Time Series Analysis of Circular Data

Author(s): N. I. Fisher and A. J. Lee

Source: *Journal of the Royal Statistical Society. Series B (Methodological)*, 1994, Vol. 56, No. 2 (1994), pp. 327-339

Published by: Wiley for the Royal Statistical Society

Stable URL: https://www.jstor.org/stable/2345903

J. R. Statist. Soc. B (1994)
56, *No.* 2, *pp.* 327–339

Time Series Analysis of Circular Data

By N. I. FISHER† and A. J. LEE

CSIRO, North Ryde, Australia *University of Auckland, New Zealand*

[Received January 1992. Final revision January 1993]

SUMMARY

The paper considers the problem of modelling and analysing a time series of measurements which take the form of unit two-dimensional vectors. Several general classes of existing and new models are studied in terms of feasibility and flexibility. An approach is recommended which uses two of these classes, and which takes advantage of several standard time series algorithms that are available in modern software packages. An application is given to the analysis of a series of data on wind directions.

Keywords: AUTOREGRESSIVE MOVING AVERAGE PROCESSES; CIRCULAR AUTOCORRELATION; EM ALGORITHM; MAXIMUM LIKELIHOOD; VON MISES DISTRIBUTION; WRAPPED NORMAL DISTRIBUTION

1. INTRODUCTION

Methods for analysing time series of circular data (or *angular* data) have not received much attention. Such methods as have been proposed can be classified into two broad groups, depending on which type of underlying model is appropriate: Markov process models (Watson and Beran, 1967; Wehrly and Johnson, 1979; Accardi *et al.*, 1987, Breckling, 1989) and models based on wrapping a linear autoregressive (AR) process around the circle (Breckling, 1989). Time series of directions having associated magnitudes (e.g. velocities of wind or ocean currents) can be handled by standard bivariate time series methods (see for example Gonella (1972)).

In this paper, we revisit the wrapped AR process developed by Breckling (1989) in the context of a major study of wind patterns in Perth, Western Australia. We also propose two new models, one based on a projection method and the other based on a link function concept adapted to the time series context from regression ideas introduced in an unpublished paper by Downs (1986) and by Fisher and Lee (1992). The projection method provides a way of modelling circular times series for which the marginal distributions are the circular uniform distribution. The link function approach allows description and analysis of a wide variety of processes with non-uniform marginals.

In Section 2, the old and new models are described in detail. The discussion is largely couched in terms of AR processes. We deal with model selection and identification of time series in Section 3 and with model fitting in Section 4. An application is given in Section 5. Several technical derivations have been relegated to Appendix A.

†*Address for correspondence*: CSIRO Division of Mathematics and Statistics, Locked Bag 17, North Ryde, NSW 2113, Australia.

 0035–9246/94/56327

2. SOME MODELS FOR TIME SERIES OF CIRCULAR DATA

In this section we discuss four models for circular time series and describe the form of the circular autocorrelation for each. We begin with the projection method, which allows circular time series with uniform marginal distributions to be analysed. None of the other models can cope with this circumstance.

2.1. *Projected Gaussian Processes*

Let $\{X_t\}$ and $\{Y_t\}$ be two independent realizations of the same stationary zero-mean Gaussian process having unit variance. Then, if (R_t, Θ_t) is the representation of the point (X_t, Y_t) in polar co-ordinates, $\{\Theta_t\}$ is a stationary circular time series with the property that *its marginal distributions are uniform*. If $\{X_t\}$ and $\{Y_t\}$ have *non-zero* means, then the resulting projected process does not have uniform marginals: they are, in fact, offset normal (angular Gaussian). The details in this case appear to be mathematically intractable.

The fitting of such models amounts to a missing data problem, since we observe the 'incomplete' data Θ_t rather than the complete data X_t and Y_t. The EM algorithm can be used for analysis, as is shown in Section 3.

To examine the correlation patterns of $\{\Theta_t\}$, we require a measure of correlation between two circular random variables Θ and Φ.

Let (Θ_1, Φ_1) and (Θ_2, Φ_2) be independent realizations of (Θ, Φ), and define

$$\rho_T = \frac{E \sin(\Theta_1 - \Theta_2) \sin(\Phi_1 - \Phi_2)}{\{E \sin(\Theta_1 - \Theta_2)^2 E \sin(\Phi_1 - \Phi_2)^2\}^{1/2}}.$$

This is the circular correlation coefficient introduced by Fisher and Lee (1983). It takes values between -1 and 1, and is 0 if Θ and Φ are independent. See Fisher and Lee (1983) for other properties. Using this correlation, the following result can be derived (see Appendix A).

Lemma. Let X_1, Y_1 and X_2, Y_2 be independent random vectors from a bivariate normal distribution with variances both σ^2 and correlation ρ. Let Θ_1 and Θ_2 be defined by $(X_i, Y_i) = R_i(\cos\Theta_i, \sin\Theta_i)$, $i = 1, 2$. Then the circular correlation between Θ_1 and Θ_2 is given by

$$\rho_T = \frac{\pi^2}{16} \rho^2 (1-\rho^2)^2 \left\{ {}_2F_1\left(\frac{3}{2}, \frac{3}{2}, 2; \rho^2\right)\right\}^2 \tag{2.1}$$

where ${}_2F_1$ is the usual hypergeometric function notation.

Hence, if $\rho(k)$ is the common autocorrelation function of the X_t- and Y_t-processes, then the Θ_t-process has autocorrelation function

$$\rho_T(k) = \frac{\pi^2}{16} \rho^2(k) \{1-\rho^2(k)\}^2 \left\{ {}_2F_1\left(\frac{3}{2}, \frac{3}{2}, 2; \rho^2(k)\right)\right\}^2. \tag{2.2}$$

The variances of the X_t- and Y_t-processes play no part in the definition of the projected process, and we have arbitrarily taken them to be 1 in what follows.

2.2. *Wrapped Processes*

Suppose that $\{X_t\}$ is a univariate time series of (linear) observations. The corresponding wrapped circular time series $\{\Theta_t\}$ is given by

$$\Theta_t = X_t\,[\mathrm{mod}(2\pi)].$$

Following the approach taken by Breckling (1989), we can think of X_t as being decomposed as

$$X_t = \Theta_t + 2\pi k_t$$

where k_t is an unobserved integer. Viewed in this way, we see that the problem of fitting such a model again reduces to a missing data problem (see Section 2.1)—Θ_t being observed instead of the complete data X_t—so that the EM algorithm again suggests itself as a suitable approach. In practice, the algorithm takes a reasonably simple form for low order AR processes. The details are given in Section 4 for a wrapped Gaussian process.

The correlation function of a wrapped Gaussian AR(p) (WAR(p)) process is easily found. From Fisher and Lee (1983) we see that, if (X, Y) has a bivariate normal distribution with variances σ_X^2 and σ_Y^2 and correlation ρ, then the circular correlation ρ_T of $\Theta = X[\mathrm{mod}(2\pi)]$ and $\Phi = Y[\mathrm{mod}(2\pi)]$ is

$$\rho_T = \frac{\sinh(2\rho\sigma_X\sigma_Y)}{\{\sinh(2\sigma_X^2)\sinh(2\sigma_Y^2)\}^{1/2}}.$$

Hence, if $\{X_t\}$ is an AR(p) process, the circular autocorrelation function of the WAR(p) process θ_t is given by

$$\rho_T(k) = \frac{\sinh\{2\rho_k\sigma^2/(1-\phi_1\rho_1-\ldots\phi_p\rho_p)\}}{\sinh\{2\sigma^2/(1-\phi_1\rho_1-\ldots\phi_p\rho_p)\}}$$

where ρ_k is the k-lag autocorrelation of $\{X_t\}$ and σ^2, $\phi_1, \ldots, \phi_p$ are its AR(p) parameters.

2.3. *Processes Derived using Link Functions*

Let $g(x)$ be an odd monotone function mapping the real line onto the interval $(-\pi, \pi)$ and such that $g(0)=0$. The particular choice

$$g_T(x) = 2\tan^{-1}x$$

was used by Downs (1986) to formulate a regression model for an angular response and an angular covariate. Fisher and Lee (1992) formulated a general class of regression models for an angular response variate and linear covariates using a general link g, and then gave an example using g_T.

For such link functions g, if X is a real variate then

$$\Theta = g(X)$$

is a circular variate and conversely, for a circular variate Θ,

$$X = g^{-1}(\Theta)$$

is a real variate.

Using these transformations, we say that a circular stationary process $\{\Theta_t\}$ with mean direction μ is a linked autoregressive moving average (LARMA) process if its *linked linear process* $g^{-1}(\Theta_t-\mu)$ is an ARMA(p, q) process.

Suppose that θ_t is a LARMA(p, q) process with mean direction μ. Then $X_t = g^{-1}(\theta_t - \mu)$ is an ARMA(p, q) process, and the circular autocorrelation of θ_t is

$$\rho_T(k) = \rho_T\{g(X_t),\ g(X_{t+k})\},$$

where ρ_T denotes the circular correlation coefficient of Section 2.1. Graphs of this function in the LAR(1) case are shown in Fig. 1.

For AR processes, an alternative model can be defined by using conditional distributions. In this case, we can define a circular AR(p) process as one for which the conditional distribution of Θ_t given $\theta_{t-1}, \theta_{t-2}, \ldots$ is von Mises VM(μ_t, κ), or any other circular distribution symmetric about its mean, where the mean direction μ_t is given by

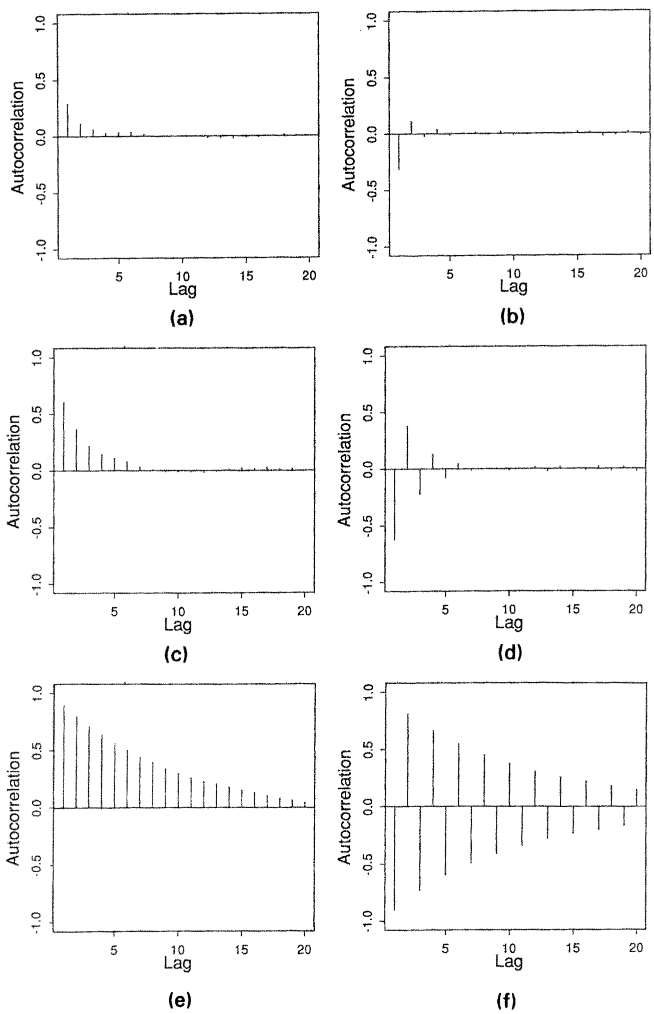

Fig. 1. Prototype autocorrelation functions for an LAR(1) model using a probit link (all classes of model have similar patterns for the autocorrelation function, and the patterns are similar to those of a linear AR(1) process): (a) $\phi = 0.3$, $\sigma = 0.2$; (b) $\phi = -0.3$, $\sigma = 0.2$; (c) $\phi = 0.6$, $\sigma = 0.2$; (d) $\phi = -0.6$, $\sigma = 0.2$; (e) $\phi = 0.9$, $\sigma = 0.2$; (f) $\phi = -0.9$, $\sigma = 0.2$

$$\mu_t = \mu + g\{\omega_1 g^{-1}(\theta_{t-1} - \mu) + \ldots + \omega_p g^{-1}(\theta_{t-p} - \mu)\} \tag{2.3}$$

and κ is the constant concentration parameter. We thus have two circular forms of AR(p) model, the *direct or linked form* LAR(p) where the transformed angles follow an ordinary linear AR(p) model and the *inverse form* IAR(p) where the conditional distributions have the conditional means (2.3).

What are the connections between the error structures of Θ_t and X_t? Suppose that the density of Θ_t is f_θ. Then the density of $X_t = g^{-1}(\Theta_t)$ is

$$f_X(x) = f_\theta\{g(x)\}g'(x).$$

If for example $g(x) = \tan(x/2)$ and f_θ is not too concentrated, the density of f_X behaves in the tails like $(1+x^2)^{-1}$, so that methods based on the normal distribution for the analysis of the transformed series X_t are unlikely to be successful. A better choice for g is a scaled version of the probit transformation, i.e.

$$g(x) = 2\pi\{\Phi(x) - 0.5\}, \tag{2.4}$$

where $\Phi(x)$ is the distribution function of the standard normal. This will induce in the distribution of X_t tail behaviour that is roughly equivalent to that of the normal distribution. Consequently a package implementing the standard Box–Jenkins methodology should be adequate for the analysis of the transformed series.

The advantages of the linked linear model are

(a) if the process $\{\Theta_t\}$ is stationary then so is the resulting process X_t and
(b) the parameters are easily estimated by using standard software.

The inverse model allows the direct specification of the circular error structure and is suitable for dispersed data, but the fitting of the inverse models becomes complicated for p much beyond 2 or 3 if methods for maximizing the relevant likelihoods by using exact derivatives are used. In the AR(1) case, there is a unique marginal distribution for which the process is stationary. This distribution is in fact the asymptotic stationary distribution of the (Markov) AR(1) process, and is not von Mises.

The autocorrelation function in this case is very similar in appearance to that of the linked process depicted in Fig. 1.

3. MODEL SELECTION AND IDENTIFICATION

Unlike the situation for linear time series models, where we typically use a single family of autoregressive integrated moving average models (namely, those with Gaussian errors), in the present case we have four families of models at our disposal. In Section 3.1 we discuss ways of making an appropriate choice of one of these models. In Section 3.2, we discuss how to choose a particular model within a given class.

3.1. *Model Selection*

In the absence of contextual information that might indicate a specific class of models on physical grounds, we must base our choice of class on the observable features of the data. It appears that all four classes have similar correlation patterns: selected autocorrelation functions for the linked LAR(1) process are shown in Fig. 1 and are similar in form to the typical exponential decay pattern in linear AR(1) models. Note that the projected AR(1) process always has positive lag 1 autocorrelation, and that its autocorrelations are of the order of the squares of those shown in Fig. 1.

Thus, in making a choice of the class of model, we must inspect other features of the time series.

The projection model is characterized by having uniform marginals. If the linear processes being projected have high correlation, the projected process will tend to occupy changing arcs of the circle, corresponding to the level shifts observed in highly correlated linear processes. For low degees of correlation, the circular process will wander about the whole circle in a rapidly changing manner.

For the wrapped normal process, both the correlation and the variance of the wrapping (linear) process affect the appearance of the circular process. For larger values of the variance, the circular process will tend to be spread uniformly around the circle, and for data of this type either the projection model or the wrapped normal model will be appropriate. For small variances, the circular (wrapped) process will tend to occupy an arc of the circle, rather than spread uniformly around the whole circumference.

The inverse processes can capture similar behaviour. The smaller the value of the κ-parameter, the more uniformly spread will be the data around the circle. The larger the value of ω is, the greater the tendency to remain for short periods in a small arc about μ.

Finally, the linked process is characterized by a gap in the way that the data distribute themselves around the circle, caused by the fact that only very large values of the underlying process can project into an arc around $\mu+\pi$. The spread of the circular process around the rest of the arc depends on the specific link function chosen and the variation of the underlying linear process.

On the basis of these observations, we can make the following recommendations. If the circular data are distributed uniformly around the circle, the projection model or the wrapped normal model is appropriate, the latter being expected to have a large variance for the wrapping (linear) process.

If, in contrast, the data tend to cluster in an arc, the indirect and linked methods will be more appropriate. If a definite gap in the data is observed, the direct model is the model of choice.

However, all four models are quite flexible and capable of capturing a wide range of behaviours.

3.2. *Model Identification*

As in linear time series analysis, the selection of an appropriate model for a circular times series is aided by inspection of the correlogram. For the projected, wrapped normal and inverse models, we can calculate the sample correlogram of the series efficiently using an alternative expression for the sample autocorrelations (see Fisher and Lee (1986)),

$$\hat{\rho}_T(k) = \frac{\det\left(\sum_{t=1}^{T-k} X_t X'_{t+k}\right)}{\left\{\det\left(\sum_{t=1}^{T-k} X_t X'_t\right)\det\left(\sum_{t=k+1}^{T} X_t X'_t\right)\right\}^{1/2}} \tag{3.1}$$

where $X_t = (\cos\theta_t, \ \sin\theta_t)'$.

The sample correlogram (3.1) is then compared with the theoretical autocorrelations of the various models. A tentative model is then identified and fitting proceeds as in the next section.

In the projection model, we can obtain a rough estimate of the autocorrelation function $\rho(k)$ of the linear process by solving equation (2.2) for $\rho(k)$, using the sample autocorrelation (3.1) in place of $\rho_T(k)$.

Similarly, the autocorrelation of the wrapping process in the wrapped normal model can be estimated by using

$$\hat{\rho}_T(k) = \sinh(2c_0\hat{\rho}_k)/\sinh(2c_0).$$

Here, c_0 is the variance (i.e. the autocorrelation at lag 0) of the wrapping process. It can be estimated from the relationship $\bar{R} = \exp(-c_0/2)$ where $\bar{R}$ is the mean resultant length of $\{\theta_t\}$.

To identify the direct model, a preliminary estimate $\hat{\mu}$ of μ is made, e.g. by calculating the mean direction of the circular series. The transformed series $X_t = g^{-1}(\Theta_t - \hat{\mu})$ is then identified in the usual way.

4. MODEL FITTING

4.1. *Wrapped Normal Model*

We discuss the fitting of the WAR(p) model. For a detailed discussion of the AR(1) model see Breckling (1989). We can parameterize the wrapped linear process X_t by the mean μ, the process variance c_0, say, and the first p lagged covariances $c_1, \ldots, c_p$. A vector s of statistics that is (almost) sufficient for $(\mu, c_0, \ldots, c_p)$ is $(\Sigma X_t, \Sigma X_t^2, \ldots, \Sigma X_t X_{t-p})$. Given a circular time series $\theta_1, \ldots, \theta_T$, the EM algorithm takes the form

(a) E-step—given an estimate γ_N of $(\mu, c_0, \ldots, c_p)$ compute

$$s_N = E[s|\gamma_N, \theta_1, \ldots, \theta_T]; \tag{4.1}$$

(b) M-step—compute the updated estimate γ_{N+1} as the solution of

$$s_N = E[s|\gamma_{N+1}]. \tag{4.2}$$

The M-step is trivial. The major difficulty is in the computation of the conditional expectation (4.1). We illustrate with the calculation of $E[\Sigma X_t X_{t-1}|\theta_1, \ldots, \theta_T]$. The conditional expectation of $X_t X_{t-1}$ given $\theta_1, \ldots, \theta_T$ is

$$\frac{\sum_{m_1=-\infty}^{\infty} \cdots \sum_{m_T=-\infty}^{\infty} (\theta_t+2\pi m_t)(\theta_{t-1}+2\pi m_{t-1}) f(\theta_1+2\pi m_1, \ldots, \theta_t+2\pi m_T)}{\sum_{m_1=-\infty}^{\infty} \cdots \sum_{m_T=-\infty}^{\infty} f(\theta_1+2\pi m_1, \ldots, \theta_t+2\pi m_T)} \tag{4.3}$$

where f is the density of the AR(p) process. The amount of computation involved in evaluating this expression is prohibitive. However, an approximation can be obtained by observing that the influence of θ_{t-j} on θ_t declines quickly with increasing j so that $E[X_tX_{t-1}|\theta_1, \ldots, \theta_T]$ is reasonably well approximated for the AR(1) process by $E[X_tX_{t-1}|\theta_t, \theta_{t-1}]$ which is given by the more tractable expression

$$\frac{\sum_{m_t=-\infty}^{\infty} \sum_{m_{t-1}=-\infty}^{\infty} (\theta_t+2\pi m_t)(\theta_{t-1}+2\pi m_{t-1}) f_{t,t-1}(\theta_1+2\pi m_t, \theta_{t-1}+2\pi m_{t-1})}{\sum_{m_t=-\infty}^{\infty} \sum_{m_{t-1}=-\infty}^{\infty} f_{t,t-1}(\theta_1+2\pi m_1, \theta_{t-1}+2\pi m_{t-1})} \tag{4.4}$$

where $f_{t,t-1}$ is the joint density of (X_t, X_{t-1}). For general ARMA models, the covariances are no longer sufficient, and the EM equations take an even more intractable form. Thus, the wrapped normal approach is only feasible for low order AR models. Breckling (1989) uses the approximation (4.4) rather than the exact expression (4.3).

4.2. *Projection Model*

We discuss the AR(p) case only. As before, we parameterize the linear processes X_t and Y_t in terms of the mean μ, the process variance c_0 and the first p lagged covariances $c_1, \ldots, c_p$. The EM algorithm is again applied. As in the wrapped normal case, the M-step is trivial. The major difficulty is again in the computation of the conditional expectation (4.1). Consider (writing $(R_t \cos \Theta_t, R_t \sin \Theta_t)$ for the polar representation of (X_t, Y_t))

$$\begin{aligned} E[X_tX_{t-j}|\theta_1, \ldots, \theta_T] &= E[R_t \cos \Theta_t R_{t-j} \cos \Theta_{t-j}|\theta_1, \ldots, \theta_T] \\ &= \cos \theta_t \cos \theta_{t-j} E[R_tR_{t-j}|\theta_1, \ldots, \theta_T]. \end{aligned}$$

We now appeal to the following fact: if X and Y have independent T-variate normal distributions with 0 means and covariance matrices Σ, and if (R_i, θ_i) is the polar representation of (X_i, Y_i), then the conditional expectation of R_tR_{t-j} given $\theta_1, \ldots, \theta_T$ is

$$E[R_tR_{t-j}|\theta_1, \ldots, \theta_T] = \frac{\int_0^\infty \cdots \int_0^\infty r_t r_{t-j} \prod_{s=1}^T r_s \exp(-\tfrac{1}{2}r'Ar) \mathrm{d}r_1 \ldots \mathrm{d}r_T}{\int_0^\infty \cdots \int_0^\infty \prod_{s=1}^T r_s \exp(-\tfrac{1}{2}r'Ar) \mathrm{d}r_1 \ldots \mathrm{d}r_T} \tag{4.5}$$

where $A=(a_{ij})$, $a_{ij}=\sigma^{ij}\cos(\theta_i-\theta_j)$ and $\Sigma^{-1}=(\sigma^{ij})$. Thus to evaluate the conditional expectation (4.1) we need to evaluate equation (4.5) for the matrix Σ representing the covariance of the appropriate ARMA model. As for the wrapped normal model, the amount of computation required to do this is prohibitive. However, for low order AR processes we can approximate $E[R_t R_{t-j} | \theta_1, \ldots, \theta_T]$ by $E[R_t R_{t-j} | \theta_t \theta_{t-j}]$. In the case $p=1$, for example, we can approximate $E[R_t R_{t-1} | \theta_1, \ldots, \theta_T]$ by

$$\frac{\int_0^\infty \int_0^\infty r_1^2 r_2^2 \exp\left[-\frac{1}{2\sigma^2} \{r_1^2 + r_2^2 - 2\phi_1 r_1 r_2 \cos(\theta_t - \theta_{t-1})\}\right] \mathrm{d}r_1 \, \mathrm{d}r_2}{\int_0^\infty \int_0^\infty r_1 r_2 \exp\left[-\frac{1}{2\sigma^2} \{r_1^2 + r_2^2 - 2\phi_1 r_1 r_2 \cos(\theta_t - \theta_{t-1})\}\right] \mathrm{d}r_1 \, \mathrm{d}r_2}.$$

To evaluate the integrals, we expand the term involving $r_1 r_2$ in the exponential and obtain for the integral in the numerator

$$2\sigma^6 \sum_{n=0}^{\infty} \frac{\Gamma^2\{(n+3)/2\}}{n!} \{2\phi_1 \cos(\theta_t - \theta_{t-1})\}^n$$

with a similar expression for the integral in the denominator.

4.3. *Direct and Inverse Models*

Fitting the direct model can be accomplished with standard software once the mean has been estimated, say by the sample mean direction of the series. For a more refined analysis, we may estimate the parameters of the direct model by using Gaussian likelihoods or robust regression for a fixed value of μ. The resulting maximum value of the likelihood (or minimum value of the robust criterion) is thus a function of μ, and we may estimate μ by the value that maximizes this function. This yields maximum likelihood estimates. Standard errors can be obtained by numerical differentiation of the log-likelihood.

To fit AR(p) processes by using the inverse form of the model, we need to assume a marginal distribution for $\Theta_1, \ldots, \Theta_p$. We have taken this to be the product of independent VM(μ, κ) distributions to avoid introducing another parameter. This assumption does not seem to affect the estimation procedure provided that the series is not too short. The joint density of $\theta_1, \ldots, \theta_T$ is then

$$\prod_{t=p+1}^{T} f(\theta_t - \mu - g[\omega_1 g^{-1}\{(\theta_{t-1} - \mu)/2\} \ldots + \omega_p g^{-1}\{(\theta_{t-p} - \mu)/2\}]) \prod_{t=1}^{p} f(\theta_t - \mu)$$

where f is the VM(0, κ) density The fitting then proceeds by a routine application of the Newton–Raphson algorithm.

5. EXAMPLE

We shall illustrate the fitting of the inverse and linked time series models by studying a time series of 72 wind directions, comprising hourly measurements for 4 days at a site on Black Mountain, Australian Capital Territory, Australia. The data were collected as part of an experiment to calibrate three anemometers. Further information about the experiment can be found in Cameron (1983).

We begin by looking at the circular autocorrelation function, which is shown in Fig. 2(a). Comparison with Fig. 1 suggests that an IAR(1) model may be appropriate. To use the indirect method, assume that the conditional distributions of $\Theta_t|\theta_{t-1}$ are von Mises with concentration parameter κ. Parameter estimates and standard errors can be obtained as described in Section 4.3. Starting values μ and κ in the model fitting algorithm can be obtained from treating the data as a circular sample and a starting value for ω can be estimated from one of the prototype circular autocorrelation function graphs, giving

$$\hat{\mu}_0 = 291.2°, \qquad \hat{\kappa}_0 = 1.9, \qquad \hat{\omega}_0 = 0.5.$$

Maximizing the likelihood, we obtain the parameter estimates

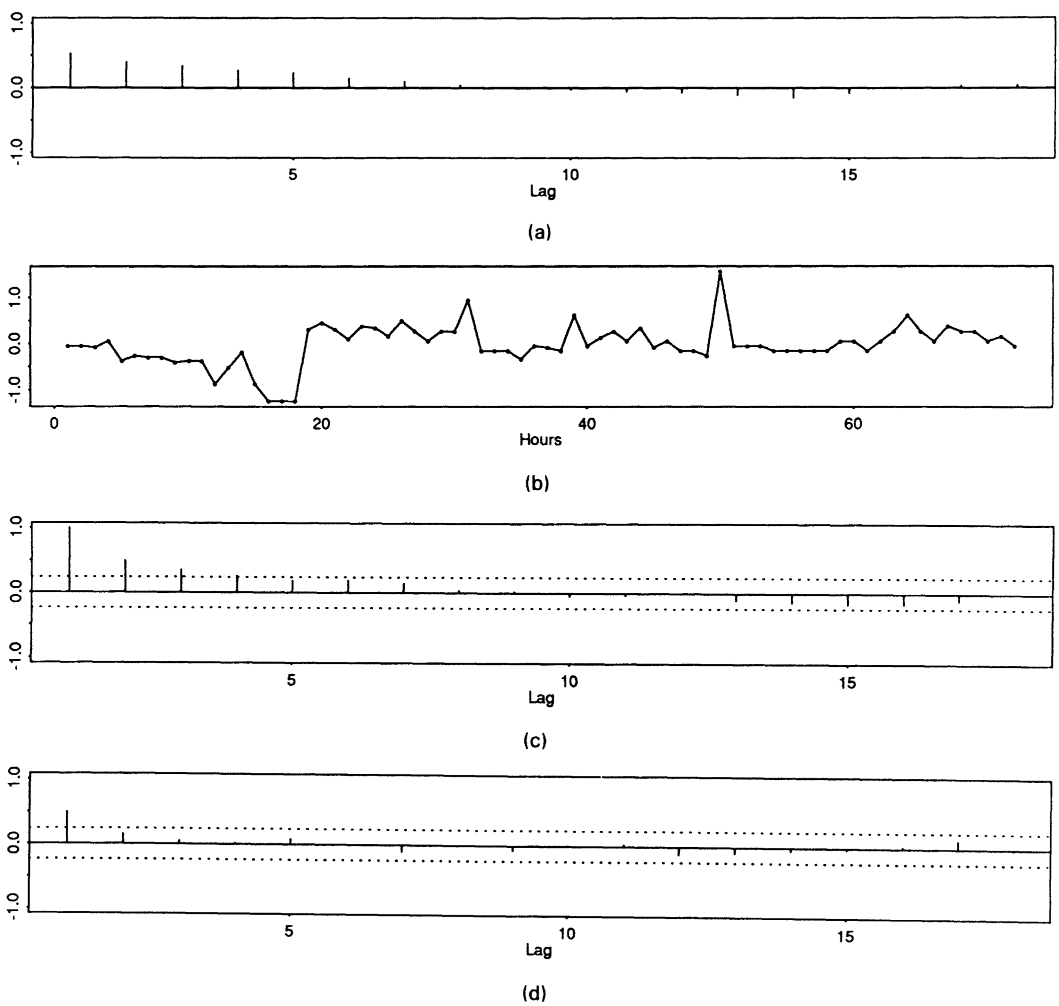

Fig. 2. Analysis of the series of 72 wind direction measurements, using a probit link to fit an LAR(1) model (the autocorrelation and partial autocorrelation functions suggest that an LAR(1) model may be reasonable): (a) circular autocorrelation function; (b) linked linear time series; (c) autocorrelation function for linked series; (d) partial autocorrelation function for linked series (the dotted lines in (c) and (d) are an indication about which correlations are significantly different from 0)

$$\hat{\mu} = 289.5°, \qquad \hat{\kappa} = 2.467, \qquad \hat{\omega} = 0.678$$

with standard errors 13.8°, 0.352 and 0.138 respectively. The covariances between the estimates are negligible.

Turning to the probit link model, consider fitting an AR(1) model to the transformed data

$$Y_t(\mu) = \Phi^{-1}\left(\frac{\theta_t - \mu}{2\pi} + \frac{1}{2}\right).$$

Graphs of the linked linear time series and the associated autocorrelation and partial autocorrelation functions are shown in Figs 2(b), 2(c) and 2(d) respectively. Again the evidence is in favour of a first-order autoregressive model. The resulting likelihood function is a function of μ. Maximizing it over μ yields the estimates

$$\hat{\mu}_0 = 297.2, \qquad \phi = 0.52, \qquad \sigma^2 = 0.146.$$

Note that the parameters μ in the two forms of the model have different definitions and are not directly comparable.

ACKNOWLEDGEMENTS

We are most grateful to Professor Tom Downs for supplying us with a preprint of his report on angular–angular regression, and to Dr Murray Cameron for supplying the data. Valuable comments by the Editor and two referees have led to a considerable clarification of the properties of the models.

APPENDIX A

In this appendix we collect a few technical details. We begin by giving formulae for the derivatives of the log-likelihood for the inverse model. Let $G(w, \theta)$ denote the function

$$G(w, \theta) = g\{w g^{-1}(\theta)\}$$

and let G_w, G_θ etc. denote its partial derivatives. The log-likelihood for the inverse AR(1) model of Section 2.3 is

$$l = -n \log I_0(\kappa) + \kappa \left[\sum_{t=2}^{n} \cos\{\theta_t - \mu - G(w, \theta_{t-1} - \mu)\} + \cos(\theta_1 - \mu) \right]$$

where I_0 is the modified Bessel function of order 0. The derivatives are (writing $c_t = \cos\{\theta_t - \mu - G(w, \theta_{t-1} - \mu)\}$ and $s_t = \sin\{\theta_t - \mu - G(w, \theta_{t-1} - \mu)\}$)

$$\frac{\partial l}{\partial \kappa} = -n A(\kappa) + \sum_{t=2}^{n} c_t + \cos(\theta_1 - \mu),$$

$$\frac{\partial l}{\partial \mu} = \kappa \left\{ \sum_{t=2}^{n} s_t (1 - G_\theta^t) + \sin(\theta_1 - \mu) \right\},$$

$$\frac{\partial l}{\mathrm{d} w} = \kappa \sum_{t=2}^{n} s_t G_w^t,$$

$$\frac{\partial^2 l}{\partial \kappa^2} = -n A'(\kappa),$$

$$\frac{\partial^2 l}{\partial\kappa\,\partial\mu} = \sum_{t=2}^{n} s_t(1-G_\theta^t) + \sin(\theta_1-\mu),$$

$$\frac{\partial^2 l}{\partial\kappa\,\partial w} = \sum_{t=2}^{n} s_t G_w^t,$$

$$\frac{\partial^2 l}{\partial\mu^2} = \kappa\left[\sum_{t=2}^{n} \{s_t G_{\theta\theta}^t - c_t(1-G_\theta^t)^2\} - \cos(\theta_1-\mu)\right],$$

$$\frac{\partial^2 l}{\partial\mu\,\partial w} = -\kappa\left[\sum_{t=2}^{n} \{c_t(1-G_\theta^t)G_w^t + s_t G_{w\theta}^t\}\right],$$

$$\frac{\partial^2 l}{\partial w^2} = \kappa\left[\sum_{t=2}^{n} \{s_t G_{ww}^t - c_t(G_w^t)^2\}\right]$$

where, for example, $G_\theta^t = G_\theta(w, \theta_{t-1}-\mu)$.

We next prove the lemma of Section 2. Let (R_i, Θ_i) be the polar representation of (X_i, Y_i). The marginal distributions of Θ_1 and Θ_2 are uniform so we can write ρ_T as

$$4(E[\sin\Theta_1\sin\Theta_2]\,E[\cos\Theta_1\cos\Theta_2] - E[\sin\Theta_1\cos\Theta_2]\,E[\sin\Theta_1\cos\Theta_2]).$$

We need to evaluate the expectations in this expression. The joint density of $(R_1, R_2, \Theta_1, \Theta_2)$ is

$$Cr_1 r_2 \exp\left[-\frac{1}{2(1-\rho^2)}\{r_1^2+r_2^2-2\rho r_1 r_2\cos(\theta_1-\theta_2)\}\right]$$

where

$$C^{-1} = (2\pi)^2(1-\rho^2),$$

so that for example $E[\cos\Theta_1\cos\Theta_2]$ is given by

$$C\sum_{n=0}^{\infty}\frac{1}{n!}\int_0^\infty r_1^{n+1}\exp\left\{-\frac{r_1^2}{2(1-\rho^2)}\right\}\int_0^\infty r_2^{n+1}\exp\left\{-\frac{r_2^2}{2(1-\rho^2)}\right\}$$
$$\times\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\cos\theta_1\cos\theta_2\left\{\frac{\rho}{1-\rho^2}\cos(\theta_1-\theta_2)\right\}^n \mathrm{d}r_1\,\mathrm{d}r_2\,\mathrm{d}\theta_1\,\mathrm{d}\theta_2.$$

Using the facts that

$$\int_0^\infty r^{n+1}\exp\left\{-\frac{r^2}{2(1-\rho^2)}\right\}\mathrm{d}r = \frac{1}{2}\{2(1-\rho^2)\}^{(n+2)/2}\Gamma\left(\frac{n+2}{2}\right)$$

and

$$\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\cos\theta_1\cos\theta_2\cos^n(\theta_1-\theta_2)\,\mathrm{d}\theta_1\,\mathrm{d}\theta_2 = \begin{cases} 0, & n \text{ even} \\ \dfrac{\pi^2}{2^{2m}}\dbinom{2m+1}{m}, & n \text{ odd}, \ n=2m+1, \end{cases}$$

the series above can be written after some simplification as

$$\frac{1}{2}\rho(1-\rho^2)\sum_{m=0}^{\infty}\frac{\Gamma^2(m+\frac{3}{2})\rho^{2m}}{\Gamma(m+2)m!} = \frac{\pi}{8}\rho(1-\rho^2)\,{}_2F_1\left(\frac{3}{2},\frac{3}{2},2;\rho^2\right).$$

Similarly, we can show that $E[\sin\Theta_1 \sin\Theta_2]$ also has this value and that the other expectations are 0. Hence

$$\rho_T = \frac{\pi^2}{16}\rho^2(1-\rho^2)^2\left\{{}_2F_1\left(\frac{3}{2}, \frac{3}{2}, 2; \rho^2\right)\right\}^2.$$

REFERENCES

Accardi, L., Cabrera, J. and Watson, G. S. (1987) Some stationary Markov processes in discrete time for unit vectors. *Metron*, **45**, 115–133.

Breckling, J. (1989) *The Analysis of Directional Time Series: Applications to Wind Speed and Direction*. London: Springer.

Cameron, M. A. (1983) The comparison of time series recorders. *Technometrics*, **25**, 9–22.

Downs, T. D. (1986) Projective angular regression. Unpublished.

Fisher, N. I. and Lee, A. J. (1983) A correlation coefficient for circular data. *Biometrika*, **70**, 327–332.

——— (1986) Correlation coefficients for random variables on the sphere or hypersphere. *Biometrika*, **73**, 159–164.

——— (1992) Regression models for an angular response. *Biometrics*, **48**, 665–677.

Gonella, J. (1972) A rotary-component method for analysing meteorological and oceanographic time series. *Deep Sea Res.*, **19**, 833–846.

Watson, G. S. and Beran, R. J. (1967) Testing a sequence of unit vectors for serial correlation. *J. Geophys. Res.*, **72**, 5655–5659.

Wehrly, T. E. and Johnson, R. A. (1979) Bivariate models for dependence of angular observations and a related Markov process. *Biometrika*, **66**, 255–256.