46 317731255, fax +46 317731274, Marianne.Frisen@Statistics.GU.se

Abstract

Statistical and practical aspects on methods for on-line turning point detection in business cycles are discussed. When a method is used on a real data set, there are a number of special data problems to be considered. Among these are: the effect of smoothing, seasonal variation, autoregression, the presence of a trend and problems with multivariate data. Different approaches to these data problems are reviewed and discussed. In a practical situation, another important aspect is the estimation procedure for the parameters of the monitoring system. Three likelihood based methods for turning point detection are compared, one based on a Hidden Markov Model and another including a non-parametric estimation procedure. The three methods are used to analyze an actual data set of a period of the Swedish industrial production. The relative merits of comparing methods by one real data set or by simulations are discussed.

Keywords: Business cycles; decision rules; sequential signals; turning points; non-parametric; smoothing; seasonality; autoregressive; optimal; likelihood ratio; Markov switching; regime switching model; Swedish industrial production

Biographies:

Eva ANDERSSON is postdoc research associate in Statistics at Göteborg University. Her research interests include business cycles and statistical surveillance.

David BOCK is graduate student in Statistics at Göteborg University. His research interests include financial statistics and statistical surveillance.

Marianne FRISÉN is professor of Statistics at Göteborg University. Her research interests include statistical surveillance and nonparametric methods.

1. Introduction

In order to predict the turns of the business cycles leading economic indicators can be used. By timely detection of the turns in one or several leading indicators the turning point time of the general business cycle can be predicted. Timely prediction of a change in phase of the business cycle is important for both government and industry.

In this paper predictions of the actual value of the business cycle are not discussed. Instead the methods discussed here concern detecting the turning point time, i.e. the time of a change from a recession phase to an expansion phase (or vice versa). The tardiness of stabilization policies induces unintended effects (see Öller (1987)). For reviews and general discussions see e.g. Neftci (1982), Zarnowitz and Moore (1982), Westlund and Zackrisson (1986), Hackl and Westlund (1989), Zellner, Hong and Min (1991), Jun and Joo (1993), Lahiri and Wang (1994), Li and Dorfman (1996), Birchenhall, Jessen, Osborn and Simpson (1999) and Layton and Katsuura (2001).

As pointed out by e.g. Diebold and Rudebusch (1996), Kim and Nelson (1998) and Birchenhall et al. (1999) two distinct but related approaches to the characterization and dating of the business cycle can be found. One approach emphasizes the common movements of several variables. This approach is pursued by e.g. Stock and Watson (1991) and Stock and Watson (1993) and is briefly discussed in Section 3.2.5 on multivariate approaches. The other approach, the regime shift, is the one pursued in this paper, as also in the works by Neftci (1982), Diebold and Rudebusch (1989), Hamilton (1989), Jun and Joo (1993), Lahiri and Wang (1994), Layton (1996), Koskinen and Öller (1998), Layton (1998) and Birchenhall et al. (1999).

In recent years methods based on likelihood or posterior distributions have been in focus. In the general theory on statistical surveillance there are proofs for their optimality properties (see e.g. Shiryaev (1963) and Frisén and de Maré (1991)). Many of the suggested methods are based on a hidden Markov model (HMM), also referred to as a Markov-switching or regime switching model, see Hamilton (1989). In Andersson, Bock and Frisén (2002) three different likelihood based methods are compared in detail with respect to the delay of motivated alarms and the predictive value of an alarm at different times. In this report, the same methods are compared, but here the emphasis is on the situation where the methods are used on a real data set of quarterly data and the problems connected with this.

In surveillance the inference situation is one of repeated decisions. Continual observation is made of a time series with the goal of detecting the turning point in the underlying process as soon as possible. The surveillance system consists of an alarm statistic and an alarm limit. For general reviews on statistical surveillance, see Shiryaev (1963), Frisén and de Maré (1991), Wetherhill and Brown (1991), Srivastava and Wu (1993), Lai (1995), Frisén and Wessman (1999) and Frisén (1999).

The purpose of this paper is to discuss the particular data problems that are present in methods for on-line turning point detection in cyclical, economic processes. Some of the suggested approaches for dealing with seasonal variation, autocorrelated data, trend and multivariate problems are reviewed. Furthermore, the effects of smoothing the data are investigated by a Monte Carlo study.

Another important aspect in an applied situation is the assumptions made about the process and hence the estimation procedure used. We discuss different assumptions and the estimation procedures connected with them.

A frequently used approach for evaluating the performance of a method for turning point detection is to use the out-of-sample performance. Sometimes evaluation is made by Monte Carlo methods, as in Andersson et al. (2002). In this paper the evaluation is made on a set of real data, the Swedish Industrial Production (IP).

The paper is organized as follows. Section 2 contains a description of different likelihood based approaches. In Section 3 special data problems and estimation procedures are discussed. Also a Monte Carlo study regarding the effects of smoothing is presented in this section. The statistical properties of the three methods are summarized in Section 4 where also a period of the Swedish IP is analyzed. The pros and cons of this way to evaluate methods are discussed. Section 5 contains a summarizing discussion.

2. Concepts of likelihood based surveillance for detection of turning points

In this section the basic concepts of likelihood based surveillance for online detection of turning points is given.

2.1 Model within each expansion- and recession phase

The situation under study is one where X is a leading economic indicator. By monitoring X we want to detect a regime shift (a turning point) as soon as possible. Here the model for X at time t is:

$$X(t) = \mu(t) + \varepsilon(t), \tag{1}$$

where $\varepsilon(t)$ ~iid N[0; σ^2] and $\mu(t)$ is a cyclical stochastic process described below.

The aim is to detect a change in μ, from expansion state to recession state (or vice versa). The definition of a turning point (or regime change) is that the regression of μ on time is monotonic within each regime. That is, for a peak we have

$$\mathrm{E}[X_t]=\mu_t : \begin{cases} \mu(1) \leq \mu(2) \leq \ldots \leq \mu(t), & t < \tau \\ \mu(1) \leq \ldots \leq \mu(\tau-1) \text{ and } \mu(\tau-1) \geq \ldots \geq \mu(t), & t \geq \tau \end{cases} \tag{2}$$

where t=1 is in a period of expansion, τ is the random time of a turning point (the time of change from the expansion to a recession) and X_t={$X(1)$, $X(2)$, ..., X(t)}. The process μ is stochastic since τ is stochastic. The monotonicity restrictions for a trough are the opposite of those in (2).

Finding a model for μ can be difficult since the cycles differ over time, both in length and level. One recently evaluated approach (Andersson (2002)) is to use only the monotonicity restrictions in (2) and thus avoiding to use any parametric assumptions regarding the shape of the turning point.

Parametric assumptions regarding μ make the method more powerful if the assumptions are valid. One example of a parametric assumption is that the regression on time is linear within each phase, with known parameters. For the situation when the turn is a peak, we have

$$\mathrm{E}[X(t)]=\mu(t) = \begin{cases} \beta_0 + \beta_1 \cdot t, & t < \tau \\ \beta_0 + \beta_1 \cdot (\tau - 1) - \beta_2 \cdot (t - \tau + 1), & t \geq \tau \end{cases}, \tag{3}$$

where t ={1, 2, ...}. The expected value in (3) holds for a random walk with drift where the value of the drift parameter changes from β_1 to -β_2 at time t=τ.

In many HMM approaches the series under observation, X, is differentiated and the expected value of the differentiated process is assumed to be constant, conditional on the state, see e.g. Layton (1996), Ivanova, Lahiri and Seitz (2000) and Layton and Katsuura (2001). If the process is assumed to be a random walk with drift, or another process with an expected value as in (3), then the differentiated series will have a constant expected value, conditional on the state. For the situation when the turn is a peak we have

$$\mathrm{E}[X(t)-X(t-1)] = \begin{cases} \beta_1, & t<\tau \\ \beta_2, & t\geq\tau \end{cases}, \tag{4}$$

where $\beta_1 \geq 0$ and $\beta_2 < 0$. When the observations are independent over time, as in (1), the expected values in (4) imply the linear functions in (3). If the process is a random walk with drift, or another process with an expected value as in (3), then (4) is valid. The business cycle can be assumed to be symmetric or asymmetric in different aspects, see e.g. Neftci (1984), Falk (1986) and McQueen and Thorley (1993). One kind of asymmetry is differences in slopes between phases of recession and expansion (i.e. $|\beta_1| \neq |\beta_2|$). The effect of a non-symmetric turning point for a monitoring method where only monotonicity restrictions are used is studied in Andersson (2001).

The standard deviation σ can be assumed to be different for recession and expansion, as in e.g. Koskinen and Öller (1998). However, this might not always be the case, for example in Hussey (1992) it is demonstrated that the standard deviations are different for one indicator but not for another. Macroeconomic time series are sometimes considered to have a continuously varying standard deviation. If there is evidence of considerable heteroscedasticity, then the observations in the alarm statistic should have different weights. Sometimes the logarithm transformation is used for variance stabilization. This is the case here, where the observation X is the logarithm of the original observation. After this transformation, the variance is here assumed equal, as also in Andersson (2001). The surveillance is conducted and evaluated for the transformed variable X.

2.2 Event to be detected

An alarm system is developed in order to detect the turning point. At decision time s an alarm statistic is used to discriminate between $D(s) = \{\tau > s\}$ and $C(s) = \{\tau \leq s\}$, where τ is the unknown time when the underlying process μ changes from expansion to recession, or vice versa. Knowledge of whether the next turn will be a trough or a peak is assumed. The solutions for peak- and trough-detection are equivalent, as everything is symmetrical. It is the knowledge per se which is important. Henceforward, for simplicity, the turning point will be expressed as a peak (a transition from expansion to recession). This is not, however, a restriction in the methods. Thus, for a method that relies on the monotonicity restrictions only, the aim is to discriminate between the following two events:

$$\begin{aligned} &D(s): \mu(1) \leq \ldots \leq \mu(s) \\ &C(s): \mu(1) \leq \ldots \leq \mu(\tau\text{-}1) \text{ and } \mu(\tau\text{-}1) \geq \mu(\tau) \geq \ldots \geq \mu(s) \end{aligned} \tag{5}$$

where $\tau = \{1, 2, \ldots, s\}$ and at least one inequality is strict in the second part.

Under the assumption that the regression consists of linear functions where the slopes are symmetrical for the two phases, the aim is to discriminate between D and C, such that

$$D(s): \mu(s) = \beta_0 + \beta_1 \cdot s \qquad (6)$$
$$C(s) = \{\cup C(\tau)\},$$

where $C(\tau)$: $\mu(s) = \beta_0 + \beta_1 \cdot (\tau\text{-}1) - \beta_1 \cdot (s\text{-}\tau\text{+}1)$ and where $\tau=\{1, 2, ..., s\}$ and β_0 and β_1 are known constants.

If an HMM is assumed then, at decision time s, an alarm statistic is used to discriminate between

$$D(s): \mu(s\text{-}1) \leq \mu(s), \qquad (7)$$
$$C(s): \mu(s\text{-}1) > \mu(s).$$

The difference between D and C in (5) and (6) is only the assumptions regarding $\mu(t)$. However, when an HMM is assumed, as in (7), the events are different also in another aspect. The apparently simpler event in the HMM approach is combined with a more complicated situation for the information of previous states. No knowledge of previous states is utilized in an HMM approach. Thus the probabilities for the history of those earlier states will have an effect. The two events in (7) correspond to families of histories of states. Because of Markov dependence the probabilities for the histories of those earlier states will have an effect and earlier observations carry information of the history of states.

In this paper, the alarm statistics are based on the likelihood ratio between the events D and C, i.e.

$$\frac{f(x_s|C)}{f(x_s|D)},$$

where $x_s=\{x(1), \ldots, x(s)\}$ and f is the likelihood function. This is described further in Section 2.4.

2.3 Assumptions about transition probabilities

The probability of a transition from recession to expansion (or vice versa) are, in most approaches in the HMM framework, assumed to be constant with respect to time, see e.g. Hamilton (1989), Layton (1996) and Ivanova et al. (2000). But the transition probability can also be assumed time varying, as by Neftci (1982), Diebold, Lee and Weinbach (1994), Filardo (1994) and Layton and Katsuura (2001). One approach for the time varying transition probability is to use a model where the transition probability is a function of economic indicators, see e.g. Layton and Katsuura (2001). The assumption of a time invariant transition probability is made for all three methods investigated in this paper.

When the information about the time and type (peak or trough) of the last turning point is not utilized (as in (7) above), it is necessary to make a probability statement regarding the type of the next turning point as well as inference about whether the turning point has occurred or not. For this purpose, it is necessary to consider all previous possible turns (both peaks and troughs) and hence two transitions probabilities are needed in the monitoring system. The (common) situation when these two probabilities have different values is sometimes also referred to as asymmetry. The probabilities of transitions can also be expressed as intensities of the occurrences of peaks and troughs.

Contrary to the assumption of an HMM, the type of the next turn (peak or trough) is sometimes assumed known (e.g. Neftci (1982)). This is assumed in two of the methods here (see Section 2.4). When information about the type of the next turning point is used, then the events D and C are declared corresponding to that information and then it is sufficient with one measure of intensity in the monitoring system. The intensity, when not presented as Markov transition probabilities, is hereafter denoted ν. Note that the value of the intensity may differ for peak detection and trough detection. The single transition probability is

$$\nu = P(C(t)|D(t-1)) = P(\tau = t|\tau \geq t). \tag{8}$$

The assumption of a constant transition probability, and thus a geometric distribution for the turning point time, τ, is not very realistic in the business cycle application. The lengths of the cycles vary more than for a geometrical distribution and the probability of small values of the time for the turning point is much smaller for the business cycle than for a geometrically distributed variable. Kim and Nelson (1998) investigate duration dependence (if the tendency to switch state depends on the time spent in the current state). A test is carried out on four economic indicators in the US economy and the result is that there is a tendency to duration dependence.

In Fig. 1 the observed sample density function for the Swedish IP for the period 1970Q1 to 1987Q1 is compared to geometric density functions using different intensity estimates.

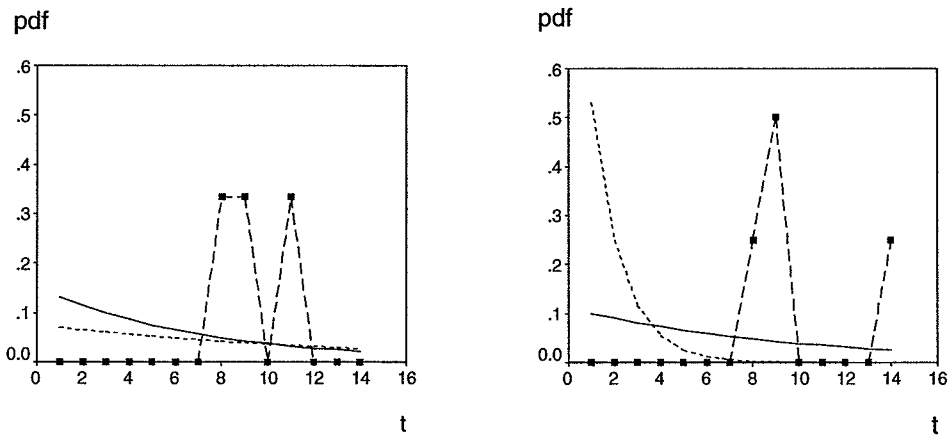

Fig. 1. Left: density for peak with observed sample density function (— — — ■) compared to geometric density function with intensities 0.07 (• • •) and 0.13 (——). Right: density for trough with observed data (— — — ■) compared to geometric density with intensities 0.53 (• • •) and 0.10 (——).

Since the observed distribution of τ is far from geometric, it could be argued that the empirical distribution should be used. However a prior based on the observed data above would result in a high prior probability for a turning point after about 9 quarters. The consequence would be that the influence of the actual data is reduced and the probability of an alarm after 9 quarters would be very high, only due to prior information.

For methods where the transition probabilities are assumed to be known and constant, estimation of the probabilities can be made using data from an earlier period. Since there can be several years between transitions in the business cycle (see Fig. 1), a very long time series is required to get reliable estimates.

In order to avoid the risk of misspecification, a non-informative prior for the turning point time can be used. When a non-informative prior for τ is used, no transition probability needs to be estimated. One approach that uses a non-informative prior is the so called Shiryaev-Roberts (SR) approach, suggested by Shiryaev (1963) and Roberts (1966). The likelihood ratio method is optimal for the intensity that is used, see Frisén and Wessman (1999), therefore a non-informative prior is suboptimal. However, the non-informative prior makes the method robust against assumptions regarding the intensity. Thus the risk of errors due to erroneous assumption, uncertain estimates or uncertain assumptions is avoided. The SR approach is used for two of the methods in this paper.

2.4 Alarm rules

All methods considered here use a likelihood ratio based alarm statistic. The likelihood ratio (LR) method has several optimal properties, see Frisén and de Maré (1991). The expected utility, based on very general functions of the gain of an alarm and the loss of a false alarm, is maximized. The LR method yields a minimum expected delay of an alarm signal conditional on a fixed probability of false alarm. In Frisén and Wessman (1999) several properties of the LR method are investigated and compared with other methods of surveillance, e.g. the Shewhart method and the CUSUM method, for the case of a shift in level.

The LR method signals an alarm for the first time s for which

$$\mathrm{LR}(s) = \frac{f(x_s|C)}{f(x_s|D)} > k_s,$$

where $k_s = k/(1-k) \cdot P(D(s))/P(C(s))$. Frisén and de Maré (1991) showed that the posterior probability approach is equivalent to the LR approach for the situation where C is the complement of D.

At the unknown time τ there is a turning point, i.e. there is a change in μ from expansion to recession (or visa versa). For the situation where μ are known functions under D and C (see (6)), the alarm rule for the likelihood ratio at time s is written as

$$\sum_{j=1}^{s} \frac{w_s(j) \cdot f(x_s|\mu = \mu^{C(j)})}{f(x_s|\mu = \mu^{D})} > k_s,$$

where $w_s(j) = \mathrm{P}(\tau=j)/\mathrm{P}(\tau \le s)$, $k_s = k/(1-k) \cdot \mathrm{P}(\tau > s)/\mathrm{P}(\tau \le s)$ and $\mu^{C(j)}$ and μ^{D} are parameters under restriction $C(j) = \{\tau = j\}$ and $D = \{\tau > s\}$.

The alarm rule using the non-informative prior approach, SR, is

$$\sum_{j=1}^{s}\frac{f(x_s|\mu=\mu^{C(j)})}{f(x_s|\mu=\mu^{D})} > k_{SRlin} \tag{9}$$

where k_{SRlin} is a constant alarm limit. The method where μ^D and $\mu^{C(j)}$ are modelled using known linear functions with a symmetric turning point and where the SR approach is used regarding the intensity is hereafter referred to as the *SRlin* method.

A method that assumes no knowledge of μ, other than the monotonicity restrictions in (5), and uses the SR approach regarding the intensity was suggested by Frisén (1994) and evaluated by Andersson (2002) and Andersson et al. (2002). The alarm rule is

$$\sum_{j=1}^{s}\frac{f(x_s|\mu=\hat{\mu}^{C(j)})}{f(x_s|\mu=\hat{\mu}^{D})} > k_{SRnp}, \tag{10}$$

where k_{SRnp} is a constant alarm limit and $\hat{\mu}^D$ and $\hat{\mu}^{C(j)}$ are the maximum likelihood estimators of the vector μ under monotonicity restrictions D and $C(j)$, described in section 3.1.1. This method is hereafter referred to as the *SRnp* method. In previous evaluations, this method has been referred to as the MSR method (Maximum likelihood Shiryaev-Roberts).

With an HMM approach (e.g. Hamilton 1989), the posterior probability is used in order to classify time points to the different regimes (expansion and recession), the classification being "recession" if the posterior probability exceeds a constant. Often the constant is chosen to be 0.5 (see e.g. Hamilton (1989) and Ivanova et al. (2000)). That is

$$P(C(s)|x_s) > 0.5. \tag{11}$$

The classification by the posterior probability can be used in prospective monitoring (see e.g. Neftci (1982)). By "the *HMlin* method" we hereafter refer to a monitoring method where the rule in (11) is used, together with the definition of the event C in (7) and the assumption that the differentiated series is constant in each regime according to (4). These conditions agree with those used by Koskinen and Öller (1998).

The time of alarm, t_A, is the first time for which the alarm statistic (in (9), (10) and (11) respectively) exceeds its specified alarm limit.

The approach by Birchenhall et al. (1999) is similar to both the HMM approach and the likelihood ratio method of surveillance in two respects: i) Birchenhall's approach is based on Bayes theorem and the likelihood and ii) a classification is made of the type of regime. A major difference, however, is that the classification into different regimes is based on explaining variables and not on the earlier state. This difference is discussed further in Andersson et al. (2002).

The constants in the alarm rules have to be determined. These constants are often set so that the false alarms are under control. The most common way in the general theory and practice of surveillance is to control the ARL^0, (the Average Run Length to the first alarm if the process has no turn). It is suggested by Hawkins (1992), Gan (1993) and Andersson (2002) that the control instead is made the MRL^0 (the median run length). This has several advantages, such as easier interpretations for the skewed distributions and much shorter computer time for calculations. One objection to using the limit 0.5 for the posterior probability by the HMM methods is that no direct conclusion can be made regarding the false alarms.

3. Estimation and special data problems

The discussion so far has been concerned with inferential problems relevant for all studies of detection of turning points in business cycles. However, for a specific application there are many important problems before you have the ideal data set to analyze. Some such problems will now be briefly discussed.

3.1 Estimation

An important aspect is which assumptions that are made about the process, since these assumptions determine how the parameters of the model are estimated. The parameters are often estimated using previous data. If the parameters are estimated from a short period the variance of the estimates will be large and the parameters might be severely misspecified. As a result the method will produce misguiding results, leading to wrong conclusions and decisions. Here we discuss assumptions and estimation of μ and the transition probabilities.

3.1.1 Estimation of μ

Different assumptions can be made about μ. In practice, when μ is assumed to be known, it is in fact estimated from a large enough set of data from an earlier period. If it is assumed that the differentiated series has a constant expected value, conditional on the state, and that the expected values are constant over the cycles, then the estimation can be made using previous data. This assumption is used by e.g. Neftci (1982), Layton (1996), Ivanova et al. (2000) and Layton and Katsuura (2001).

One suggested estimation procedure, under the model assumptions in (1), is to first classify each time point as belonging to either the expansion state or the recession state. Then the parameters in (3) can be estimated as

$$\hat{\beta}_j = \bar{d}_j = \frac{1}{n_j}\sum_{i=1}^{nj} d_{ji} \text{ ,}$$

where n_j = # time points classified as state j and d_j is the differentiated series classified as belonging to state j.

In order to avoid the rather strong assumption of a specific parametric function for μ, or when reliable information on the parametric function is not available, an approach based only on the knowledge that the monotonicity of μ changes at a turning point can be used. The result is the non-parametric estimation, used in the *SRnp* method, which is evaluated by e.g. Andersson (2002). The estimation is made using a least square criterion under the monotonicity restrictions C and D in (5). Under the model assumptions in (1), the estimates are also the maximum likelihood estimates.

For the peak detection situation, the event D implies a monotonically increasing μ-vector

$\{\mu(t) \leq \mu(t+1)\}$, $t \geq 1$.

In the likelihood expression $f\{x_s | \mu = \hat{\mu}^D\}$ in (10), $\hat{\mu}^D$ is the estimated parameter vector which corresponds to

$$\max_{\mu \in F^D} f(x_s | \mu),$$

where F^D is the family of μ-vectors such that $\{\mu(1) \leq \mu(2) \leq ... \leq \mu(s)\}$. Thus, $\hat{\mu}^D$ is the maximum likelihood estimator of μ under the monotonicity restriction D. For a trough, the estimation is made under the restriction $\{\mu(1) \geq ... \geq \mu(s)\}$. This estimator is described by e.g. Robertson, Wright and Dykstra (1988), p. 1-58.

The event C is composite, $C=\{\tau \leq s\}$, and thus we have $C = \{C(1), C(2), ..., C(s)\}$. In the peak detection situation $C(j)$ implies a vector that is first increasing and then decreasing, so that
$\{\mu(1) \leq ... \leq \mu(j\text{-}1), \mu(j\text{-}1) \geq \mu(j) \geq ....\}$, $j \in \{1, 2, ..., s\}$.

In the likelihood expression $f\{x_s | \mu = \hat{\mu}^{C(j)}\}$ in (10), $\hat{\mu}^{C(j)}$ is the estimated parameter vector which corresponds to

$$\max_{\mu \in F^{C(j)}} f(x_s | \mu),$$

where $F^{C(j)}$ is the family of μ-vectors such that $\{\mu(1) \leq ... \leq \mu(j\text{-}1)$ and $\mu(j\text{-}1) \geq \mu(j) \geq ....\}$, where $j = \{1, 2, ..., s\}$ and where at least one inequality is strict in the second part. Thus, $\hat{\mu}^{C(j)}$, $j \in \{1, 2, ..., s\}$, is the maximum likelihood estimator of μ under the monotonicity restriction $C(j)$. This estimator is given by Frisén (1986). For a trough, the estimation is made under the restriction $\{\mu(1) \geq ... \geq \mu(j\text{-}1)$ and $\mu(j\text{-}1) \leq \mu(j) \leq ...\}$.

When the information regarding μ is restricted to only monotonicity assumptions, then no previous data is required for estimation. The alarm statistic can be constructed directly from the data to be analyzed. One advantage of the non-parametric approach is that it can successfully be used also when reliable information on the parametric function is not available. Also important is that the non-parametric method does not assume that all phases of the same type have the same level and parametric shape. In practice, this varies a lot. The *SRnp* method, (10), only uses the monotonicity change and not the level. Wrong specifications of the slope lead to very bad properties (see Section 4.1.2), as was demonstrated by Andersson et al. (2002). The safe way by the non-parametric approach might be preferred.

3.1.2 Estimation of transition probabilities

When considering different methods of estimation, simultaneous maximum likelihood estimation of all parameters in the model is an obvious choice. However, if the whole parameter set is estimated using maximum likelihood criterion then the rareness of the turning points can lead to large errors around turning points, compensated by high accuracy within phases, as pointed out by Lahiri and Wang (1994) and Koskinen and Öller (1998). For that reason, the transition probabilities are sometimes estimated using some other criterion than maximum overall likelihood.

For the transition probabilities, maximum likelihood estimates, based only on the events of transitions, are a natural choice. For the data on the Swedish IP, the transition probabilities, p_{12} and p_{21}, are estimated using

$$\hat{p}_{12} = \frac{n_{12}}{n_{11}+n_{12}} = \frac{5}{34+5} = 0.13\ (0.054), \qquad (12)$$

$$\hat{p}_{21} = \frac{n_{21}}{n_{21}+n_{22}} = \frac{4}{4+36} = 0.10\ (0.047),$$

where n_{ij} is the number of transitions from state i to state j and the standard errors are given in parenthesis.

When the posterior probability for *HMlin* is calculated in the simulation study below, the maximum likelihood estimates of (12) are used.

Some authors, e.g. Koskinen and Öller (1998), use smoothing to reduce the stochastic error. They estimate the transition probabilities and the smoothing parameter simultaneously from historical information, with the criterion of minimizing a cost-function based on the sum of two measures of error (the Brier probability score and the proportion of wrongly classified states). The Brier probability score, also referred to as the Quadratic probability score, is the mean square error for the posterior probability, i.e. the average squared deviation between the true state (0 or 1) and the posterior probability. The effect of smoothing is investigated further in Section 3.2.1.

3.2 Special data problems

Always when using real data, there can be concerns about the data quality in general. In this section we discuss the problems connected with analyzing data that has been transformed (e.g smoothed, adjusted for seasonality and adjusted for trend). The section is also concerned with the problems when data exhibit autoregression and using multivariate data.

3.2.1 Effect of smoothing

One drawback with using the alarm limit 0.5 for the posterior probability is that the value 0.5 is not chosen to yield a certain ARL^0 or MRL^0. In other words, we have no control of the false alarm rate. The variability of the process affects the false alarm rate. In order to reduce the false alarm rate some authors, e.g. Koskinen and Öller (1998), recommend that the observations should be smoothed after differentiation, see also Öller (1986). The objective of the smoothing, in Koskinen and Öller (1998), is both reduction of white noise and lagging of the turning point. The former purpose is motivated by the desire to reduce the false alarm rate whereas the latter purpose is motivated by the use of multivariate data, where the turning points of the different processes are not always synchronized. The differentiated observations $y(t)$ are smoothed according to

$$\tilde{Y}(t) = \lambda Y(t) + (1-\lambda)\tilde{Y}(t-1),$$

where $\lambda \in \{0, 1\}$.

Smoothing by kernel estimators is used by e.g. Hall, Marron and Titterington (1995). The smoothing of observations reduces the variance and hence reduces the false alarm probability. However, there are also disadvantages as will be seen in the

Monte Carlo study below. Results are given on the distribution of the alarms, both conditional of no change and of a turn, and the probability of detecting a turning point within a specified time.

The effect of smoothing on the performance of the *HMlin* method is investigated. In order to evaluate the false alarm properties, a model for the event D (no turn) has to be specified. Data on the (logarithm of the) Swedish IP was used to get a reasonable simulation model. A linear function was fitted to the officially dated (National Institute of Economic Research (1992)) expansion phase 1987Q2 to 1989Q3. The observations on X, under event D, are simulated using the following model

$$X^D(t) = \mu^D(t) + \varepsilon(t), \tag{13}$$

where $\mu^D(t) = 11.194 + 0.0069 \cdot t$ and $\varepsilon(t) \sim$ iid N[0; 0.016]. In order to evaluate the alarm properties under event C the following model is used:

$$X^C(t) = \mu^{C\tau}(t) + \varepsilon(t), \tag{14}$$

where $\mu^{C\tau}(t) = 11.194 + 0.0069 \cdot t - 2D_1 \cdot 0.0069 \cdot (t-\tau+1)$, $t = \{1, 2, ...\}$, and

$$D_1 = \begin{cases} 1, t \geq \tau \\ 0, \text{otherwise} \end{cases}$$

and $\varepsilon(t) \sim$ iid N[0; 0.016].

Both the *SRlin* method and the *HMlin* method use information from previous expansion and recession phases in the alarm statistic and for parameter estimation. Thus, a sequence of three expansion phases and four recession phases are simulated by the following model

$$X_i(t) = \mu_i(t) + \varepsilon_i(t)\,,$$

$$\text{where } \mu_i(t) = \begin{cases} \mu_{1j}(t), & j = \{1, 2, 3\} \\ \mu_{2j}(t), & j = \{1, 2, 3, 4\}\} \end{cases}$$

$$\text{and } \varepsilon_i(t) \sim iid \begin{cases} N(0; \sigma_{1j}^2), & j = \{1, 2, 3\} \\ N(0; \sigma_{2j}^2), & j = \{1, 2, 3, 4\}\}. \end{cases}$$

The function $\mu_{1j}(t)$ and parameter σ_{1j} represent the expected value and the standard deviation in expansion phase j, respectively and $\mu_{2j}(t)$ and σ_{2j} represent the expected value and the standard deviation in recession phase j, respectively (see Appendix 1). The evaluation period (t=1) starts 4 time points after the last turning point (a trough).

The effect of smoothing on the *HMlin* method is shown below. The alarm limit for the smoothed data (λ=0.3) is adjusted to yield the same MRL^0 as for the unsmoothed data (λ=1).

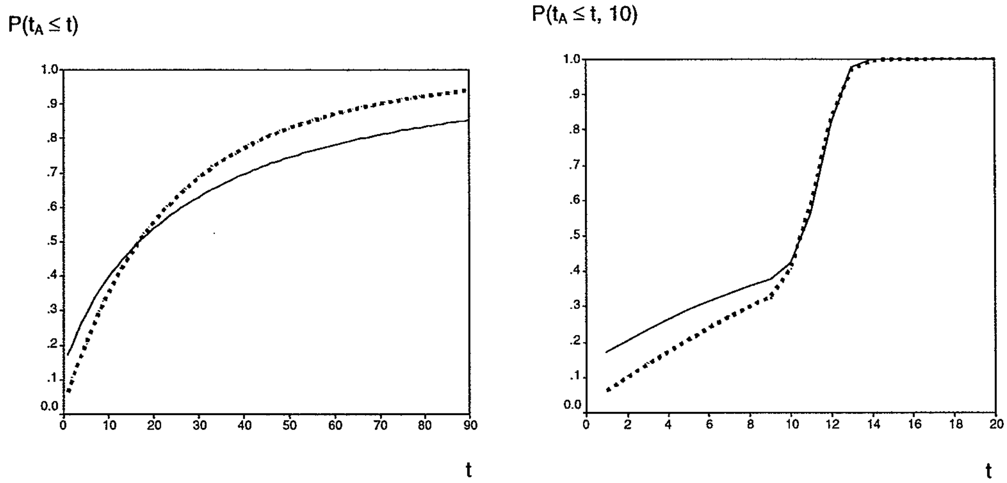

Fig. 2. HMlin method, λ=1 (···) and λ=0.3 (——). Left: Distribution of the time of an alarm, conditional on event D (no turn). Right: Distribution of the time of an alarm, conditional on τ=10..

The common value of the MRL^0 is 17. For both cases, λ={1, 0.3} we have a large enough sample so that the standard error of the median is less than 0.15. We can see in Fig. 2 that the density of the time of the alarm has different skewness for λ=1 and λ=0.3.

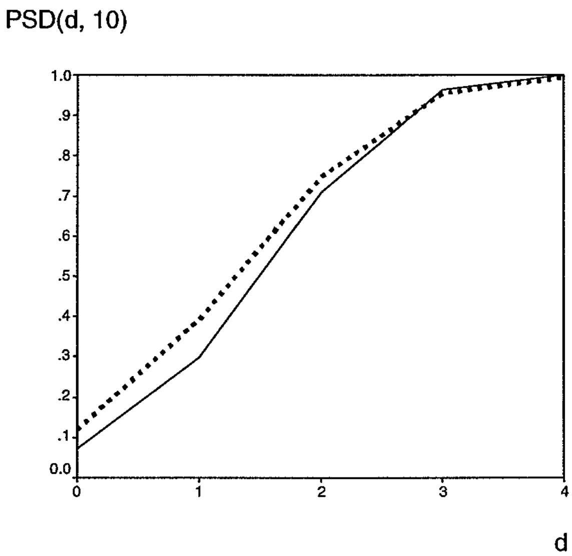

Fig. 3. Probability of successful detection within d time points for τ=10. HMlin method, λ=1 (···) and λ=0.3 (——).

The probability of successful detection (PSD) is the probability of detecting a turn within d time units, that is

$$P\left((t_A - \tau) \le d \middle| t_A > \tau = \tau_0\right).$$

For both cases, λ={1, 0.3}, the number of replicates is large enough to yield a standard error of PSD of less than 0.0030. The reduced distinctness of the turning point, due to smoothing, decreases the probability of successful detection (see Fig. 3).

3.2.2 Seasonal variation

The variables that are considered to be leading economic indicators are measured monthly or quarterly and thus often contain seasonal variation, which could complicate the monitoring. The seasonal variation can be considerable, as is seen in Fig. 4.

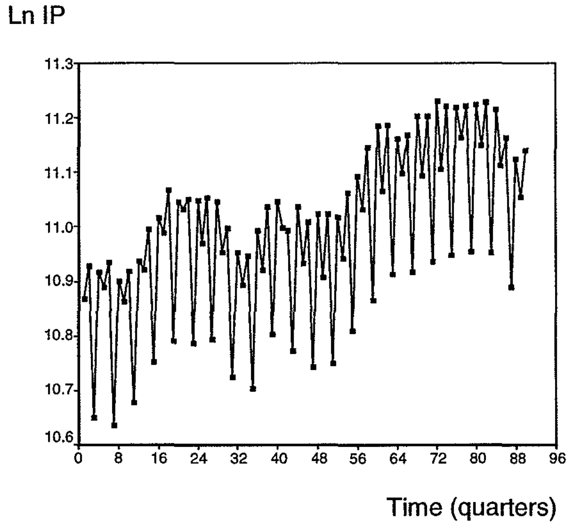

Fig. 4. Swedish Industrial Production, quarterly data (1970Q1 to 1992Q2).

If seasonality is neglected in the modeling and in the monitoring, it could lead to seriously wrong conclusions. It is important that the structure of the original series is not disturbed by the seasonal adjustment. In the monitoring situation here, it is important that the time of the turns is preserved after the adjustment. The problem of altered change points by seasonal adjustment has been briefly discussed, see e.g. Ghysel and Perron (1996) and Franses and Paap (1999).

The effect of using different filters in order to adjust for seasonality is analyzed in Andersson and Bock (2001) and it is demonstrated that the detection is delayed when a data transformation such as differentiating or moving average is used. The largest reduction in probability of detection is caused by the moving average.

3.2.3 Autoregression

Economic time series often exhibit strong autocorrelation. In Ramjee, Crato and Ray (2002) the performance of the EWMA method for detecting a shift in level in a process with long range dependence is discussed. Lahiri and Wang (1994) evaluate the performance of a monitoring system where a model with autoregressive errors is assumed and where the posterior probability is used together with an alarm limit. The same alarm limit is used for models with autoregressive errors of different order. They find that the introduction of autoregression in the errors leads to a smaller forecast error within phases but increases the risk of wrong inference concerning turning points. The effect of time dependent observations can be dealt with by adjusting the alarm limit. If the assumption of an independent process is used, when it is in fact dependent over time, the result is an increased false alarm probability. The consequence of autoregression in the process is examined in the general theory of surveillance where also remedies are suggested. For a review, see Pettersson (1998) and Frisén (2002).

Ivanova et al. (2000) argue that the effect of the autoregressive parameters will largely be captured by the probabilities of remaining in the current state (p_{11} or p_{22}). Many of the suggested approaches for turning point detection assume that the possible autoregression is not a problem, as is also assumed here.

3.2.4 Adjusting for trend

Many macroeconomic variables can be characterized as cyclical movements around a trend. In order to distinguish the movements and make the time series stationary it is sometimes necessary to adjust for the trend. In model (1) no separation between the trend and the cycle is made. In most HMM approaches, and also the one considered here (*HMlin*), differentiation is used, whereas in *SRlin* and *SRnp* no such adjustment is made.

The choice of a method for trend adjustment should depend on the assumptions regarding the trend-component. Whether the trend is assumed to have components that are deterministic, stochastic or both, has implications for the appropriate method. Adjusting for trend implies a data transformation, which may result in a distortion of the characteristics of the original series. Gordon (1997) studies the effect of trend removal for predictive densities of the US GDP and warns against using other information from the data than that which is directly associated with the business cycle turning points. Canova (1998) discusses trend removal and evaluates the effect using several different approaches, among them first order differentiating. One conclusion from the study is that linear trend removal does result in turning point times which do not correspond to the official turning point times of the National Bureau of Economic Research, USA. In another paper Canova (1999) points out that previous research has shown that the trend may interact with the cyclical component and is therefore difficult to isolate. The general conclusion is that statements concerning the turning points are not independent of the statistical assumptions needed to extract trends. Boone and Hall (1999) compare different methods for separating trend and cycle by a Monte Carlo study. The mean squared differences are used as a measure of evaluation. A difficulty with this measure, in a turning point setting, is that it does not primarily indicate whether the turning point times are preserved.

The removing of the trend has less effect on the possibility to distinguish the turning points when analyzing short time series. Since the *SRlin* and *SRnp* methods are applied to a part of the time series that contains one turning point at most, no attempt to separate the trend from the cycle is made.

3.2.5 Multivariate problems

By the common movement approach, a business cycle is characterized as the cyclical movement of many economical activities. This demonstrates that important information is contained in the relation between the turns of different indices. This information can be utilized, either by transforming the problem to a univariate one (by using a composite index of leading indicators) or by applying a multivariate method of surveillance.

Stock and Watson (1991) and Stock and Watson (1993) model the common movements of coincident variables as arising from an unobservable common factor (the overall state of the economy). The key elements are the selection of variables and the estimation of the common factor. Leading indicators are added to the model to help predict future values of the common factor.

Diebold and Rudebusch (1996) discuss the relation between the common movement approach and the regime approach and attempt to encompass both approaches by considering the common movements of coincident variables where the common factor is assumed to be governed by a two-state HMM.

Kim and Nelson (1998) use the same approach as in Diebold and Rudebusch (1996). However, estimation is here made by Gibbs sampling in a Bayesian framework. They find that the main cause of increase in forecast accuracy is the ability to capture the common movement among several variables instead of just one, whereas the prior assumptions concerning the transition probabilities had a minor influence. Hamilton and Perez-Quiros (1996) compare the accuracy in predicting the phases of U.S. real gross national product using univariate and bivariate linear models, where the latter included a composite leading index (CLI), and corresponding HMM. Adding a CLI to the linear model was found to result in the greatest increase in accuracy, whereas using HMM makes no substantial increase in accuracy.

Koskinen and Öller (1998) utilize multivariate information by monitoring a joint vector of leading indicators with a common time of turn. When used for turning point detection in the Swedish business cycle the following three series are used: the Swedish IP, the Swedish Business Tendency Survey and Stockholm Stock Exchange Index. When the method is applied to the U.S. economy, the (first difference of) Gross National Product and the Composite Index of Leading indicators are used.

Birchenhall et al. (1999) exploit the feature of a business cycle, of common movements across variables, by extracting a business-cycle index from a vector of time series. As in the works by Stock and Watson, the selection of variables is an important element.

Kontolemis (2001) uses an HMM approach to identify turning points, using four different time series. A comparison is made between turning point identification based on the individual series and a multivariate approach. It is shown that the business cycle chronology based on the latter approach is closer to that of NBER than the turning points obtained from individual series.

Wessman (1998) demonstrates that the minimal sufficient alarm statistic for changes in several variables with the same change point (or known time-lag), is univariate. In this paper the investigation has been made for the turning point of one leading index, possibly constructed as a function of many different indices. This is, in fact, the situation also for most of the earlier studies since a reduction to a univariate statistic is possible. However, procedures that are more efficient might be constructed by using the indices separately in the method of surveillance.

For reviews on multivariate surveillance, see Wessman (1999) and Frisén (2002). A topic for future research might be to use more of the theory of optimal multivariate surveillance for building a system for multivariate monitoring of business cycles.

4 A comparison of some likelihood based approaches for detecting a turning point in a leading index

4.1 Statistical properties

In Andersson et al. (2002) an evaluation was made of the properties of the three methods by means of a Monte Carlo study. The major results are summarized here. Since two of the methods, *HMlin* and *SRlin*, rely on assumptions of known parameter values, evaluation was made both for the situation when the correct parameter values were used and for the situation when the values were misspecified.

4.1.1 Correct parameter values used

For all three methods the alarm limits are set to yield $MRL^0 = 17$ (quarters), but the distribution of the alarm times differs. The *HMlin* method has more frequent alarms at early time points, but low alarm probability later on, compared to the others.

The expected delay measures the delay time for a motivated alarm (i.e. the time between τ and t_A). This measure depends on τ. For a turning point within a year from the last turn, the *SRnp* method has the longest expected delay. For turning points that occur later, the *SRnp* method has a (slightly) shorter delay than the *HMlin* method. *SRlin* has the shortest delay time for every value of τ.

In a practical situation it is important to have a strategy for what action to take when a turning point is signalled. The predictive value (see Frisén (1992)) reflects the trust you should have in an alarm. For the *HMlin* method the high alarm probability at the first time point results in that alarms at t=1 are of little value, whereas the predictive value for *SRnp* and *SRlin* at this time have much higher predictive values. The alarms that come at time points t=4 and hence forward have predictive values of (at least) 0.75 for all three methods.

4.1.2 Incorrect parameter values used

The effect of using wrong parameter values for μ was evaluated both for the situation when only the slope after the turn was misspecified and for the situation when both slopes (pre-turn and post-turn) were misspecified. Also in this evaluation the alarm limits are set to yield $MRL^0 = 17$ (quarters).

When only the post-turn slope is misspecified, it has very little effect on the conditional expected delay and the predictive value.

When both slopes (pre-turn and post-turn) are misspecified, the effect on the conditional expected delay and the predictive value is major. For small and moderate values of τ, (τ<10), the delay time is longer when the slopes are specified as being too steep. Thereafter the delay time is shorter for misspecified slopes. The price for the short delay times is however that the predictive value of those alarms is low.

In view of these results, using a method that does not require any parametric values for μ is a safe way, particularly since the properties of the *SRnp* method are almost as good as those of the *SRlin* method.

4.2 Evaluation by data on the Swedish Industrial Production

The most common way to evaluate methods for detection of turning points in business cycles is by using one set of data. We use quarterly Swedish data on IP (1987Q2 to 1992Q2) with one turning point, to evaluate the three methods (*SRnp*, *HMlin* and *SRlin*). The aim is to detect the peak in the time series.

According to official records (National Institute of Economic Research (1992)), the peak occurs at time 1989Q3 (t = 10), implying that the time of change is 1989Q4 (τ=11).The period is displayed in Fig. 5. The official turning point times can often be based on more information than data. This other information might make the official time different than it should have been, if only the IP data had been used. Fig. 5 indicates that the turning point in the data is earlier than the official time. Thus the methods, using only the IP data, can not be expected to be good at indicating the recorded official time for this realization. All three methods give alarms earlier than the official times for this set of data.

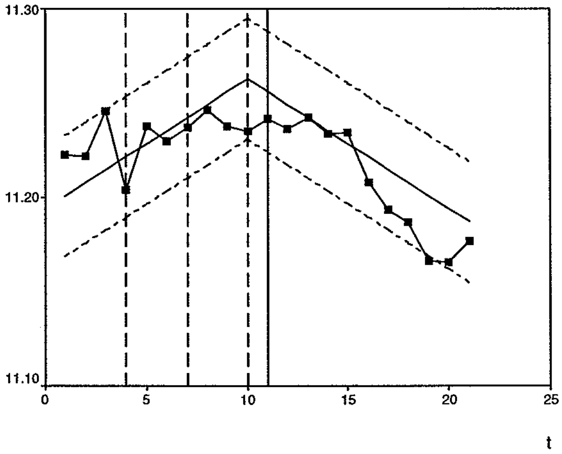

Fig. 5. Seasonally adjusted values of the (logarithm of) Swedish industrial production, quarterly data for the period 1987Q2:1992Q2. The official time of change (11) is marked with a solid vertical line. The alarm times (t_A={4, 7, 10} for SRnp, HMlin and SRlin, respectively) are marked with dashed vertical lines. The model for μ, used by SRlin and HMlin, with τ=11 is marked with a solid curve (see (14)). The 2.5th and 97.5th percentiles of values of the observations according to the model is marked as dotted curves (• • •).

Both the *SRlin* method and the *HMlin* method use the assumption of a piecewise linear model for μ (see (3)). The piecewise linear model fits less well at the turning point as we have a plateau. McQueen and Thorley (1993) argue that it is reasonable that recessions tend to be preceded by plateaus. A plateau will result in a tendency to give alarms just before the turn. It can be discussed whether this is a drawback or not. An early indication of a coming recession is a plateau. In this light, alarms just before the turning point can be considered to be good, since they can be seen as warnings.

SRlin and *HMlin* rely on the assumption that the parameters (slopes and standard deviations) are known or possible to estimate with great certainty. Here we use the data on Swedish IP from the period 1970Q1:1987Q1 and the estimation procedure described in Appendix 1 to estimate the parameters. The resulting signal-noise ratios are: $\hat{\beta}_1 / \hat{\sigma}_1$ = 0.47 (expansion phase) and $\hat{\beta}_2 / \hat{\sigma}_2$ = 0.40 (recession phase). These estimates are used for the *HMlin* method when calculating the posterior probability. The *SRlin* method assumes a symmetric turning point and homoscedasticity. We use pooled (by their frequency) estimates for both parameters, resulting in $\hat{\beta} / \hat{\sigma}$ = 0.41, when calculating the alarm statistic. The HMlin method also includes two transition probabilities, whose values are estimated according to (12).

A drawback with a simulation study is that the model used in the simulations might not be representative of the process we want to study. An actual data set is certainly representative of the specific time period and situation at hand. However, the real data set might deviate stochastically from the process of interest. When a method is intended to be used on future data, then it is the properties of the process that are important for that method. In order to be able to make a general statement about, for example, the average delay time of a method, it is necessary to replicate the performance of the method at a turning point. Then Monte Carlo methods are a valuable tool. However, it is important that the model, used in the simulations, is based on a set of real data.

5. Discussion

The user of a monitoring system is faced with many difficult decisions. Issues that must be considered when analyzing a real data set are: choice of estimation procedure, how to handle seasonal variation and whether to apply different kinds of transformations of data (e.g. trend removal and smoothing).

The smoothing of the observations before applying a method of monitoring will reduce the variation and hence reduce the false alarm probability. However, the smoothing will introduce autocorrelation and, as pointed out by Öller and Tallbom (1996), will lead to a delayed signal. In this paper it is confirmed that the expected delay of a motivated alarm is increased by smoothing (when the false alarms are controlled). One method of surveillance, where the smoothing procedure is included in the alarm statistic and not made separately, is the *EWMA* method (see e.g. Crowder (1987), Domangue and Patch (1991) and Frisén and Sonesson (2001)). This method allows for a controlled false alarm rate at the same time as the variability is reduced by smoothing.

The transition probabilities can be estimated separately, using only the observed transition frequencies. It has also been suggested that both transition probabilities and smoothing parameters are estimated simultaneously. The parameters then compensate for each other. If data is much smoothed then the transition probability must be set very high, otherwise the delay will be too long. Thus it might be difficult to interpret the parameters separately. If the Brier probability score is used as a criterion when estimating the transition probabilities, it must be borne in mind that this measure does not take into account the order of the observations. As a result the transition probabilities might again be difficult to interpret.

Often a constant transition probability is assumed, which implies that the time of the turn has a geometric distribution. This might not be in accordance with reality for business cycles, but can be interpreted as a way of avoiding to use strong assumptions regarding the intensity of turns. Good estimates of the transition probabilities are useful if the pattern is constant over time and will remain the same, even in the future. Technically, the inclusion of transition probability estimates in the monitoring system is easily done by likelihood ratio methods. However, it is important that the monitoring system has the ability to detect a turning point also when this happens at an unexpected time. Thus, it might be preferred to use a non-informative prior for the time of the turn in the suggested *SRlin* and *SRnp* methods, so as to avoid the risk of errors due to wrong assumptions or uncertain estimates.

When estimating the parameters of the monitoring system, historical data is often used. Then there has to be a balance between, on one hand, the risk of using data sets that are too small, which results in estimates with large variation and, on the other hand, the risk of using historical data which might be out of date. Sarlan (2001) examines the change in intensity and duration of US business cycles and concludes that the modern business cycle is different from the historical one.

Parametric models contain information, which should be used whenever it is reliable. However, a non-parametric approach works also when such reliable information is not available. At early time points the non-parametric method (SRnp) is sensitive to extreme observations, since this method only relies on data and uses no prior information regarding μ. Andersson et al. (2002) demonstrated the danger of misspecifying μ: even a small misspecification of the parameter value of μ may cause largely misguiding results.

Economic time series often exhibit seasonal variation. Since most data-driven filters can have serious effects on the turning point times (see e.g. Andersson, 2001), information from historical data or other prior knowledge, which makes the seasonal adjustment independent of the data to be analyzed, is very valuable.

There are different methods for evaluating a method of surveillance. In this paper, the effect of smoothing is evaluated by simulation, whereas the monitoring methods are evaluated with a set of real data; a period of the Swedish IP. Evaluation of the properties of a method by one sample of real data is difficult. One difficulty is to know whether the turnout of the sample is typical. Evaluations by several real data sets (instead of just one) would decrease some of the stochastic variation in the measures of evaluation. However, if these analyses are not totally independent (for example if the same parameter estimates are used) then some of the stochastic components would keep their variance.

The results from the application of the three methods on a period of the Swedish IP did not contradict the conclusion by the statistical properties that the non-parametric method is a safe way without much loss of efficiency.

Acknowledgements

Constructive comments by Lars-Erik Öller, National Institute of Economic Research, Stockholm, Sweden, are gratefully acknowledged. This work was supported by the Swedish Council for Research in the Humanities and Social Sciences.

References

Andersson, E. (2001). Turning point detection using non-parametric statistical surveillance: Evaluation of some influential factors, *Research Report 2001:3,* Department of Statistics, Göteborg University, Sweden.

Andersson, E. (2002). Monitoring cyclical processes - a nonparametric approach, *Journal of Applied Statistics*, 29, (7), 973-990.

Andersson, E. and Bock, D. (2001). On seasonal filters and monotonicity, *Research Report 2001:4,* Department of Statistics, Göteborg University, Sweden.

Andersson, E., Bock, D. and Frisén, M. (2002). Statistical surveillance of cyclical processes: Detection of turning points in business cycles, *Research Report 2002:8,* Research report 2002:8, Department of Statistics, Göteborg University, Sweden.

Birchenhall, C. R., Jessen, H., Osborn, D. R. and Simpson, P. (1999). Predicting US business-cycle regimes, *Journal of Business & Economic Statistics*, 17, (3), 313-323.

Boone, L. and Hall, S. G. (1999). Signal extraction and estimation of a trend: A Monte Carlo study, *Journal of Forecasting*, 18, 129-137.

Canova, F. (1998). Detrending and business cycle facts, *Journal of Monetary Economics*, 41, 475-512.

Canova, F. (1999). Does detrending matter for the determination of the reference cycle and the selection of turning points?, *The Economic Journal*, 109, (January), 126-150.

Crowder, S. (1987). A simple method for studying run-length distributions of exponantially weighted moving average charts, *Technometrics*, 29, 401-407.

Diebold, F. X., Lee, J.-H. and Weinbach, G. C. (1994). Regime Switching with Time-varying Transition Probabilites. In: C. Hargreaves (ed.), *Nonstationary Time Series Analysis* Oxford; Oxford University Press, pp. 283-302.

Diebold, F. X. and Rudebusch, G. D. (1989). Scoring the leading indicators, *Journal of Business*, 62, 369-91.

Diebold, F. X. and Rudebusch, G. D. (1996). Measuring business cycles: A modern perspective, *Review of Economics and Statistics*, 78, (1), 67-77.
Domangue, R. and Patch, S. C. (1991). Some omnibus exponentially weighted moving average statistical process monitoring schemes, *Technometrics*, 33, 299-313.
Falk, B. (1986). Further Evidence on the Asymmetric Behavior of Economic Time Series over the Business Cycle, *Journal of Political Economy*, 94, (51), 1097-1109.
Filardo, A. J. (1994). Business-Cycle Phases and Their Transitional Dynamics, *Journal of Business & Economic Statistics*, 12, (3), 299-308.
Franses, H. P. and Paap, R. (1999). Does Seasonality Influence the Dating of Business Cycle Turning Points?, *Journal of Macroeconomics*, 21, (1), 79-92.
Frisén, M. (1986). Unimodal regression, *The Statistician*, 35, (4), 479-485.
Frisén, M. (1992). Evaluations of Methods for Statistical Surveillance, *Statistics in Medicine*, 11, 1489-1502.
Frisén, M. (1994). Statistical Surveillance of Business Cycles, *Research Report 1994:1 (Revised 2000),* Department of Statistics, Göteborg University, Sweden.
Frisén, M. (1999). Characterization of methods for surveillance by optimality, *Research Report 1999:13,* Department of Statistics, Göteborg University, Sweden.
Frisén, M. (2002). Statistical surveillance: Optimality and methods, *Research Report 2002:2,* Department of Statistics, Göteborg University, Sweden.
Frisén, M. and de Maré, J. (1991). Optimal Surveillance, *Biometrika*, 78, 271-80.
Frisén, M. and Sonesson, C. (2001). Optimal surveillance by EWMA, *Research Report 2000:7,* Department of Statistics, Göteborg University, Sweden.
Frisén, M. and Wessman, P. (1999). Evaluations of likelihood ratio methods for surveillance: Differences and robustness, *Communications in Statistics. Simulations and Computations*, 28, (3), 597-622.
Gan, F. F. (1993). An optimal design of EWMA control charts based on median run-length, *Journal of Statistical Computation and Simulation*, 45, (3-4), 169-184.
Ghysel, E. and Perron, P. (1996). The Effect of Linear Filters on Dynamic Time Series With Structural Change, *Journal of Econometrics*, 70, 69-97.
Gordon, S. (1997). Stochastic trends, deterministic trends, and business cycle turning points, *Journal of Applied Econometrics*, 12, 411-434.
Hackl, P. and Westlund, A. H. (1989). Statistical analysis of "Structural Change", *Empirical Economics*, 14, 167-192.
Hall, P., Marron, J. S. and Titterington, D. M. (1995). On partial local smoothing rules for curve estimation, *Biometrika*, 82, (3), 575-587.
Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle, *Econometrica*, 57, 357-384.
Hamilton, J. D. and Perez-Quiros, G. (1996). What Do the Leading Indicators Lead?, *Journal of Business*, 69, (1), 27-49.
Hawkins, D. L. (1992). Detecting shifts in functions of multivariate location and covariance parameters, *Journal of Statistical Planning and Inference*, 33, 233-244.
Hussey, R. (1992). Nonparametric Evidence on Asymmetry in Business Cycles Using Aggregate Employment Time Series, *Journal of Econometrics*, 51, (1-2), 217-231.
Ivanova, D., Lahiri, K. and Seitz, F. (2000). Interest rate spreads as predictors of German inflation and business cycles, *International Journal of Forecasting*, 16, (1), 39-58.
Jun, D. B. and Joo, Y. J. (1993). Predicting turning points in business cycles by detection of slope changes in the leading composite index, *Journal of Forecasting*, 12, 197-213.

Kim, C.-J. and Nelson, C. R. (1998). Business cycle turning points, a new coincident index, and tests of duration dependence based on a dynamic factor model with regime switching, *Review of Economics and Statistics*, 80, (2), 188-201.
Kontolemis, Z. G. (2001). Analysis of the US business cycle with a vector-Markov-switching model, *Journal of Forecasting*, 20, (1), 47-61.
Koskinen, L. and Öller, L.-E. (1998). A Classifying Procedure for Signalling Turning Points, *Research Report 59,* National institute of economic research. Konjunkturinstitutet, Ekonomiska rådet, Stockholm, Sweden.
Lahiri, K. and Wang, J. G. (1994). Predicting cyclical turning points with a leading index in a Markov Switching model, *Journal of Forecasting*, 13, 245-263.
Lai, T. L. (1995). Sequential Changepoint Detection in Quality-Control and Dynamical-Systems, *Journal of the Royal Statistical Society B*, 57, (4), 613-658.
Layton, A. P. (1996). Dating and predicting phase changes in the U.S. business cycle, *International Journal of Forecasting*, 12, 417-428.
Layton, A. P. (1998). A further test of the influence of leading indicators on the probability of US business cycle phase shifts, *International Journal of Forecasting*, 14, (1), 63-17.
Layton, A. P. and Katsuura, M. (2001). Comparison of regime switching, probit and logit models in dating and forecasting US business cycles, *International Journal of Forecasting*, 17, (3), 403-417.
Li, D. T. and Dorfman, J. H. (1996). Predicting turning points through the integration of multiple models, *Journal of Business and Economic Statistics*, 14, (4), 421-428.
McQueen, G. and Thorley, S. (1993). Asymmetric business cycle turning points, *Journal of Monetary Economics*, 31, 341-362.
National Institute of Economic Research (1992). Konjunkturläget Maj 1992, *Research Report* National Institute of Economic Research, Stockholm, Sweden.
Neftci, S. (1982). Optimal prediction of cyclical downturns, *Journal of Economic Dynamics and Control*, 4, 225-41.
Neftci, S. (1984). Are Economic Time Series Asymmetric over the Business Cycle?, *Journal of Political Economy*, 92, (2), 307-328.
Pettersson, M. (1998). Evaluation of some methods for statistical surveillance of an autoregressive process, *Research Report 1998:4,* Department of Statistics, Göteborg University, Sweden.
Ramjee, R., Crato, N. and Ray, B. K. (2002). A note on moving average forecasts of long memory processes with an application to quality control, *International Journal of Forecasting*, 18, (2), 291-297.
Roberts, S. W. (1966). A Comparison of some Control Chart Procedures, *Technometrics*, 8, 411-430.
Robertson, T., Wright, F. T. and Dykstra, R. L. (1988). *Order Restricted Statistical Inference*, John Wiley & Sons Ltd.
Sarlan, H. (2001). Cyclical aspects of business cycle turning points, *International Journal of Forecasting*, 17, (3), 369-382.
Shiryaev, A. N. (1963). On optimum methods in quickest detection problems, *Theory of Probability and its Applications*, 8, 22-46.
Srivastava, M. S. and Wu, Y. (1993). Comparison of EWMA, CUSUM and Shiryayev-Roberts Procedures for Detecting a Shift in the Mean, *Annals of Statistics*, 21, 645-670.
Stock, J. H. and Watson, M. W. (1991). A Probability Model of the Coincident Economic Indicators. In: K. Lahiri and G. H. Moore (eds.), *Leading Economic Indicators: New Approaches and Forecasting Records* Cambridge; Cambridge University Press, pp. 63-89.

Stock, J. H. and Watson, M. W. (1993). A Procedure for Predicting Recessions with Leading Indicators: Econometric Issues and Recent Experience. In: J. H. Stock and M. W. Watson (eds.), *Business Cycles, Indicators and Forecasting* 28, Chicago; University of Chicago Press, pp. 95-156.

Wessman, P. (1998). Some Principles for surveillance adopted for multivariate processes with a common change point, *Communications in Statistics. Theory and Methods*, 27, 1143-1161.

Wessman, P. (1999). Studies on the surveillance of univariate and multivariate processes, *PhD Thesis*, Department of Statistics, Göteborg University, Sweden.

Westlund, A. H. and Zackrisson, U. (1986). On the prediction of structurally varying systems, *Technological Forecasting and Social Change*, 30, 63-72.

Wetherhill, G. B. and Brown, D. W. (1991). *Statistical Process Control: Theory and Practice*, London: Chapman and Hill.

Zarnowitz, V. and Moore, G. H. (1982). Sequential signals of recessions and recovery, *Journal of Business*, 55, 57-85.

Zellner, A., Hong, C. and Min, C.-K. (1991). Forecasting Turning Points in International Output Growth Rates Using Bayesian Exponentially Weighted Autoregression, Time-Varying Parameter, and Pooling Techniques, *Journal of Econometrics*, 49, (1-2), 275-304.

Öller, L.-E. (1986). A note on exponentially smoothed seasonal differences, *Journal of Business and Economic Statistics*, 4, 485-489.

Öller, L.-E. (1987). Stabiliseringspolitik och prognoser, *Ekonomiska samfundets tidskrift*, 40, 23-27.

Öller, L.-E. and Tallbom, C. (1996). Smooth and timely business cycle indictors for noisy Swedish data, *International Journal of Forecasting*, 12, 389-402.

Research reports from Department of Statistics, Gothenburg, Sweden can be ordered from the department (address as for corresponding author). Working papers from National institute of Economic research can be ordered from Konjunkturinstitutet, Stockholm, Sweden.

The data (industrial production) and computer programs for HMlin, SRlin and SRnp can be obtained from the corresponding author.

Appendix 1. Estimation for the HMlin method, using previous data

For the *HMlin* method several parameters are estimated from previous data, for which a simulation model is used. For each expansion and recession phase, a second degree polynomial regression function with seasonal dummy variables is fitted to data on the ln(IP). After that, the polynomials are intercept-adjusted, not to expose any jumps.

Results are given in Table 1A for each of the four recession phases and each of the three expansion phases, according to the model (without seasonal variation)

$$y_{ij}(t) = \mu_{ij}(t) + \varepsilon_{ij}(t) = (\theta_{0ij} + \theta_{1ij} \cdot t + \theta_{2ij} \cdot t^2) + \varepsilon_{ij}(t),$$

where i = {expansion} or {recession} and j = {1, 2, 3, 4}.

For the dating of recession and expansion phases, official records are used (National Institute of Economic Research (1992)), see Table 1A.

Table 1A
Dating of recession and expansion phases, official records are used (National Institute of Economic Research (1992))

i	J	θ_0	θ_1	θ_2	sd[$\varepsilon(t)$]
Exp	1	10.707	0.023	-0.0002	0.004
	2	11.056	-0.019	0.0005	0.024
	3	12.678	-0.075	0.0008	0.013
Rec	1	10.920	0.009	-0.0009	0.007
	2	10.550	0.046	-0.001	0.018
	3	8.426	0.117	-0.001	0.013
	4	11.615	-0.016	0.0001	0.020

The (simulated) data from the estimation period is used in the estimation procedure, described below.

Each time point is classified as being either expansion or recession by the same procedure as in Koskinen and Öller (1998) where the following definition of a turning point is used:
The differentiated series has kept the same sign for at least two consecutive time points when it changes sign. If the new sign is kept during at least the next time point, then a turning point is said to have occurred.

The method for estimating $\{\beta_1, \beta_2, \sigma_1^2, \sigma_2^2\}$, described below, is one component in the estimation procedure used by Koskinen and Öller (1998):

$$\hat{\beta}_j = \frac{\bar{d}_j}{4} = \frac{1}{4 \cdot n_j} \sum_{i=1}^{nj} d_{ji},$$

$$\hat{\sigma}_j^2 = \frac{1}{2 \cdot (n_j - 1)} \sum_{i=1}^{n_1} (d_{ji} - \bar{d}_j)^2,$$

where n_j= # time point classified as state j and d_j is the differentiated series classified as belonging to state j, where j = {expansion} or {recession}.

Research Report

2001:1	Holgersson, H.E.T.:	On assessing multivariate normality.
2001:2	Sonesson, C. & Bock, D.:	Statistical issues in public health monitoring – A review and discussion.
2001:3	Andersson, E.:	Turning point detection using non-parametric statistical surveillance. Evaluation of some influential factors.
2001:4	Andersson, E. & Bock, D.:	On seasonal filters and monotonicity.
2001:5	Andersson, E., Bock, D. & Frisén, M.:	Likelihood based methods for detection of turning points in business cycles. A comparative study.
2001:6	Sonesson, C.:	Evaluations of some exponentially weighted moving average methods.
2001:7	Sonesson, C.:	Statistical surveillance. Exponentially weighted moving average methods and public health monitoring.
2002:1	Frisén, M. & Sonesson, C.:	Optimal surveillance based on exponentially weighted moving averages.
2002:2	Frisén, M.:	Statistical surveillance. Optimality and methods.
2002:3	Jonsson, R. & Persson, A.:	Bayes prediction of binary outcomes based on correlated discrete predictors.
2002:4	Persson, A.:	Prediction of work resumption among men and women with lower back- and neck pain in a Swedish population.
2002:5	Persson, A.:	Prediction of work resumption in theory and practice.
2002:6	Andersson, E., Bock, D. & Frisén, M.:	Detection of turning points in business cycles.