Structural Time Series Models and Trend Detection in Global and Regional Temperature Series

XIAOGU ZHENG AND REID E. BASHER

National Institute of Water and Atmospheric Research, Wellington, New Zealand

(Manuscript received 20 January 1998, in final form 20 October 1998)

ABSTRACT

A unified statistical approach to identify suitable structural time series models for annual mean temperature is proposed. This includes a generalized model that can represent all the commonly used structural time series models for trend detection, the use of differenced series (successive year-to-year differences), and explicit methods for comparing the validity of no-trend nonstationary residuals models relative to trend models. Its application to Intergovernmental Panel on Climate Change global and latitude-belt temperature series reveals that a linear trend model (starting in 1890, with Southern Oscillation index signal removal and a red noise residuals process) is the optimal model for much of the globe, from the Northern Hemisphere Tropics to the Southern Hemisphere midlatitudes, but that a random stationary increment process (with no deterministic trend) is preferred for the northern part of the Northern Hemisphere. The result for the higher northern latitudes appears to be related to the greater climate variability there and does not exclude the possibility of a trend being present. The hemispheric and global series will contain a mixture of the two processes but are dominated by and best represented by the linear trend model. The latitudinal detectability of trends is oppositely matched to where GCMs indicate greatest anthropogenic trend, that is, it is best for the Tropics rather than for the high latitudes. The results reinforce the view that the global temperatures are affected by a long-term trend that is not of natural origin.

1. Introduction

The central task of climate change detection studies is to determine whether an observed change or trend is ‘‘significant,’’ that is, highly unusual relative to the rich background of natural variability, and unlikely to have occurred by chance alone (Santer et al. 1996). This requires a statistical model that appropriately represents the variability, both natural and anthropogenic, and that allows the testing of significance of the model for the data series being examined. A number of structural time series models (e.g., linear trend plus red noise) have been proposed (Wigley and Jones 1981; Jones 1989; Gordon 1991; Bloomfield 1992; Bloomfield and Nychka 1992; Galbraith and Green 1992; Woodward and Gray 1995; Visser and Molenaar 1995; Zheng et al. 1997), but unfortunately the different models give different conclusions, not just in the value and statistical significance of a trend, but also in the validity of the trend model itself.

The structural time series model usually used to model global-mean temperature comprises a deterministic trend plus random residuals about the trend, where the residuals are assumed to represent the natural variability, including the autocorrelation structure of the series, and the trend is the putative response of temperature to anthropogenic forcings (Bloomfield 1992; Zheng et al. 1997). Such models by themselves cannot distinguish between sources of trend; this problem of attribution, of linking trends to causes, is an additional issue (Santer et. al. 1996). Only the problem of detection is considered in the present paper.

Global-mean temperature has also been modeled as a structural time series having nonstationary residuals but no deterministic trend component (Gordon 1991; Woodward and Gray 1995; Gordon et al. 1996). If such models were indeed correct, then the fitting of a model comprising a trend and stationary residuals could result in the erroneous detection of a trend. These authors are cautious in their conclusions, none claiming that global warming does not exist, but they argue that statistically we cannot rule out the possibility that the past century’s rise in global temperatures is simply a result of the natural variability and thus may not continue in the future. Their results present a disconcerting challenge to the conventional view that global temperatures are rising because of some external forcing, probably associated with anthropogenic factors, and are likely to continue to rise.

In the work reported here, the main currently used

Corresponding author address: Dr. Xiaogu Zheng, National Institute of Water and Atmosphere Research, P.O. Box 14-901, Kilbirnie, Wellington, New Zealand.
E-mail: x.zheng@niwa.cri.nz

structural time series models for detecting anthropogenic forcing in annual mean temperatures are reviewed and are developed into a proposed unified statistical model. Two key features are 1) the modeling of residuals as stationary increment processes, which covers all the stationary and nonstationary residual models commonly used in trend analysis, and 2) the use of differenced series, which results in stationary residuals and hence allows the use of the procedures for parameter estimation and model choice described by Zheng et al. (1997). Differencing also partly deals with the technically difficult question (Bloomfield 1992) of how to separate a linear trend from low-frequency natural variability, since the latter is largely removed in the differenced series and the linear trend term transforms to a constant.

By applying these methods to the various global, hemispheric, land, and ocean Intergovernmental Panel on Climate Change (IPCC) temperature series (Nicholls et al. 1996; Jones 1996), and to similarly constructed temperature series for latitude belts, we have been able to clarify the circumstances under which the deterministic trend model and the nontrend model are each most appropriate. We find that the model choice varies geographically, but that this occurs in a systematic way and is consistent with the hypothesis that rising temperatures represent a trend that is associated with external forcing and is likely to continue. In particular, we find that the deterministic trend model is optimal over much of the globe, but becomes nonoptimal in the higher Northern Hemisphere latitudes, possibly as a result of higher natural variability there.

2. Statistical models

The proposed statistical model for trend detection in an annual mean temperature series $\{y_t\}$ of T years is

$$y_t = \mu + f_t(\lambda_1, \ldots, \lambda_r) + \beta_1 x_{1,t} + \cdots + \beta_k x_{k.t} + \varepsilon_t, \quad t = 1, \ldots, T, \tag{1}$$

where μ is the temperature normal over T years; $f_t(\lambda_1, \ldots, \lambda_r)$ is a function of t with parameters $\{\lambda_1, \ldots, \lambda_r\}$; r is the number of parameters; $\{x_{i,t}, t = 1, \ldots, T\}$ is the ith explanatory variable (such as the Southern Oscillation index), which is observable and noncolinear with f_t; k is the number of explanatory variables; β is the coefficient for the ith explanatory variables; and $\{\varepsilon_t, t = 1, \ldots, T\}$ is the residual time series, which is an autoregressive integrated moving average (ARIMA) $(p, 1, q)$ process (see next paragraph). The function $f_t(\lambda_1, \ldots, \lambda_r)$ represents the deterministic trend to be detected, while the functions $\beta_1 x_{1,t}, + \cdots + \beta_k x_{k.z} + \varepsilon_t$ represent the components of natural variability that can be explicitly modeled.

A time series $\{\xi_t\}$ is called an autoregressive integrated moving-average process of order (p, d, q) [ARIMA(p, d, q), Brockwell and Davis 1996]; if

$$\nabla^d \xi_t - \phi_1 \nabla^d \xi_{t-1} - \cdots - \phi_p \nabla^d \xi_{t-p} = \eta_t + \theta_1 \eta_{t-1} + \cdots + \theta_q \eta_{t-q}, \tag{2}$$

where p, d, and q are nonnegative integers; $\{\phi_1, \ldots, \phi_p\}$ are the autoregressive coefficients; $\{\theta_1, \ldots, \theta_q\}$ are the moving average coefficients; $\{\eta_t\}$ is a white noise process with variance σ^2 (the innovation variance); and ∇ is the difference operator, that is,

$$\nabla y_t = y_t - y_{t-1}, \qquad \nabla^2 y_t = \nabla(\nabla y_t), \tag{3}$$

etc. In the case where $d = 0$ (∇^0 represents the identity operator), $\{\xi_t\}$ is an autoregressive moving-average process (ARMA) of order (p, q); [ARMA(p, q), Brockwell and Davis 1996].

A differenced series of the annual mean temperature can be constructed by applying ∇ to the both sides of model (1):

$$\nabla y_t = \nabla f_t(\lambda_1, \ldots, \lambda_r) + \beta_1 \nabla x_{1,t} + \cdots + \beta_k \nabla x_{k.t} + \nabla \varepsilon_t, \quad t = 2, \ldots, T, \tag{4}$$

where the residuals $\{\nabla \varepsilon_t\}$ are necessarily an ARMA(p, q) process, because it is assumed that $\{\varepsilon_t\}$ is an ARIMA$(p, 1, q)$ process. This equation (4) for the differenced series is therefore in the same format as that presented by Zheng et al. (1997), and its trend $\nabla f_t(\lambda_1, \ldots, \lambda_r)$ may be estimated by the approach described by Zheng et al. (1997). The trend in the original series, $f_t(\lambda_1, \ldots, \lambda_r)$, is then determined up to a constant, which can be neglected as only the derivatives (slopes) are of interest in trend analysis. A critical advantage of differencing is that it enables the direct quantitative testing of the relative quality of competing *trend + ARMA(p, q)* and *no-trend + ARIMA(p, 1, q)* models. Previously, this was not possible (Zheng et al. 1997). Further discussion of this issue is given in section 3d.

More detailed explanations for each term of model (1) are given below.

a. Trend term

The three simplest options for the trend function are as follows.

1) Linear trend

$$f_t(\lambda_1) = \lambda_1 t \tag{5}$$

is the most commonly used form for detecting enhanced greenhouse gas forcing (Wigley and Jones 1981). The corresponding difference function is

$$\nabla f_t(\lambda_1) = \lambda_1. \tag{6}$$

2) Bilinear trend

$$f_t(\lambda_1, \lambda_2) = \begin{cases} \lambda_1 t & \text{when } t < t_0, \\ \lambda_1 t_0 + \lambda_2(t - t_0) & \text{when } t \geq t_0, \end{cases} \tag{7}$$

where t_0 is a turning point, may be used to test for a significant change in trend at some point in time.

Examples include testing whether the rate of global temperature rise increased around the turn of the century (Santer et al. 1996), or identifying the year from which the decline in diurnal temperature range becomes significant (Zheng et al. 1997). The corresponding difference function is

$$\nabla f_t(\lambda_1, \lambda_2) = \lambda_1, \quad \text{when } t \le t_0,$$
$$= \lambda_2, \quad \text{when } t > t_0. \tag{8}$$

3) Quadratic trend

$$f_t(\lambda_1, \lambda_2) = \lambda_1 t + \lambda_2 t^2 \tag{9}$$

can be introduced to test for a nonlinear response (Woodward and Gray 1995). Nonlinear types of increases are present both in projected greenhouse gas concentrations (Schimel et al. 1996) and in GCM-modeled global temperature (Kattenberg et al. 1996), and in future decades accelerating increases of temperature may become more evident. The difference function for the quadratic trend is

$$\nabla f_t(\lambda_1, \lambda_2) = (\lambda_1 - \lambda_2) + 2\lambda_2 t. \tag{10}$$

An exponential trend function is also a possible nonlinear form (Zheng et al. 1997). However, the quadratic and other linear models can be expressed as a linear combination of a group of determined functions of t, that is,

$$f_t(\lambda_1, \ldots, \lambda_r) = \lambda_1 f_{1,t} + \cdots + \lambda_r f_{r,t}, \tag{11}$$

which gives them a computational advantage over the exponential trend in respect to estimating trends and confidence intervals.

b. Explanatory variables

Some part of the natural variability in a temperature series can be described by explanatory variables that are measurable and have a clear physical sense. The Southern Oscillation index (SOI) is known to have a strong association with the global temperature series (Jones 1989) and with some regional temperature series (Zheng et al. 1997). Other variables, such as sunspot counts and Lamb's volcanic dust veil index, may be physically plausible sources of natural variability, and their potential roles can be tested by including them in model (1) as explanatory variables.

All such explanatory variable series must be subjected to high-pass filtering (with a filter cutoff of about 10 yr) prior to doing the regression model development and analysis, to ensure the estimated trend is uniquely represented by f_t [see Zheng et al. (1997) for more detailed discussion]. Explanatory variables therefore play no role in representing any natural sources of trend; their role is to remove known short-term variability that otherwise would contaminate the residuals.

c. Residuals

Residuals characterize the natural variability not explained by the explanatory variables. Galbraith and Green (1992), Richards (1993), and Woodward and Gray (1995) introduced the ARIMA(p, 1, q) form for modeling nonstationary residual time series $\{\varepsilon_t, t = 1, \ldots, T\}$. The random walk residuals model introduced by Gordon (1991) is a special form of this, where $p = q = 0$.

If the ARIMA(p, 1, q) moving-average coefficients $\theta_1, \ldots, \theta_q$ satisfy the restriction

$$\theta_1 + \cdots + \theta_q = -1, \tag{12}$$

then the residual series $\{\varepsilon_t, t = 1, \ldots, T\}$ reduces to the stationary ARMA(p, $q - 1$) process, with the same autoregressive coefficients $\phi_1, \ldots, \phi_p$ but different moving-average coefficients $\{1 + \Sigma_{i=1}^{j} \theta_i, j = 1, \ldots, q - 1\}$. Equation (12) is thus a condition of stationarity. This is the residuals model introduced by Bloomfield (1992) and used by Zheng et al. (1997). The proposed model (1) is therefore more general than the model described in Zheng et al. (1997). In particular, if $q = 1$ and $\theta_1 = -1$, $\{\varepsilon_t, t = 1, \ldots, T\}$ is an order-p autoregressive process [AR(p) = ARMA(p, 0); Brockwell and Davis 1996]. Furthermore, in this situation, if $p = 0$, then $\{\varepsilon_t, t = 1, \ldots, T\}$ is a white noise process (e.g., as used by Wigley and Jones 1981); while if $p = 1$, it is a red noise process (e.g., as used by Wigley et al. 1989).

Visser and Molenaar (1995) introduced a residuals model ARIMA(p, d, q) with selected $d > 1$ for specific data series. However, they gave no rule for choosing d and hence there is no way of knowing whether the chosen d are optimal.

In our earlier paper (Zheng et al. 1997) we were unable to statistically directly compare the utility of the ARIMA(p, 1, q) model with its subset ARMA(p, q) form, but for the series considered we were able to reject the ARIMA(p, 1, q) model by means of indirect tests and simulations. This problem was also addressed by Galbraith and Green (1992) and Richards (1993) using different methods. However, their methods can only be applied to series without trend, and hence their results on annual mean temperatures are dependent on the prior detrending procedure used. In the present paper, the use of differenced series allows direct statistical comparison of the residuals models and no detrending procedure is required.

3. Fitting procedure

For each chosen structural time series model to be examined (e.g., linear trend plus residuals), the fitting procedure we propose comprises the following five steps. This procedure includes several improvements in concept and technique over the earlier set of steps described in Zheng et al. (1997).

a. Remove low-frequency variability from candidate explanatory variable series

Filter the candidate explanatory variable series, for example, by the S-PLUS routine ''filter'' (StatSci 1993), where a 13-term high-pass Gaussian filter (with a 50% response at 10 yr) is constructed by the S-PLUS routine ''demod.''

b. Select optimal trend model and optimal no-trend model

Four types of model were considered: no trend + stationary residual, trend + stationary residuals, trend + nonstationary residuals, and no trend + nonstationary residual. In each case, the optimal explanatory variables and order (p, q) were determined using the Bayesian Information Criteria (BIC), where the smallest value of BIC implies the best fit.

For an integer m $(\geq p)$, define the BIC (Brockwell and Davis 1996) for (1) as

$$\begin{aligned}\mathrm{BIC}(\nabla y_{m+1}, \ldots, \nabla y_T \mid \nabla y_1, \ldots, \nabla y_m) \\ = -2 \ln \max\{L(\nabla\varepsilon_{m+2}, \ldots, \nabla\varepsilon_T \mid \nabla\varepsilon_2, \ldots, \nabla\varepsilon_{m+1})\} \\ + (1 + r + k + p + q - s(\theta_1, \ldots, \theta_q)) \ln(T - 1),\end{aligned} \tag{13}$$

where $\{\varepsilon_t\}$ is a realization of an ARIMA$(p, 1, q)$ process; the function $s(\theta_1, \ldots, \theta_q)$ takes a value of 1 or 0 depending on whether $\theta_1, \ldots, \theta_q$ satisfy the stationarity restriction described above in (12). Here $L(\nabla\varepsilon_{m+2}, \ldots, \nabla\varepsilon_T \mid \nabla\varepsilon_2, \ldots, \nabla\varepsilon_{m+1})$ is the conditional likelihood of $\nabla\varepsilon_{m+2}, \ldots, \nabla\varepsilon_T$ given $\nabla\varepsilon_2, \ldots, \nabla\varepsilon_{m+1}$ and is a function of $\lambda_1, \ldots, \lambda_r$; $\beta_1, \ldots, \beta_k$; $\phi_1, \ldots, \phi_p$, $\theta_1, \ldots, \theta_q$ [for the detailed form of L, see StatSci (1993), Eq. (16.46)]. Its maximum is taken over all the parameters. The trend coefficients, the coefficients for the selected explanatory variables, the autoregressive coefficients, the moving-average coefficients, and the innovation variance are estimated simultaneously and are denoted as $\hat{\lambda}_1, \ldots, \hat{\lambda}_r$; $\hat{\beta}_1, \ldots, \hat{\beta}_{k'}$; $\hat{\phi}_1, \ldots, \hat{\phi}_p$; $\hat{\theta}_1, \ldots, \hat{\theta}_q$; $\hat{\sigma}^2$, respectively.

If the selected trend function is in the form of (11), then the S-PLUS routine ''arima.mle'' can be applied to estimate the BICs. The routine's inputs are the differenced temperature series, the differenced selected explanatory variables, and the order (p, q). However, the routine only works for invertible ARIMA(p, d, q) models, that is, where all the roots of the polynomial equation

$$\theta_1 z + \cdots + \theta_q z^q = -1 \tag{14}$$

lie outside of the unit cycle. If the differenced series meets the stationarity restriction (12), Eq. (14) for the residuals of the differenced series will have the unit root. Therefore the ARMA model for the residuals of differenced series will be noninvertible and the S-PLUS routine arima.mle will fail. To deal with this, we modify (12) as follows:

$$\theta_1 + \cdots + \theta_q = -0.999\,999 \tag{15}$$

and apply the invertible transformation proposed by Jones (1980) (which has been set as a default for arima.mle).

While an ARIMA$(p, 1, q)$ process with the restriction (12) is equal to the stationary ARMA$(p, q - 1)$ process, this is not mathematically true when the restriction is changed to (15), especially for the asymptotic behavior. To further examine this, we have simulated the processes for 100–200-yr series and find that the results are very similar. For example, when the trend + SOI + ARMA(1, 0) model and the trend + SOI + ARIMA(1, 1, 1) model with restriction (15) are fitted to the temperature series studied in this paper, the estimated regression coefficients and autoregressive coefficients are exactly the same and the relative errors for the estimated innovation variances are only about 0.5%. On this basis, we accept the validity of this use of ARIMA$(p, 1, q)$ residuals with restriction (15).

If the selected trend function is not in the form (11) (e.g., it is an exponential trend as described by Zheng et al. 1997), the S-PLUS routine ''nlminb'' can be used together with arima.mle for the estimation (Zheng et al. 1997).

Visser and Molenaar (1995) introduced a residuals model ARIMA(p, d, q) with selected $d > 1$ for specific data series, but it is difficult to accommodate such a model using the methods proposed in the present paper, because the second-order differencing required [see (3)] would eliminate any linear trend present in such series. It is therefore desirable to check whether $d > 1$ is preferred or not. If the derived optimal model has $d = 0$, that is, $\hat{\theta}_1, \ldots, \hat{\theta}_q$ satisfy the stationary restriction (12), the chance of the real d being greater than 1 is very low. Otherwise, the possibility of the real d being greater than 1 can be checked through the characteristic equation (Brockwell and Davis 1996)

$$1 - \hat{\phi}_1 z - \cdots - \hat{\phi}_p z^p = 0. \tag{16}$$

The equation can be solved by the S-PLUS routine ''polyroot.'' An ad hoc rule is that if the characteristic equation (16) has a near-unit root x_1, (e.g., $|x_1 - 1| < 0.2$), then $d > 1$ is likely. Otherwise, $d = 1$ is optimal. More details on tests for unit roots may be found in Woodward and Gray (1995). To date, we have not encountered any global or regional temperature series for which $d > 1$ occurs or is likely to occur.

c. Test goodness of fit for optimal models

To ensure that the optimal trend models obtained in section 3b are not illposed statistical models, their goodness of fit has to be checked. The single most important diagnostic is the noise characteristics of the standardized prediction error series. If the model is correct, then this series should behave approximately like a white noise process with zero mean and unit variance (Brockwell and Davis 1996). The whiteness can be tested by the following goodness-of-fit statistic:

$$Q_K = (T - 1 - p) \sum_{k=1}^{K} \hat{\gamma}_k^2, \tag{17}$$

where $\hat{\gamma}_k$ is the estimated kth order autocorrelation of the prediction error series, and K is a fixed maximum number of lags between 10 and 20. If the correct model is fitted, then the goodness-of-fit statistic should be approximately distributed as a $\chi^2(K - r - k - p - q)$. The p values of the goodness-of-fit statistic can be estimated by the S-PLUS routine ''arima.diag.'' If there exists a K between 10 and 20 such that Q_K is significant at the 5% level, then the optimal model is illposed and therefore cannot represent the temperature series. If this is the case, the model must be discarded.

d. Discriminate between models

The final choice of model is based principally on which has the least BIC value, subject to the goodness-of-fit test and with consideration to the value of near-unit roots if any. In some cases, the choice may not be clear cut. It should be noted that the BIC approach can only be applied to ''nested'' models, that is, models of the same statistical family. An important reason for using differenced series is to nest models.

The simulation study described in section 3b above indicates that the BIC approach works well under the assumption that the true residual model is either ARIMA(p, 1, $q + 1$) with restriction (12) [i.e., ARMA(p, q)] or ARIMA(p, 1, q) without near-unit root for (14). This represents a weaker restriction (wider range of possibility) than the traditional restriction that the true residual model is only ARMA(p, q). Past results found under the traditional restriction may not necessarily remain valid under this weaker restriction.

Our simulation study indicates that the roots of Eq. (14) for the trend + ARIMA(p, 1, q) model are useful for judging the confidence that may be placed in the optimal model. If the true model is trend + ARMA(p, q), then Eq. (14) for trend + ARIMA(p, 1, q) is likely to have a very near-unit root (e.g., 1.01). Our experience is that the closer the roots are to unity, the more likely it is that the true model is trend + ARMA(p, q). However, if the true model is no trend + ARIMA(p, 1, q), with no near-unit root of (14), then including a trend component will only add a nuisance parameter with little likelihood of introducing a near-unit root to Eq. (14). Therefore, the closer the roots are to unity, the more unlikely the no trend + ARIMA(p, 1, q) model is.

The simulation study indicates that when the true residual model is ARIMA(p, 1, q) with a near-unit root of (14), the methods described here cannot adequately distinguish between competing models and hence may choose the wrong model. Unfortunately, to our best knowledge, no complete theory is yet available for discriminating between the structural time series models studied in this paper, which is why we have developed the current practical approach. Some recent developments on the topic have been reported by Arellano and Pantula (1995) and Davis and Dunsmuir (1996), but their models do not fully cover the regression part of the structural time series models studied in this paper and their discrimination procedures rely on particular nonstandard probability distributions.

As a further check on the discrimination procedure, for each estimated optimal model, a set of 100 independent time series with length T was simulated and the proposed fitting approach was applied to these simulated series. It was found that in every case, the original model was correctly discriminated to very high probability. This indicates that if the estimated optimal models are true, the proposed discrimination procedure can effectively identify them.

Last, spectral analysis can be useful to check the goodness of fit of the models for the differenced series. We have used the S-PLUS routine ''spec.pgram'' with a sequence of length 5 Daniell smoothers for the periodogram. The ($\alpha \times 100$)% of confidence interval at frequency λ is

$$\left[\frac{\nu\hat{f}(\lambda)}{\chi^2_{(1+\alpha)/2}(\nu)}, \frac{\nu\hat{f}(\lambda)}{\chi^2_{(1-\alpha)/2}(\nu)}\right], \tag{18}$$

where $\hat{f}$ is the smoothed periodogram and ν, the degree of freedom of the chi-square distribution, is directly calculated by spec.pgram. To map the spectral behavior of the residual models, the formula

$$\frac{\hat{\sigma}|1 + \hat{\theta}_1 e^{-i2\pi\lambda} + \cdots + \hat{\theta}_q e^{-i2\pi\lambda q}|^2}{2\pi|1 - \hat{\phi}_1 e^{-i2\pi\lambda} - \cdots - \hat{\phi}_P e^{-i2\pi\lambda p}|^2}, \quad 0 \le \lambda \le 1 \tag{19}$$

(Brockwell and Davis 1996) is used.

e. Calculate confidence intervals of trends

The choice of optimum model is based on the BIC difference between trend model and no-trend model, and not on the significance of a trend estimate, but the confidence intervals of the estimated trends remain of interest. Here we confine ourselves to trend functions that have a linear combination form (11). Consider the trend model

$$\nabla y_t = \lambda_1 \nabla f_{1,t} + \cdots + \lambda_r \nabla f_{r,t} + \beta_1 \nabla x_{1,t} + \cdots + \beta_{k'} \nabla x_{k',t} + \nabla \varepsilon_t, \qquad t = 2, \ldots, T, \tag{20}$$

where $\{x_{1,t}, \ldots, x_{k',t}\}$ is the explanatory variables set selected in step b. The covariance matrix of $\{\hat{\lambda}_1, \ldots, \hat{\lambda}_r; \hat{\beta}_1, \ldots, \hat{\beta}_{k'}\}$ can be estimated as

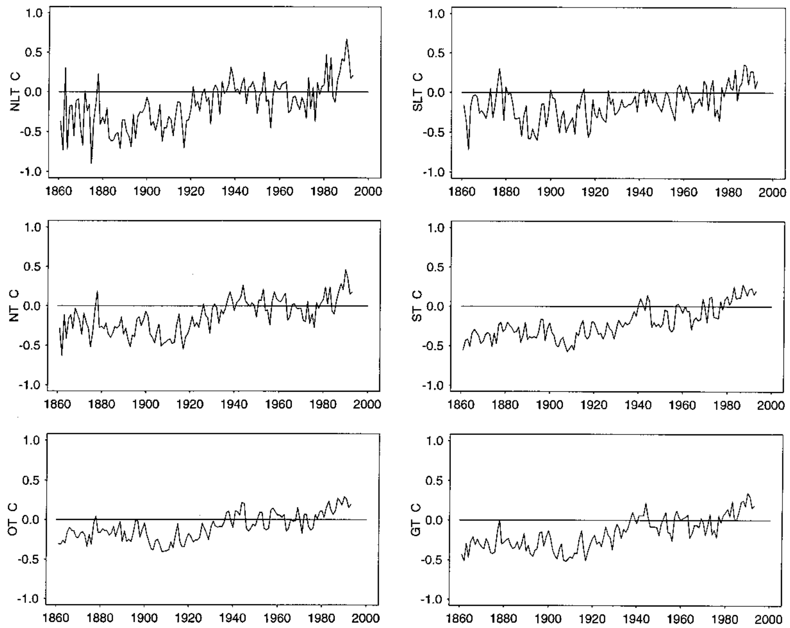

FIG. 1. IPCC global, global ocean, and hemispheric temperature series.

$$\hat{\Sigma} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1} \tag{21}$$

(Jones 1993), where $\mathbf{X}$ is a $T \times (r + k')$ matrix whose (t, i)th entry is $\nabla f_{i,t}$ if $i \leq r$, or $\nabla x_{i-r,t}$ if $i > r$ and $\mathbf{V}$, the $T \times T$ asymptotic autocovariance matrix of residuals, can be estimated by solving Yule–Walker equations (see appendix) with coefficients $\{\hat{\phi}_1, \ldots, \hat{\phi}_p; \hat{\theta}_1, \ldots, \hat{\theta}_q\}$ and innovation variance $\hat{\sigma}^2$. The Yule–Walker equations can be solved using the S-PLUS routine ''solve.'' This routine can also be used for calculating the inverse of a matrix with small dimension, such as $\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}$. However, the S-PLUS routine ''choleski'' should be used to calculate the inverse of a matrix with large dimension, such as $\mathbf{V}$.

The $(100 - \alpha)$% confidence interval for the estimated trend coefficient $\hat{\lambda}_1$ can be constructed as $\hat{\lambda}_1 \pm \hat{\sigma}_1 t_{\alpha/2}(T - 1)$, where $t_{\alpha/2}(T - 1)$ is the $100 - \alpha/2$ percentile of the t distribution with $T - 1$ degrees of freedom and $\hat{\sigma}_1$ is the standard deviation of the estimated variance of $\hat{\lambda}_1$. The confidence intervals for other coefficients can be constructed in the same way.

4. Data

The proposed approach was applied to the set of six annual temperature anomaly series 1861–1993 (Fig. 1) used in the IPCC Second Assessment Report (Houghton et al. 1996; Nicholls et al. 1996; Jones 1996), and to annual temperature anomaly series for six latitude belts (Fig. 2) corresponding to those used in Fig. 3.11 in Houghton et al. (1996). The IPCC series comprise the Northern Hemispheric land air temperature (NLT), Southern Hemispheric land air temperature (SLT) and global sea surface temperature (OT), and the combined Northern Hemispheric temperature (NT), Southern Hemispheric temperature (ST), and global temperature (GT).

The series for the latitude belts are formed from a 5° lat × 5° long gridded monthly mean temperature dataset archived at the Climate Research Unit, University of East Anglia (Jones 1996). The belts span 70°–55°N, 55°–30°N, 30°–10°N, 10°N–10°S, 10°–30°S, and 30°–55°S. The areas north of 70°N (which comprises only 3% of the global surface) and south of 55°S (9% of the global surface) are excluded owing to data sparseness. The latitude belt temperature series are plotted in Fig. 2. In the differenced series (not plotted) the high-frequency variability remains evident but the low-frequency variability is largely absent, and any trend is present as a constant bias.

Figure 3 shows the explanatory variables considered. The correlation between the SOI and hemispheric and global temperature is strongest when the index leads global temperatures by 6 months (Jones 1989), and therefore we have used the annual mean series of SOI for 12-month periods starting in the July of the year preceding the year under analysis. Series of annual sunspot count for 1880–1988 and Lamb's volcanic dust veil index for 1870–1983, obtained from National Center for Atmospheric Research archives, were also considered.

5. Results

We mainly consider the hypothesis of a trend starting in 1890, following suggestions that temperatures have increased mainly from the late nineteenth century (Nich-

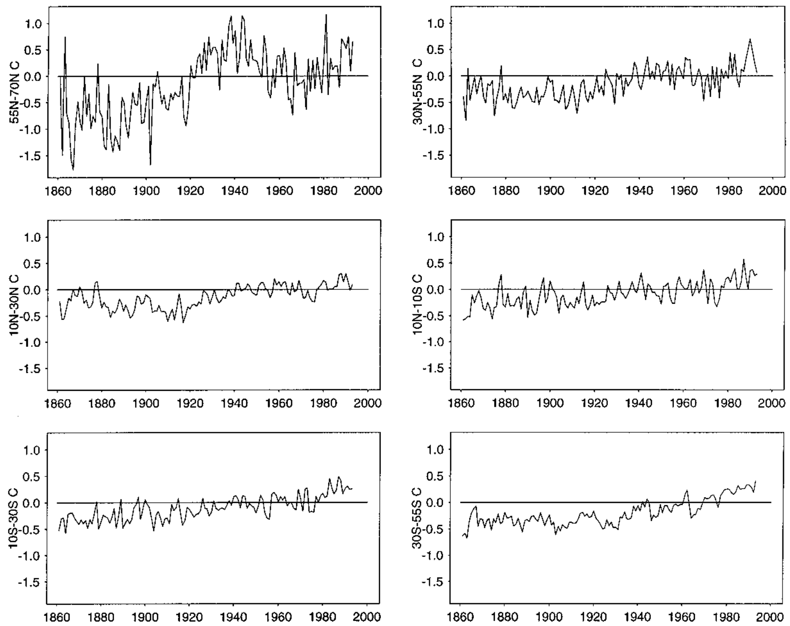

FIG. 2. Latitude-belt temperature series.

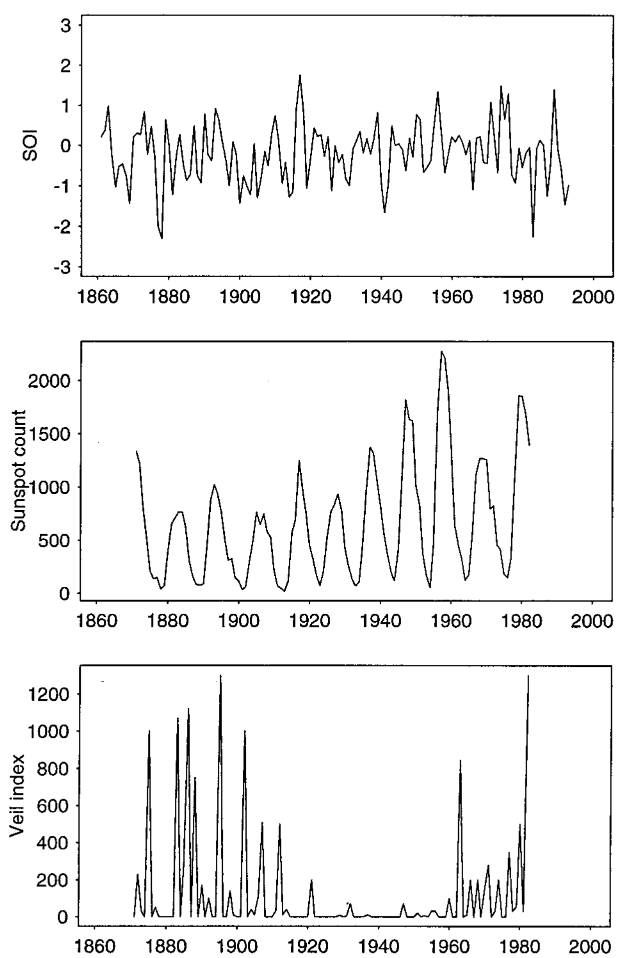

FIG. 3. Time series for explanatory variables.

olls et al. 1996; Santer et al. 1996). Our generic trend model is f_t + SOI + ARIMA(p, 1, q), where f_t is the bilinear function (7) with turning point $t_0 = 1890$ and $\lambda_1 = 0$. Table 1 presents the comparison of results for the various models and data sets. Because the BIC of the no trend + stationary residual model was always higher than those of the other options, its results are not listed. The finally selected models and associated trend estimates are listed in Table 2.

The SOI (with six-month lead to temperature) is found to be an explanatory variable in all series except those for the two latitude belts between 70° and 30°N. The use of the SOI as an explanatory variable significantly reduces the autoregressive order of the residual series in most cases. The SOI series is nonstationary (Gu and Philander 1995) and explains significant variance in the natural variability (Jones 1989), which means that if its influence is left in AR(p) [=ARMA(p, 0)] residuals, it will cause large-order p, as was found by Bloomfield (1992) and Woodward and Gray (1995). The sunspot count series is not an explanatory variable. This is similar to the results of Richards (1993) and Visser and Molenaar (1995), despite the different statistical model used by these authors. Lamb's dust veil index is not an explanatory variable, this result being similar to that of Richards (1993).

a. Latitude belts 70°–55°N, 55°–30°N

None of the models considered for these two belts are illposed (goodness-of-fit statistics not significant at 5% level). The BICs are relatively high, especially for the northernmost belt, owing to the high amounts of unmodeled natural variability for these belts. The BICs

TABLE 1. Structural models for latitude belt and IPCC temperature series. AR(p) = ARMA(p, 0) = ARIMA(p, 1, 1) with $\theta_1 = -1$. The BIC values pertain to the differenced series. The asterisk indicates models that fail the goodness-of-fit test.

Temperature series	Explanatory variable	Stationary residual form of trend model	BIC for trend + stationary residual	Nonstationary residual form for no-trend model	BIC for no-trend + nonstationary residual	Nonstationary residual form for trend model and its root of Eq. (14)	BIC for trend + nonstationary residual	BIC for random walk
70°–55°N		AR(4)	181.32	ARIMA (3, 1, 1)	169.35	ARIMA (3, 1, 1), 1.37	172.51	239.60*
55°–30°N		AR(4)	−13.45	ARIMA (3, 1, 1)	−15.63	ARIMA (3, 1, 1), 1.39	−13.48	52.10*
30°–10°N	SOI	AR(2)	−186.10	ARIMA (2, 1, 1)	−187.12	ARIMA (2, 1, 1), 1.16	−184.24	−177.90*
10°N–10°S	SOI	AR(1)	−153.32	ARIMA (1, 1, 1)	−144.54	ARIMA (1, 1, 1), 1.01	−148.44	−122.74*
10°–30°S	SOI	AR(1)	−178.28	ARIMA (1, 1, 1)	−168.70	ARIMA (1, 1, 1), 1.01	−173.41	−143.92*
30°–55°S	SOI	AR(1)	−187.72	ARIMA (1, 1, 1)	−186.37	ARIMA (1, 1, 1), 1.10	−187.04	−177.52*
NLT	SOI	AR (4)*	−35.56	ARIMA (4, 1, 0)*	−36.69	ARIMA (4, 1, 0)* —	−33.18	—
SLT	SOI	AR (1)	−145.63	ARIMA (1, 1, 1)	−141.96	ARIMA (1, 1, 1), 1.06	−140.15	−126.93*
NT	SOI	AR (4)*	−160.13	ARIMA (5, 1, 0)*	−169.32	ARIMA (6, 1, 0), —	−166.09	−128.31*
ST	SOI	AR (1)	−244.99	ARIMA (1, 1, 1)	−241.27	ARIMA (1, 1, 1), 1.04	−240.66	−233.75*
OT	SOI	AR (1)	−260.23	ARIMA (1, 1, 1)	−259.42	ARIMA (1, 1, 1), 1.18	−256.43	−253.03*
GT	SOI	AR (1)	−236.48	ARIMA (1, 1, 1)*	−233.45	ARIMA (1, 1, 1), 1.16	−231.58	−226.14*

favor the no trend + nonstationary residual model, these being lower by 2 than the BICs for the trend + stationary residual model. The roots of Eq. (14) for the trend + nonstationary residual models are not close to unity (>1.35). We conclude that the no-trend model is preferred and that a linear trend cannot be detected in the temperature series for these belts.

b. Latitude belt 30°–10°N

The situation for this latitude belt is less clear. None of the models is illposed. The BICs slightly favor the no trend + nonstationary residual model, whose BIC is about 1 higher than the BIC for the trend + stationary residual model. However, the root of Eq. (14) for trend + nonstationary residual is 1.16, which is relatively close to unity, which reduces confidence in the no-trend model. We therefore conclude that the presence of trend in this belt is a possibility, though at a relatively low confidence level.

c. Latitude belt 10°N–10°S and 10°–30°S

These belts exhibit clear trends. None of the models is illposed. The BICs of the trend + stationary residual model are about 10 lower than those for the no trend + nonstationary residual model, and hence are significantly in favor of the trend model. The roots of (14) for trend + nonstationary residual are very close to 1.0, which provides added confidence in the trend models. We conclude with high confidence that a linear upward trend is detected in the temperature series for this pair of latitude belts.

d. Latitude belt 30°–55°S

A trend is just detectable in this belt. None of the models is illposed. The BICs are slightly in favor of the trend + stationary residual model whose BIC is about 1 lower than the BIC for the no trend + nonstationary residual model. The root of (14) for the trend + nonstationary residual model is 1.10, but this is not suffi-

TABLE 2. Trends for latitude belt and IPCC temperature series. The dagger indicates the no-trend model is preferred over the trend model. The trend coefficients for models in which confidence is high or moderate are in bold type.

Temperature series	Structural time series model	Trend coefficient, from 1890 °C decade^{-1}	Std dev of trend estimate °C decade^{-1}
70°–55°N	†f_t + ARIMA (3, 1, 1)	0.128	0.6053
55°–30°N	†f_t + ARIMA (3, 1, 1)	0.070	0.2347
30°–10°N	†f_t + SOI + AR (2)	**0.040**	0.0094
10°N–10°S	f_t + SOI + AR (1)	**0.041**	0.0053
10°–30°S	f_t + SOI + AR (1)	**0.048**	0.0045
30°–55°S	f_t + SOI + AR (1)	**0.059**	0.0072
NTL	—	—	—
STL	f_t + SOI + AR (1)	**0.034**	0.0070
NT	f_t + SOI + ARIMA (6, 1, 0)	**0.055**	0.0021
ST	f_t + SOI + AR (1)	**0.048**	0.0056
OT	f_t + SOI + AR (1)	**0.035**	0.0060
GT	f_t + SOI + AR (1)	**0.046**	0.0060

ciently close to unity to provide additional support for the trend + stationary residual model. We therefore conclude that a linear upward trend is detected in the temperature series for this latitude belt, but with less confidence than for the 10°N–30°S region.

e. Northern Hemispheric series

The goodness of fit statistics are significant at the 5% level for all models for NLT, and for the trend + stationary residual model and the no trend + nonstationary residual model for NT, and therefore none of these models is acceptable. The trend + nonstationary residual model for NT passes the goodness-of-fit test and thus allows a trend to be determined. However, the results indicate that these particular data series are not well fitted by any of the structural time series considered here.

f. Southern Hemispheric series

None of the models for SLT and ST is illposed. For both series, the trend + stationary residual model is favored, its BIC being 3 lower than the BIC for the no trend + nonstationary residual model. The roots of (14) for the trend + nonstationary residual model (1.05) are close to 1. We conclude that linear upward trends are detected with good confidence for these series.

g. Global series

For OT, none of the models is ill-posed, and the BICs marginally favor the trend + stationary residual model (by 1) over the no trend + nonstationary residual model. For GT, the no trend + nonstationary residual model is illposed, and in any case the BICs favor the trend + stationary residual model by a good margin, of about 3, over the no trend + nonstationary residual model. For both series, the roots of (14) for the trend + nonstationary residual model are around 1.17, which is not sufficiently close to unity to provide added confidence in the trend models. We therefore conclude that linear upward trends are detected, but with only modest confidence for the OT series.

h. Other models

Results have been calculated for the SOI + random walk models (Table 1). In all cases the model fails the goodness-of-fit test, and the BICs are higher (often substantially so) than the corresponding BICs for the optimal no-trend model. This indicates that the random walk model is an illposed model for the IPCC temperature series, presumably because it fails to account for the series' temporal correlation. The random walk model without the SOI explanatory variable (Gordon 1991) gives even poorer results.

The roots of the characteristic equation for the optimal trend models were calculated. None of them is close to unity, which indicates that the residuals model ARIMA(p, d, q) with $d > 1$ is not an appropriate model.

The use of the model f_t + SOI + ARIMA(p, 1, q), where f_t is a linear trend starting at the beginning of the series in 1861 (i.e., not bilinear), was also examined. The BICs did not favor this model over the no-trend model. Likewise, a quadratic trend from 1890 was not favored over the linear trend model from 1890.

6. Discussion and conclusions

The unified statistical model and testing procedures we have proposed appear to provide a comprehensive means to choose structural time series for temperature trend detection purposes. Their application to the IPCC global temperature series has allowed the competing merits of the trend + stationary residual model and the no trend + nonstationary residual model to be clarified, and it has demonstrated that random walk models are unsatisfactory. In general, bilinear models with trend starting from 1890 and with SOI signal reduction and stationary residuals are preferred for at least half of the globe.

The results for the different latitude belts indicate a systematic latitudinal pattern of behavior. In the latitude belts surrounding the equator and the Southern Hemisphere Tropics (10°N–10°S and 10°–30°S), the bilinear trend model (trend starting in 1890, with SOI signal reduction and red noise residuals) is strongly preferred. To the south of this zone, in the southern midlatitudes (30°–55°S), it is also preferred, though less strongly so, while to the north of the zone, in the Northern Hemisphere Tropics (30°–10°N) it remains a possibility at low confidence, even though a no-trend model has the advantage of a slightly lower BIC. These belts together cover about two-thirds of the earth's surface.

Farther north, in the midlatitude and high-latitude belts above 30°N, the choice of optimum model fully switches to the no trend + nonstationary residual model, ARIMA(3, 1, 1). However, there is a significant amount of unmodeled prediction error for all the model options for these belts, as indicated by the high BIC values, especially for the 70°–55°N belt, and if a trend model is assumed, the estimated trend coefficients (Table 2) are relatively large (0.07 and 0.13°C decade^{-1}). In this situation, we therefore cannot assume that the absence of detectability means that a trend does not exist in these latitude belts. A trend, if present, may be masked by the significant natural variability that the existing trend models cannot properly represent.

It should be noted that, for these two northern belt series, if the residual model is restricted to the traditional ARMA(p, q) family (e.g., Zheng et al. 1997), a trend model would be selected by the BIC procedure and the 95% confidence interval would not include zero; that is, a trend would be ''detected.'' However, under the

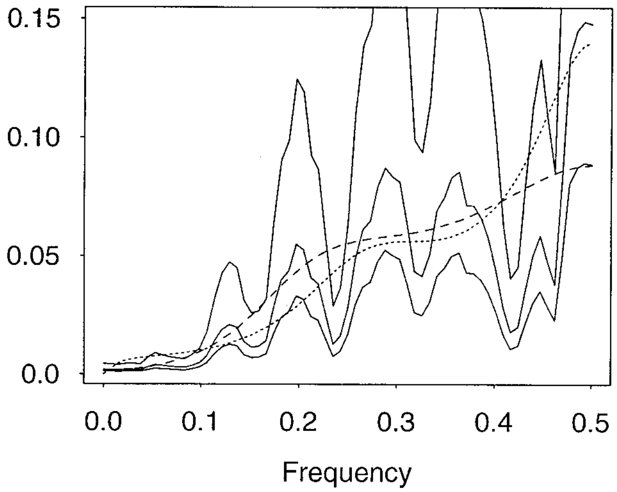

FIG. 4. Spectral density for differenced series for 70°–55°N. Smoothed periodogram with 80% confidence interval estimates (solid lines), ARIMA(3, 1, 1) model (dashed), and AR(4) [=ARIMA(4, 1, 1) with $\theta_1 = -1$] model (dotted).

weaker restriction on residual models that we use here, a completely different conclusion is derived.

It is of interest that the northern parts of the Northern Hemisphere are marked by greater variability (Fig. 2). Using singular spectral decomposition, Schlesinger and Ramankutty (1994) identified a low-frequency oscillation of period around 65–70 yr that is particularly strong at these latitudes, especially in the North Atlantic and North America zones, but is relatively weak in the Tropics and the Southern Hemisphere. They concluded that the oscillation arises from internal variability of the ocean–atmosphere system, rather than anthropogenic causes. The pattern is also apparent in the most recent IPCC summary of regional variations (Karl 1998). The preference for the no trend + ARIMA(p, 1, q) model may be associated with this variability feature.

As a further diagnostic of why the trend + ARMA(p, q) model is not preferred for the 70°–55°N belt, spectral analysis was applied to the *differenced* series and to the model options [Eq. (19)]. As expected, the differencing removes the 65–70-yr oscillation and the series shows near-zero power at low frequencies (Fig. 4). Over the low-frequency range $1/66 \leq \lambda \leq 1/15$, the spectrum for the no trend + ARIMA(3, 1, 1) model coincides very well with the (near zero) smoothed periodogram, but the spectrum for the trend + ARMA(4, 0) model lies above at the 20% significance level. Thus it appears that the trend + ARMA(p, q) model cannot model the small but crucial low-frequency natural variability in the differenced series as well as the no trend + ARIMA(3, 1, 1) model can.

A broad conclusion from the results for the latitudinal belts is that there is more than one category of statistical process present. A trend process is clearly evident in the tropical belts and Southern Hemisphere midlatitudes, while other processes that interfere in trend detection are present only at high latitudes, especially in the Northern Hemisphere. The uncertain result for the low-latitude Northern Hemisphere belt (30°–10°N) is likely to be due to the presence there of both processes. This leads to the idea that the main hemispheric and global series will comprise a mixture of processes that may not be well represented by a single structural time series.

On this basis, the NT and NLT series will require a mixture of an ARIMA(3, 1, 1) process and a trend + SOI + AR(1) process, which is the probable reason that both the optimal trend model and the no-trend model are illposed and no significant trend is detectable. However, the Southern Hemisphere counterparts, ST and SLT, which are dominated by the trend process, are not illposed and exhibit clear trends. The oceanic OT series is similarly dominated by the Southern Hemisphere and the trend process, and none of its models are illposed. The complete GT series contains a greater proportion of higher-latitude regions than the OT series, but the tropical and Southern Hemispheric part still dominates the mixture and allows a trend model to be fitted. The quality of fit of a model is thus related to the relative ''purity'' of the mixture of processes present. This indicates that the selection of regions for analysis is important for developing good structural time series models. The major regional series in Karl (1998) indicates that there may be significant variations within each latitudinal belt.

The linear trend coefficients are reasonably similar for all of the series, being about +0.04 to +0.05°C decade^{-1} (Table 2). The geographical pattern of the magnitudes of the trends is broadly consistent with GCM greenhouse warming predictions (Kattenberg et al. 1996), being generally greater for the higher latitudes and lower for the Tropics and oceans. The intrinsic detectability of the temperature trends, however, is opposite to this latitudinal variation, being best in the Tropics and poorest in high latitudes. This indicates that particular attention needs to be paid to the tropical and subtropical regions in trend detection studies. Methods of global trend pattern detection that seek to maximize the signal to noise ratio globally have been described (Santer et al. 1996; Hasselmann 1997; Hegerl et al. 1997).

A similar conclusion does not arise for the relative detectability of trends for continents, where trends are predicted to be greater, versus lower-trend oceans. The distributions and higher proportion of land and associated land–ocean interactions presumably play a key role in the different results we find for the Northern Hemisphere, especially at higher latitudes, but the relatively higher natural variability of temperature series for land areas (see Fig. 1) has little part in the detectability of trends, which is judged on the BIC advantage of the trend model over the no-trend model. For example, the BICs of the SLT models are much higher than those for the ST model, but in each case the BIC advantage of the trend model is essentially the same, at about 3.6.

We conclude that for a large part of the globe a linear trend is detected with good confidence, and that in the high northern latitudes, where trend cannot be detected by fitting a structural time series model, the existence of trend cannot be ruled out. The results depend on the assumption that the true residual model is either ARMA(p, q) or ARIMA(p, 1, q) with no near-unity root of (14), but this is a weaker constraint than previously assumed and it leads to more robust conclusions. While our study does not directly address the problem of attribution, it does reinforce the view that the globe is subject to a temperature forcing factor that is global in extent, is likely to continue, and is consistent with greenhouse warming predictions.

Acknowledgments. This work was supported by the New Zealand Foundation for Research, Science and Technology under Contract CO1621. We wish to thank Dr. Phillip Jones for providing the IPCC temperature series, and Dr. Brett Mullan and Mr. Richard Rayner for assistance in obtaining the 5° latitude–longitude temperature data and the series of sunspot counts and Lamb's volcanic dust veil index. We are indebted to Dr. Francis Zwiers and the other journal's referees for their critical comments, which have led to some important clarifications and improvement.

APPENDIX

Yale–Walker Equations

The Yale–Walker equations for $\{U(t), t = 0, \ldots, T\}$ with autoregressive coefficients $\{\phi_1, \ldots, \phi_p\}$, moving-average coefficients $\{\theta_1, \ldots, \theta_q\}$, and innovation variance σ^2 (Brockwell and Davis 1996) are

$$U(t) - \phi_1 U(t-1) - \cdots - \phi_p U(t-p) = \sigma^2 \sum_{t \le s \le q} \theta_s \psi_{s-t}, \qquad 0 \le t < \max(p, q+1)$$

$$U(t) - \phi_1 U(t-1) - \cdots - \phi_p U(t-p) = 0, \qquad t \ge \max(p, q+1),$$

where coefficients $\{\psi_s\}$ are the solution of the following equations:

$$\psi_t - \sum_{0<s\le t} \phi_s \psi_{t-s} = \theta_t, \qquad 0 \le t < \max(p, q+1)$$

$$\psi_t - \sum_{0<s\le p} \phi_s \psi_{t-s} = 0, \qquad t \ge \max(p, q+1).$$

REFERENCES

Arellano, C., and S. G. Pantula, 1995: Testing for trend stationary versus difference stationary. *J. Time Ser. Anal.,* **16,** 147–164.

Bloomfield, P., 1992: Trend in global temperature. *Climate Change,* **21,** 1–16.

——, and D. W. Nychka, 1992: Climate spectra and detecting climate change. *Climate Change,* **21,** 275–287.

Brockwell, P. J., and R. A. Davis, 1996: *Time Series: Theory and Methods.* Springer-Verlag, 577 pp.

Davis, R. A., and W. T. M. Dunsmuir, 1996: Maximum likelihood estimation for MA(1) processes with a root on or near the unit circle. *Econometric Theory,* **6,** 433–444.

Galbraith, J., and C. Green, 1992: Inference about trends in global temperature data. *Climate Change,* **22,** 209–221.

Gordon, A. H., 1991: Global warming as a manifestation of a random walk. *J. Climate,* **4,** 589–597.

——, J. A. T. Bye, and R. A. D. Byron-Scott, 1996: Is global warming climate change? *Nature,* **380,** 478.

Gu, D., and S. G. H. Philander, 1995: Secular changes of annual and interannual variability in the Tropics during the past century. *J. Climate,* **8,** 864–876.

Hasselmann, K., 1997: Multi-pattern fingerprint method for detection and attribution of climate change. *Climate Dyn.,* **13,** 601–611.

Hegerl, G. C., K. Hasselmann, U. Cubasch, J. F. B. Mitchell, E. Roeckner, R. Voss, and J. Waszkewitz, 1997: Multi-pattern fingerprint analysis of greenhouse gas, greenhouse gas-plus-aerosol and solar forced climate change. *Climate Dyn.,* **13,** 613–634.

Houghton, J. T., L. G. Meira Filho, B. A. Callander, N. Harris, A. Kattenberg, and K. Maskell, Eds., 1996: *Climate Change 1995: The Science of Climate Change. Contribution of Working Group I to the Second Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge University Press, 572 pp.

Jones, P. D., 1989: The influence of ENSO on global temperatures. *Climate Monit.,* **17,** 80–90.

——, 1996: Hemispheric and global temperatures, 1851–1996, Part 1. *Climate Monit.,* **25,** 20–30.

Jones, R. H., 1980: Maximum likelihood fitting of ARMA models to time series with missing observations. *Technometrics,* **22,** 389–395.

——, 1993: *Longitudinal Data with Series Correlation: A State-Space Approach.* Chapman and Hall, 225 pp.

Karl, T. R., 1998: Regional trends and variations of temperature and precipitation. *The Regional Impacts of Climate Change, An Assessment of Vulnerability,* R. T. Watson et al., Eds., Cambridge University Press, 411–425.

Kattenberg, A., and Coauthors, 1996: Climate models—Projections of future climate. *Climate Change 1995: The Science of Climate Change. Contribution of Working Group I to the Second Assessment Report of the Intergovernmental Panel on Climate Change,* J. T. Houghton et al., Eds., Cambridge University Press, 285–357.

Nicholls, N., G. V. Gruza, J. Jouzel, T. R. Karl, L. A. Ogallo, and D. E. Parker, 1996: Observed climate variability and change. *Climate Change 1995: The Science of Climate Change. Contribution of Working Group I to the Second Assessment Report of the Intergovernmental Panel on Climate Change,* J. T. Houghton et al., Eds., Cambridge University Press, 133–192.

Richards, G. R., 1993: Change in global temperature: A statistical analysis. *J. Climate,* **6,** 546–559.

Santer, B. D., T. M. L. Wigley, T. P. Barnett, and E. Anyamber, 1996: Detection of climate change and attribution of causes. *Climate Change 1995: The Science of Climate Change. Contribution of Working Group I to the Second Assessment Report of the Intergovernmental Panel on Climate Change,* J. T. Houghton et al., Eds., Cambridge University Press, 407–444.

Schimel, D., and Coauthors, 1996: Radiative forcing of climate change. *Climate Change 1995: The Science of Climate Change. Contribution of Working Group I to the Second Assessment Report of the Intergovernmental Panel on Climate Change,* J. T. Houghton et al., Eds., Cambridge University Press, 407–444.

Schlesinger, M. E., and N. Ramankutty, 1994: An oscillation in the global climate system of period 65–70 years. *Nature,* **367,** 723–726.

StatSci., 1993: S-PLUS guide to statistical and mathematical analysis, version 3.2. Statistical Sciences, 630 pp. [Available from StatSci Division, 1700 Westlake Ave. N., Suite 500, Seattle, WA 98109.]

Visser, H., and J. Molenaar, 1995: Trend estimation and regression analysis in climatological time series: An application of structural time series models and the Kalman filter. *J. Climate,* **8,** 969–979.

Wigley, T. M. L., and P. D. Jones, 1981: Detecting CO_2-induced climate change. *Nature,* **292,** 205–208.

——, ——, P. M. Kelly, and S. C. B. Raper, 1989: Statistical significance of global warming. *Proc. 13th Annual Climate Diagnostics Workshop,* A1–A8.

Woodward, W. A., and H. L. Gray, 1993: Global warming and the problem of testing for trend in time series data. *J. Climate,* **6,** 953–962.

——, and ——, 1995: Selecting a model for detecting the presence of a trend. *J. Climate,* **8,** 1929–1937.

Zheng, X., R. E. Basher, and C. S. Thompson, 1997: Trend detection in regional-mean temperature series: Maximum, minimum, mean, diurnal range, and SST. *J. Climate,* **10,** 317–326.