Biometrika (2001), **88**, 4, *pp.* 1186–1192

Printed in Great Britain

A simple generalised crossvalidation method of span selection for periodogram smoothing

By HERNANDO C. OMBAO

Department of Statistics, University of Pittsburgh, Pittsburgh, Pennsylvania 15260-0001, U.S.A.

ombao@stat.pitt.edu

JONATHAN A. RAZ*

ROBERT L. STRAWDERMAN

Biometrics Department, Cornell University, Ithaca, New York 14853-7801, U.S.A.

rls54@cornell.edu

and RAINER VON SACHS

Institut de Statistique, Université Catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium

rvs@stat.ucl.ac.be

Summary

A consistent estimator for the spectral density of a stationary random process can be obtained by smoothing the periodograms across frequency. An important component of smoothing is the choice of the span. Lee (1997) proposed a span selector that was erroneously claimed to be unbiased for the mean squared error. The naive use of mean squared error has some important drawbacks in this context because the variance of the periodogram depends on its mean, i.e. the spectrum. We propose a new span selector based on the generalised crossvalidation function derived from the gamma deviance. This criterion, originally developed for use in fitting generalised additive models, utilises the approximate full likelihood of periodograms, which asymptotically behave like independently distributed chi-squared, i.e. gamma, random variables. The proposed span selector is very simple and easily implemented. Simulation results suggest that the proposed span selector generally outperforms those obtained under a mean squared error criterion.

Some key words: Bandwidth selection; Deviance; Generalised additive model; Generalised crossvalidation; Smoothed periodograms; Spectrum estimation; Stationary time series.

1. Introduction

Let X_t $(t = 0, \pm 1, \pm 2 \ldots)$ be a zero stationary time series with covariance function $\gamma(\tau) = E(X_t X_{t+\tau})$ $(\tau = 0, \pm 1, \pm 2, \ldots)$. The second-order properties of $\{X_t\}$ are equivalently described by the spectrum

$$f(\lambda) = \sum_{\tau=-\infty}^{\infty} \gamma(\tau) \exp(-i2\pi\lambda\tau), \quad \lambda \in [0, 1]. \tag{1}$$

Since $f(\lambda)$ is symmetric about $\lambda = \frac{1}{2}$, it suffices to estimate $f(\lambda)$ only for $\lambda \in [0, \frac{1}{2}]$.

* Jonathan Raz died on 23 February 2000. His work on this paper was done at the Department of Biostatistics, University of Pennsylvania, Philadelphia, Pennsylvania, U.S.A.

Various nonparametric methods for estimating the spectrum use periodograms, or sample analogues of the spectrum. The periodogram at frequency $\lambda_k = k/T$ computed from a time series $\{X_0, \ldots, X_{T-1}\}$ is

$$I_k = T^{-1} \left| \sum_{t=0}^{T-1} X_t \exp(i2\pi kt/T) \right|^2 \quad (k = 0, \ldots, T-1). \tag{2}$$

The periodogram is known to be an asymptotically unbiased estimator of the spectrum. However, the I_k's are themselves inconsistent estimators as they exhibit erratic behaviour even when the underlying spectrum of the process is smooth. The traditional method for obtaining consistent estimators of the spectrum is to smooth the periodograms over frequency using a kernel smoother (Brillinger, 1981, p. 149). Let f_k be the value of the spectrum frequency $\lambda_k = k/T$. A smoothed periodogram estimator of f_k, denoted by $\hat{f}_k$, is then given by

$$\hat{f}_{k,p} = \sum_{j=-p}^{j=p} W_{p,j} I_{k+j} \quad (k = 0, 1, \ldots, T/2), \tag{3}$$

where $2p + 1$ is the smoothing span and $W_{p,j}$ are nonnegative weights that satisfy the following conditions for any fixed p:

$$W_{p,j} = W_{p,-j} \quad (j = 1, \ldots, p), \quad \sum_{j=-p}^{j=p} W_{p,j} = 1. \tag{4}$$

For $\hat{f}_{k,p}$ to be consistent, $p = p_T$ must go to infinity and $p_T/T \to 0$ as $T \to \infty$. Generally, the weights are chosen so that $W_{p,0}$ is a decreasing function of p. Various weights satisfying (4) exist. However, the form of the weights $W_{p,j}$ is generally only of secondary importance to the value of the span or bandwidth (Priestley, 1981, pp. 449, 463).

Recently, Lee (1997) proposed an automatic span selector, a method for selecting p, that minimises a purportedly unbiased estimator of the mean squared error between the smooth periodogram and the true spectrum. The use of mean squared error does not acknowledge the heteroscedasticity of the periodograms $I_0, \ldots, I_{T-1}$, which leads one to question the appropriateness of the criterion itself. Moreover, Lee derives his estimate of risk by assuming that all the periodograms $I_0, \ldots, I_{T-1}$ are independent. This assumption is incorrect; see for example Brillinger (1981, Theorem 5.2.6).

In this paper, we propose a simple alternative method for periodogram smoothing. The proposed method of span selection is based on a procedure originally developed for fitting generalised additive models (Hastie & Tibshirani, 1990). Our method employs the deviance function, a measure of distance between the smoothed spectral estimate and the raw periodogram, appropriate for the asymptotic distribution of the periodogram. The proposed method has a number of attractive properties, and provides a computationally simpler alternative to the method of Lee (1997). It has also been successfully implemented as part of an automatic method for the analysis of nonstationary time series (Ombao et al., 2001).

In § 2, we briefly review the method of Lee (1997), discuss its deficiencies and propose a bias-corrected version. In § 3, we introduce our method, and, in § 4, we use simulations to compare it to both the original method of Lee (1997) and its bias-corrected version.

2. Periodogram asymptotics and the method of Lee (1997)

Let $X_0, \ldots, X_{T-1}$ be a zero-mean stationary time series of length T, where T is even. According to Theorem 5.2.6 in Brillinger (1981), if (i) all the moments of X_t exist, (ii) $\sum |\gamma(\tau)| < \infty$, and (iii) T is large, then $I_j/f_j \sim \varepsilon_j$ $(j = 0, \ldots, T/2)$, where ε_0 and $\varepsilon_{T/2}$ are approximately distributed as χ_1^2 random variables and $\varepsilon_1 \ldots \varepsilon_{T/2-1}$ are approximately distributed as $\chi_2^2/2$ random variables. Furthermore, $\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_{T/2}$ are approximately independent. Note that $E(I_j) = f_j$ for $j = 0, \ldots, T/2$, $\text{var}(I_j) = 2f_j^2$ when $j = 0$ or $T/2$ and $\text{var}(I_j) = f_j^2$ when $j = 1, \ldots, T/2 - 1$.

Lee (1997)'s method for span selection is to choose the smoothing parameter p that minimises

$$\hat{R}(p) = \frac{1}{T}\left\{\textsc{rss}(p) - (1 - 2W_{p,0}) \sum_{j=0}^{T-1} I_j^2/2\right\}, \tag{5}$$

where $\textsc{rss}(p) = \sum_{j=0}^{T-1} (I_j - \hat{f}_{j,p})^2$. This criterion function is derived as an unbiased estimator of the $\mathscr{L}_2$-risk function $R_2(p) = (1/T)\{\sum_{j=0}^{T-1} E(f_j - \hat{f}_{j,p})^2\}$, assuming that $\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_{T-1}$ are each exactly distributed as independent $\chi_2^2/2$ random variables. In Lee (1997), the distributions of ε_0 and $\varepsilon_{T/2}$ are erroneously assumed to be the same as that of the other ε_j's. This assumption was based on the work of Pawitan & O'Sullivan (1994) who claim that the difference between using the distribution χ_1^2 and $\chi_2^2/2$ at frequencies $\lambda_k = 0$ and $\lambda_k = \frac{1}{2}$ is asymptotically negligible.

The claim that $\hat{R}(p)$ is an unbiased estimator of $R_2(p)$ relies on an incorrect assumption that the periodograms $I_0, I_1, \ldots, I_{T-1}$ are independent and ignores the fact that the periodograms are periodic and symmetric at $\lambda_k = 0$ and $\lambda_k = \frac{1}{2}$. In fact, $I_{-1} = I_1 = I_{T-1}$; $I_{-2} = I_2 = I_{T-2}, \ldots$, and thus one must be especially careful when smoothing at the edges if the smoothing window has a span that is large enough to cover identical periodograms ordinates. In an unpublished University of Pittsburgh technical report, the authors establish that $\hat{R}(p)$ is biased for $R_2(p)$, and prove that $E\{\hat{R}(p)\} = R_2(p) + B_T(p)$, where

$$B_T(p) = \left\{-4 \sum_{l=1}^{[p/2]} W_{p,2l}(f_l^2 + f_{T/2-l}^2) + (1 - 2W_{p,0})(f_0^2 + f_{T/2}^2)\right\} \Big/ T \tag{6}$$

and where $[.]$ denotes the greatest integer function. The requisite correction yields a criterion function considerably more complicated than (5).

A bias-corrected Lee span selector is

$$\hat{R}_{\mathrm{BC}}(p) = \hat{R}(p) + \frac{2}{T} \sum_{l=1}^{[p/2]} W_{p,2l}(I_l^2 + I_{T/2-l}^2) - \left(\frac{1 - 2W_{p,0}}{6T}\right)(I_0^2 + I_{T/2}^2). \tag{7}$$

This follows from $E(I_0^2/3) = f_0^2$, $E(I_{T/2}^2/3) = f_{T/2}^2$ and $E(I_j^2/2) = f_j^2$ for all $j = 1, \ldots, T/2 - 1$. We note, however, that under suitable conditions the bias $B_T(p)$ is asymptotically negligible. Since $R_2(p) = O(1)$, the Lee and the bias-corrected Lee methods are expected to select similar spans for large T.

We close this section by pointing out a potentially more significant deficiency of the use of mean squared error as a criterion for selecting the smoothing parameter in this context. Under Lee's assumptions, $E(I_j^2) = 2f_j^2 = 2\,\mathrm{var}(I_j)$ for $j = 0, \ldots, T-1$. Hence, up to a constant shift independent of p, (5) and (7) can be viewed as approximations to the generalised crossvalidated deviance function in a generalised additive model with normally distributed error terms; see Hastie & Tibshirani (1990, §§ 3.4, 6.9). The use of this criterion therefore corresponds to selecting the smoothing parameter in a homoscedastic additive model with normal errors. Since the errors are multiplicative and the variance of I_j depends on the mean f_j, this casts significant doubt on the appropriateness of the mean squared error as a fitting criterion in this context, particularly with regard to efficiency. More generally, the problem of selecting the smoothing parameters can be viewed as fitting a generalised additive model in which the explanatory variables are given by the frequency. This suggests that a more appropriate generalised crossvalidation criterion may be derived from a suitably constructed deviance function. We explore this idea further in § 3.

3. The gamma generalised crossvalidation method

We propose a simple method for selecting the smoothing parameter using a procedure developed for generalised additive models. In line with Brillinger (1981, Theorem 5.2.6), consider the model

$$I_j \sim \begin{cases} \mathrm{Ga}(1, f_j) & (j = 1, \ldots, T/2 - 1), \\ \mathrm{Ga}(\frac{1}{2}, 2f_j) & (j = 0, T/2), \end{cases}$$

where $I_0, \ldots, I_{T/2}$ are independent. The connection to the previous section is easily seen upon recalling that a χ^2_ν random variable is equivalent to a $\mathrm{Ga}(\nu/2, 2)$ random variable for all $\nu \geqslant 1$. We employ gamma random variables in our discussion because it is more natural in the context of generalised additive models.

Under the present model, $E(I_j) = f_j$ for $j = 0, \ldots, T/2$, and consequently the problem of spectral estimation can be considered in terms of a generalised additive model with a gamma distribution, the identity link and one explanatory variable, namely frequency. Hastie & Tibshirani (1990, § 6.9) discuss a number of different methods for the automatic selection of a smoothing parameter in a generalised additive model, including the crossvalidated deviance, the generalised crossvalidated deviance and the AIC statistic. All of these span selection methods employ a general form of the deviance rather than the usual sum of squared residuals. Presumably, selecting a penalised criterion function based on an appropriate loglikelihood function allows better simultaneous control of goodness of fit and smoothness; see O'Sullivan et al. (1986) and Gu (1990) for related discussions in the context of semiparametric generalised linear models.

We propose to select p by minimising the generalised crossvalidated deviance function

$$\mathrm{GCV}(p) = \frac{M^{-1}\sum_{j=0}^{M-1} D(I_j, \hat{f}_{j,p})}{\{1 - \mathrm{tr}(H_p)/M\}^2}, \tag{8}$$

where $M = T/2 + 1$, $D(I_j, \hat{f}_{j,p})$ is the deviance, and H_p is the smoother matrix with smoothing parameter p. Since the full asymptotic likelihood of the I_j's corresponds to that of a set of independent gamma random variables with means governed by the true spectral values f_j, for $j = 0, \ldots, T/2$, an appropriate choice of deviance function is

$$D(I_j, \hat{f}_{j,p}) = q_j\{-\log(I_j/\hat{f}_{j,p}) + (I_j - \hat{f}_{j,p})/\hat{f}_{j,p}\}$$

(McCullagh & Nelder, 1989, pp. 33–4). Here $q_j = 1 - 0{\cdot}5\mathscr{I}(j = 0 \text{ or } j = M - 1)$ where $\mathscr{I}$ is the indicator function. This accounts for the fact that I_0 and $I_{T/2}$ have distributions that differ from the remaining I_k. The denominator in (8), often referred to as the model degrees of freedom, can be expressed in terms of the weight at the centre of the smoothing window: $\{1 - \mathrm{tr}(H_p)/M\}^2 = (1 - W_{p,0})^2$. Substituting these various pieces into (8) results in our proposed criterion function

$$\mathrm{GCV}(p) = M^{-1} \sum_{j=0}^{M-1} q_j \left\{\frac{-\log(I_j/\hat{f}_{j,p}) + (I_j - \hat{f}_{j,p})/\hat{f}_{j,p}}{(1 - W_{p,0})^2}\right\}. \tag{9}$$

The gamma generalised crossvalidation method is both conceptually simple and easily implemented. It is based on the full asymptotic likelihood because it uses a deviance function appropriate for the asymptotic distribution of the periodograms. Hence, it is based on a deviance function that is arguably more relevant than that considered in Lee (1997). Moreover, as described in Hastie & Tibshirani (1990, § 6.8.3), it has an appealing interpretation since $\mathrm{GCV}(p)$, evaluated at its minimiser $\hat{p}_{\mathrm{GCV}}$, may be viewed as an approximately unbiased estimator of the prediction error $E\{\sum_j D(I_j^o, \hat{f}_{j,\hat{p}})/M\}$. Here, I_j^o has the same distribution as, but is independent of, I_j for $j = 0, \ldots, T/2$. Finally, the bias correction inherent in (8) is computationally straightforward and simpler than that employed in either (5) or (7).

4. Numerical experiments

4·1. *Design of the study*

A simulation study was conducted to compare the performance of the Lee (1997) method, the bias-corrected Lee method and the gamma generalised crossvalidation method using six different ARMA (α, β) processes. We define a random process $\{X_t\}$ to be ARMA (α, β) if it is generated by the equation

$$X_t + a_1 X_{t-1} + \ldots + a_\alpha X_{t-\alpha} = \xi_t + b_1 \xi_{t-1} + \ldots + b_\beta \xi_{t-\beta},$$

where ξ_t are independent $N(0, 1)$ random variables. In Example 1 we used ARMA (3, 0) with

$a_1 = -1{\cdot}5$, $a_2 = 0{\cdot}7$ and $a_3 = -0{\cdot}1$, in Example 2 we used ARMA$(0, 4)$ with $b_1 = -0{\cdot}3$, $b_2 = -0{\cdot}6$, $b_3 = -0{\cdot}3$ and $b_4 = 0{\cdot}6$, in Example 3 we used ARMA$(3, 0)$ with $a_1 = 0{\cdot}9$, $a_2 = 0{\cdot}8$ and $a_3 = 0{\cdot}6$, in Example 4 we used ARMA$(0, 3)$ with $b_1 = 0{\cdot}9$, $b_2 = 0{\cdot}8$ and $b_3 = 0{\cdot}6$, in Example 5 we used ARMA$(1, 0)$ with $a_1 = 0{\cdot}05$, and in Example 6 we used ARMA$(0, 0)$, which is just the white noise process. Examples 1–4 were presented in Lee (1997). For convenience in implementation, we used in each case a boxcar smoother with weights defined by $W_{p,j} = 1/(2p+1)$ for all $j = -p, \ldots, p$. Our simulation study was done using Matlab. In each ARMA example, 5000 independent time series datasets, each of length $T = 1024$, were generated.

4·2. *The criteria*

For the nth dataset generated in each ARMA example ($n = 1, \ldots, 5000$), we obtained the spans selected by the three methods and computed the corresponding smoothed spectral estimates. We then computed the three average losses, averaged over 5000 generated datasets under each of four criteria that represent different measures of closeness of the spectral estimates to the true spectrum.

The first criterion is the squared error, or $\mathscr{L}_2$ loss, defined to be

$$\tilde{R}_2(\hat{p}) = T^{-1} \sum_{k=0}^{T-1} (f_k - \hat{f}_{k,\hat{p}})^2.$$

The second criterion is the Kullback–Leibler loss, which we take to be

$$\tilde{G}(\hat{p}) = M^{-1} \sum_{k=0}^{M-1} q_k \{ f_k/\hat{f}_{k,\hat{p}} - \log(f_k/\hat{f}_{k,\hat{p}}) - 1 \},$$

with $q_j = 1 - 0{\cdot}5\mathscr{I}(j = 0 \text{ or } M-1)$. This measure is derived directly from the Kullback–Leibler distance between two gamma densities having a common shape parameter and different scale parameters. We also considered the absolute error, or $\mathscr{L}_1$ loss, defined by $\tilde{R}_1(\hat{p}) = T^{-1}\sum_{k=0}^{T-1} |f_k - \hat{f}_{k,\hat{p}}|$, and the supremum error, $\mathscr{L}_\infty$ loss, defined by $\tilde{R}_\infty(\hat{p}) = \sup_k |f_k - \hat{f}_{k,\hat{p}}|$. The averages and standard errors of the various measures over the 5000 simulations are reported in Table 1.

4·3. *Results*

Under the squared error criterion, the gamma generalised crossvalidation method had smaller average loss and standard error than the Lee span selectors. This was somewhat surprising because the Lee methods were specifically developed for $\mathscr{L}_2$. Moreover, correcting the bias in the Lee method did not measurably improve its performance. In fact, in a second small simulation study, not shown, we noted that even adjusting the Lee method for the theoretical bias, as opposed to the estimated bias in the bias corrected Lee method, continued to offer very little improvement. The observed lack of improvement results from the fact that the magnitude of Lee's criterion function substantially dominates that of the bias correction term. This is not entirely unexpected given the large sample size in this simulation study. A secondary factor contributing to the observed lack of improvement is the fact that p runs through integral values, implying that small perturbations in the criterion function are unlikely to produce significant changes in the span selected. Importantly, these results do not explain why Lee's method fails to dominate the gamma generalised crossvalidation method under $\mathscr{L}_2$; they only help to explain why bias correction fails to improve significantly the performance of Lee's method.

Under the Kullback–Leibler criterion, the gamma generalised crossvalidation method substantially outperformed the mean squared error methods in every example. The result was expected given the risk criterion used to select the smoothing parameter.

Under $\mathscr{L}_1$, the gamma crossvalidation method had smaller average loss and smaller standard error than the Lee methods in all examples. The same was true for $\mathscr{L}_\infty$ except for Example 2, where the mean squared error methods had smaller average loss. However, the standard errors of the gamma generalised crossvalidation method were substantially smaller than those obtained using mean squared error.

Table 1. *Average losses from* 5000 *simulated time series for the Lee* (1997), *bias-corrected Lee* (BC Lee) *and gamma generalised crossvalidation*, Gamma GCV, *methods computed from* 5000 *simulated time series datasets, with standard errors of the mean in parentheses*

Example	Lee	BC Lee	Gamma GCV	Lee	BC Lee	Gamma GCV
		(a) $\mathscr{L}_2$ loss			(b) KL loss	
1	37·286	37·049	22·815	433·0	435·0	285·0
	(0·644)	(0·629)	(0·333)	(7·6)	(7·6)	(3·4)
2	0·182	0·182	0·182	771·0	772·0	422·0
	(0·006)	(0·006)	(0·004)	(18·0)	(18·0)	(7·0)
3	0·733	0·728	0·539	461·0	460·0	331·0
	(0·009)	(0·009)	(0·005)	(6·7)	(6·6)	(2·7)
4	0·651	0·651	0·557	787·0	796·0	355·0
	(0·013)	(0·013)	(0·007)	(11·0)	(11·0)	(3·0)
5	0·010	0·010	0·008	104·0	103·0	81·0
	$(2{\cdot}0 \times 10^{-4})$	$(2{\cdot}0 \times 10^{-4})$	$(1{\cdot}2 \times 10^{-4})$	(3·0)	(3·0)	(1·7)
6	$100{\cdot}0 \times 10^{-4}$	$100{\cdot}0 \times 10^{-4}$	$79{\cdot}0 \times 10^{-4}$	102·0	102·0	80·0
	$(1{\cdot}7 \times 10^{-4})$	$(1{\cdot}7 \times 10^{-4})$	$(1{\cdot}0 \times 10^{-4})$	(2·5)	(2·5)	(1·5)
		(c) $\mathscr{L}_1$ loss			(d) $\mathscr{L}_\infty$ loss	
1	1871·2	1872·6	1595·4	3282·9	3255·7	2247·2
	(25·0)	(25·0)	(20·8)	(36·8)	(35·5)	(16·8)
2	265·8	265·8	262·3	135·7	156·6	148·8
	(2·6)	(2·6)	(2·1)	(3·1)	(3·1)	(1·8)
3	420·4	411·9	380·6	398·5	395·9	310·6
	(1·9)	(1·8)	(1·3)	(3·6)	(3·5)	(1·6)
4	383·0	383·5	357·3	291·3	290·4	298·0
	(2·6)	(2·6)	(1·8)	(3·9)	(3·8)	(2·1)
5	74·7	74·6	67·8	21·5	21·4	18·2
	(0·6)	(0·6)	(0·4)	(0·2)	(0·2)	(0·1)
6	74·1	74·1	67·8	21·1	21·1	18·2
	(0·5)	(0·5)	(0·4)	(0·2)	(0·2)	(0·1)

(a) $\mathscr{L}_2$ loss; (b) Kullback–Leibler loss, values in units of 10^{-4}; (c) $\mathscr{L}_1$ loss, values in units of 10^{-3}; (d) $\mathscr{L}_\infty$ loss, values in units of 10^{-2}.

The apparent failure of either $\mathscr{L}_2$-derived method to dominate under the $\mathscr{L}_2$ loss criterion raises serious questions about the validity of using squared error loss as a basis for developing span selectors when the random observations are non-Gaussian and have heteroscedastic mean-dependent variance. The gamma generalised crossvalidation method proposed here employs a criterion function that is tailored to the asymptotic behaviour of the periodogram estimates. Our simulation results show that this method generally gives more stable and less variable spectral estimates than those based on squared error loss.

Acknowledgement

The National Institute of Mental Health supported the research of H. C. Ombao, J. A. Raz and R. von Sachs, the last of whom was supported also by the Belgian government. The authors are thankful to the editor and referee for helpful comments.

References

Brillinger, D. (1981). *Time Series: Data Analysis and Theory*. Oakland, CA: Holden-Day.

Gu, C. (1990). Adaptive spline smoothing in non-Gaussian regression models. *J. Am. Statist. Assoc.* **85**, 801–7.
Hastie, T. & Tibshirani, R. (1990). *Generalized Additive Models.* London: Chapman and Hall.
Lee, T. C. M. (1997). A simple span selector for periodogram smoothing. *Biometrika* **84**, 965–9.
McCullagh, P. & Nelder, J. (1989). *Generalized Linear Models*, 2nd ed. London: Chapman and Hall.
Ombao, H., Raz, J., von Sachs, R. & Mallow, M. (2001). Automatic statistical analysis of bivariate non-stationary time series. *J. Am. Statist. Assoc.* **96**, 543–60.
O'Sullivan, F., Yandell, B. & Raynor, W. (1986). Automatic smoothing of regression functions in generalized linear models. *J. Am. Statist. Assoc.* **81**, 96–103.
Pawitan, Y. & O'Sullivan, F. (1994). Nonparametric spectral density estimation using penalized Whittle likelihood. *J. Am. Statist. Assoc.* **89**, 600–10.
Priestley, M. (1981). *Spectral Analysis and Time Series.* London: Academic Press.

[*Received November* 1999. *Revised April* 2001]