Ann Inst Stat Math (2008) 60:781–800
DOI 10.1007/s10463-008-0207-z

An angular–linear time series model for waveheight prediction

Tsukasa Hokimoto · Kunio Shimizu

Received: 4 January 2008 / Revised: 13 August 2008 / Published online: 1 October 2008

Abstract The modeling of the dynamic relationship between the changes of the wind and the waveheight has been an important topic from various standpoints such as oceanography, technology and navigation safety. Generally, when we apply the standard statistical models for the waveheight prediction, the observation of wind direction has been treated as the ordinary time series data, not reflecting unique properties as directional data. In this article, we develop a time series model with linear and angular–linear variables, by extending the angular–linear regression model considered by Johnson and Wehrly. Our prediction test based on extrapolation suggested the possibility that the angular–linear time series structure gave positive effect on improving the prediction accuracy of the time series model, in which the original wind direction is included as a linear variable.

Keywords Directional time series model · Angular–linear regression · ARIMA model · Prediction · Waveheight · Wind direction

1 Introduction

The modeling and prediction of the sea condition and the wind motion is of importance not only in estimating the safety of the ship navigation, but also in considering various

T. Hokimoto (✉)
Department of Economics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku,
Tokyo 113-0033, Japan
e-mail: hokimoto@e.u-tokyo.ac.jp

K. Shimizu
Department of Mathematics, Keio University, 3-14-1 Hiyoshi, Kohoku-ku,
Yokohama 223-8522, Japan
e-mail: shimizu@stat.math.keio.ac.jp

physical phenomena on the sea, such as the motion of floating objects and the diffusion of oil pollution. In the last few decades, the statistical methodologies to predict the change of sea condition, based on the observations measured on the research ship or buoy, have been researched from the various aspects.

Traditional researches on the prediction of the sea condition have been mainly focused on the statistical modeling of the changes of the physical factors, such as the significant waveheight, wave period, wind speed and wind direction. In general, it is not necessarily reasonable to assume that the original process of the sea surface movement follows Gaussian process. However, it is often possible to regard the time series to follow Gaussian process, by applying Box–Cox or Rosenblatt transformation to the original time series of the sea surface movement. For the time series data of this class, the linear stationary time series models proposed by Box and Jenkins (1976), such as autoregressive (AR) model or autoregressive moving average (ARMA) model, have been widely used to construct the predictor. Under the background that various methodologies related to their modeling have been developed as well as their statistical properties (for example, Brockwell and Davis 1991), many applications to the wind data (Brown et al. 1984; Daniel and Chen 1991), and the waveheight data (Cunha and Guedes 1999; Yim et al. 2002) have been reported by many authors. However, as for the observed time series data in the nonstationary aspect, such as the wave development process, these stationary models cannot give us reasonable interpretations.

The researches on the nonstationary time series model to predict the change of the observation in the exogenous aspect of the sea state is classified into two categories. One is to apply the standard linear nonstationary models, and the other is to develop the original model with nonstationary structure. As for the former, there are literatures on the various applications. Among them, ARIMA model proposed by Box and Jenkins (1976) is one of the widely used models in the practical aspect. The model is identified by fitting a stationary ARMA model to the differenced series of the original time series data. Also, the autoregressive model with time varying coefficients (Kitagawa and Gersch 1985) is also one of the widely used nonstationary models. Since this model has the state space representation, the predicted values can be obtained by applying Kalman filter. If the speed in changing statistical structure can be regarded to be slow, it may be reasonable to fit a stationary AR model to the time series data in the segment which can be regarded to be locally stationary (the theoretical background is given in Dahlhaus 1997; Dahlhaus and Giraitis 1998, and so forth). Hokimoto et al. (2003) developed the nonstationary spectral model of the sea surface movement in the wave development process, based on the idea of local stationarity. We can also find the application of generalized autoregressive conditional heteroscedascity (GARCH) model to the wind speed observations (Toll 1997). On the other hand, as for the latter, we can find some nonstationary time series models based on the decomposition of the trend and the other components have been proposed (Athanassoulis and Stefanakos 1995; Walton and Borgman 1990; Stefanakos et al. 2002).

When we apply these models in the practical aspect, the observation of the wind direction has been usually treated as general time series data, just the same as those of the wind speed or the sea surface movement. Angular data have a unique property that they take the values on the circle. It is expected, therefore, to be able to express the interaction relationship between the wind motion and the sea surface movement

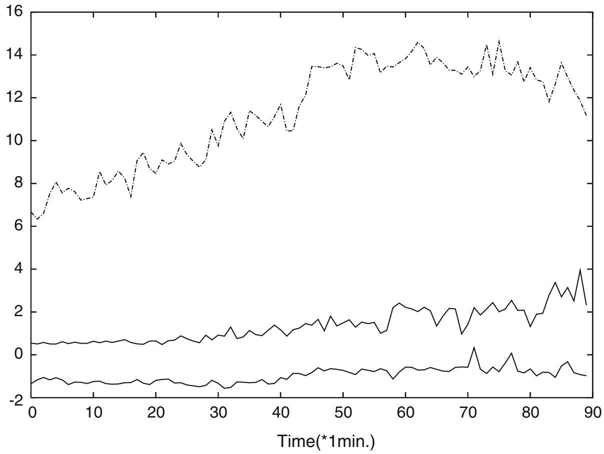

Fig. 1 Changes of the measured data per 1 min (from the top, wind speed, m/s), 1/3 significant waveheight (m), wind direction, rad.)

more reasonably, by reflecting this unique property. In the framework of directional statistics, various methodologies for statistical inferences on the directional data have been proposed (for example, Marida and Jupp 2000). Among them, multivariate regression models, including circular and linear variables, have been often proposed in environment studies. Johnson and Wehrly (1978) considered the theoretical background of the linear parametric regression, which has the linear variable and the angular variable. And the extension of their model has given in Fisher and Lee (1992); SenGupta (2004); SenGupta and Ugwuowo (2006), and so forth. However, there have been only limited attempts to model multivariate angular–linear data. In this paper, we are interested in the modeling of the dynamic dependencies on the multivariate linear variables and the circular variables. Our goal here is to develop an angular–linear time series model to express the dynamic structure among the waveheight, the wind speed and the wind direction, by extending the multiple regression model by Johnson and Wehrly (1978), and to show the effectiveness of the model structure through the test of prediction based on extrapolation (for example, Makridakis et al. 1984, Chapter 4).

Our motivation of this research is based on the measured observations of the sea surface movement and the wind motion in the wave development process. We have measured the simultaneous changes of the relative sea surface level, the wind speed and wind direction, on a research ship at Hunka-bay, Hokkaido, Japan. For the relative sea surface movement, we measured relative displacement from the mean of the sea surface movement for 10 min by an ultrasonic waveheight meter of the research ship. Also, for the changes of the wind speed and wind direction, we measured their changes at about 15 m height from the sea surface, by using an ultrasonic wind meter.

In this paper, we focus on the analysis of the changes of the mean values for every 1 min on the wind speed, the wind direction and the waveheight, which is defined as the mean value of the one-third of the amplitudes which were sorted from the maximum. Figure 1 shows a record of their changes for every 1 min, measured on 2

December, 1999. It is noted that the horizontal axis denotes the time in minutes, a dotted line denotes the wind speed (m/s), a solid line denotes the wind direction (rad.) and a bold line denotes the 1/3 significant waveheight (m), where each sample size is 90. According to the weather maps of the sampling day, as well as the days before and after, the location of atmospheric pressure formed typical pattern of winter in Japan; the high pressure area is extending over the west of Japan Islands and the low pressure area is extending over the south. Under the background of this location, the observation in Fig. 1 showed the tendency that the wind direction changed slowly from north-west to true north, and the wind speed rapidly increased approximately from 6 to 13 m/s in 40–50 min, and then changed slowly in the range approximately from 12 to 15 m/s. Also, the waveheight gradually grew up to 2.5 m under the condition that the wind direction changed slowly and the wind speed increased. The record of Fig. 1 can be regarded as a typical measurement of the wave development process, because the waveheight increases, under the situation that the wind speed becomes faster and the wind direction does not change so much.

This paper is composed as follows. In the next section, we make preliminary analyses on the correlation structure of the measured data. In Sect. 3, we propose a multivariate time series model as an extension of the model by Johnson and Wehrly (1978). The effect on prediction based on the model is examined by numerical experiment in Sect. 4, and the results are summarized to conclude in the final section.

2 Correlation structures of the wind and the waveheight data

In this section, we make basic investigations on the correlation structure of the observation shown in Fig. 1, in order to consider the class for our model. We first analyze the sample circular autocorrelation and the sample autocorrelation of the wind direction time series data. And then, we investigate the cross correlation between the observed variables.

2.1 Circular autocorrelation of the wind direction data

Let $\{\mathrm{WD}_t; t = 1, \ldots, N\}$ be a set of measurements of wind direction, where N is the sample size and t is the time point. As a basic concept of exploratory circular data analysis, we refer to a book by Fisher (1993, Chapter 2) and use the following two transformations of WD_t

$$x_t = \cos(\mathrm{WD}_t), \quad y_t = \sin(\mathrm{WD}_t)$$

In order to explore the possibility of detecting changes of direction, we use two statistics; one is the cumulative sum (CUSUM) plot displayed by the points

$$C_t = \sum_{i=1}^{t} x_i, \quad S_t = \sum_{i=1}^{t} y_i$$

and the other is the cumulative mean direction plot $\{\Theta_t^c; t = 1, \ldots\}$, such that

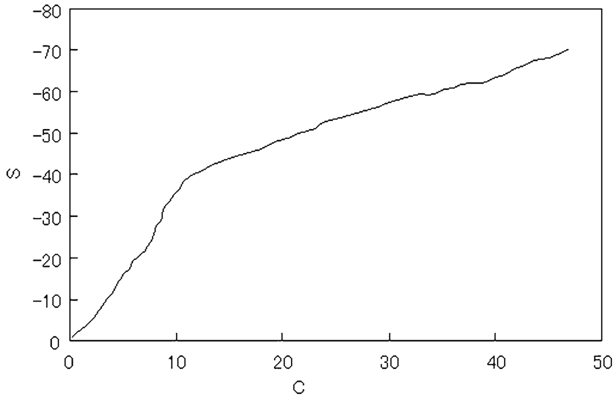

Fig. 2 Cumulative sum plot

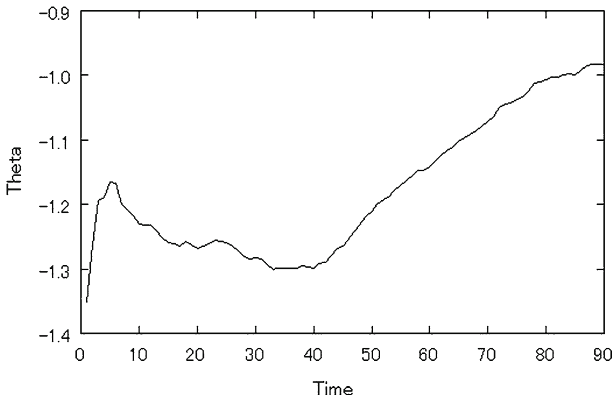

Fig. 3 Cumulative mean directional plot

$$\cos(\Theta_t^c) = C_t \Big/ \sqrt{C_t^2 + S_t^2}, \quad \sin(\Theta_t^c) = S_t \Big/ \sqrt{C_t^2 + S_t^2} \tag{1}$$

are satisfied simultaneously. CUSUM plot is shown in Fig. 2, where the horizontal axis denotes C_t and the vertical axis denotes S_t. Also, the cumulative mean directional plot is displayed in Fig. 3, where the horizontal axis denotes the time point t and the vertical axis denotes Θ_t^c. It is noted that the change in statistical structure of the directional time series data is admitted, when the trend of plots in Fig. 2 is clearly different from the straight line whose slope is one, and when the value of $\{\mathrm{WD}_t^c\}$ in Fig. 3 is clearly different from the constant value. Figure 3 suggests the possibility that the directional time series data have a change point of statistical structure at $t = 40$ roughly, and in this case the time series exhibits nonstationarity. We also checked the statistical test of change in mean direction by using CircStats (Chapter 11 of Jammalamadaka and SenGupta 2001). The result showed that there exists a change point at the time point $t = 42$, which suggested that the data exhibit nonstationarity.

Now we are interested in whether there is clear difference in the correlation structures, between the cases when we regard the wind direction data to be circular data and linear time series data. Before the calculation of correlation, it is necessary to subtract the trend of the data. For estimating trend, we obtain the smoothed series of $\{x_t\}$ and $\{y_t\}$ by using the locally weighted regression (LOWESS). And then, based on the smoothed series, say $\{x_t^*\}$ and $\{y_t^*\}$, we obtain the smoothed trend of wind direction T_t^*, such that

$$\frac{x_t^*}{\sqrt{(x_t^*)^2 + (y_t^*)^2}} = \cos(T_t^*), \quad \frac{y_t^*}{\sqrt{(x_t^*)^2 + (y_t^*)^2}} = \sin(T_t^*) \tag{2}$$

are satisfied simultaneously. On the other hand, we also estimate the trend of $\{\mathrm{WD}_t\}$ by applying the trend model,

$$\mathrm{WD}_t = T_t^{**} + \zeta_t, \quad \zeta_t \sim N(0, \sigma_\zeta^2)$$

and

$$T_t^{**} - T_{t-1}^{**} = v_t, \quad v_t \sim N(0, \sigma_v^2)$$

where T_t^{**} is the random variable to express the trend, σ_ζ^2 and σ_v^2 are unknown variances of ζ_t and v_t, respectively (The detail of this model is shown in Kitagawa and Gersch 1985). Figure 4 shows the changes of $\{T_t^*\}$ (dotted curve) and $\{T_t^{**}\}$ (bold curve). It looks that there is no clear difference between $\{T_t^*\}$ and $\{T_t^{**}\}$. So we estimate circular autocorrelation coefficient based on the subtracted series, $\mathrm{WD}_t^* \equiv \mathrm{WD}_t - T_t^*$. Based on circular–circular association (Fisher 1993, Chapter 6), the sample circular autocorrelation coefficient is given by

$$\hat{\rho}^*(\tau) = \frac{4(A_\tau B_\tau - C_\tau D_\tau)}{\left[(N^2 - E_\tau^2 - F_\tau^2)(N^2 - G_\tau^2 - H_\tau^2)\right]^{1/2}}, \quad \tau = 0, 1, \ldots$$

where τ is the time lag, and

$$\begin{aligned}
A_\tau &= \sum_{t=1}^{N-\tau} \cos \mathrm{WD}_t^* \cos \mathrm{WD}_{t+\tau}^*, \quad B_\tau = \sum_{t=1}^{N-\tau} \sin \mathrm{WD}_t^* \sin \mathrm{WD}_{t+\tau}^*,\\
C_\tau &= \sum_{t=1}^{N-\tau} \cos \mathrm{WD}_t^* \sin \mathrm{WD}_{t+\tau}^*,\\
D_\tau &= \sum_{t=1}^{N-\tau} \sin \mathrm{WD}_t^* \cos \mathrm{WD}_{t+\tau}^*, \quad E_\tau = \sum_{t=1}^{N-\tau} \cos(2\mathrm{WD}_t^*),\\
F_\tau &= \sum_{t=1}^{N-\tau} \sin(2\mathrm{WD}_t^*), \quad G_\tau = \sum_{t=1}^{N-\tau} \cos(2\mathrm{WD}_{t+\tau}^*),\\
H_\tau &= \sum_{t=1}^{N-\tau} \sin(2\mathrm{WD}_{t+\tau}^*).
\end{aligned}$$

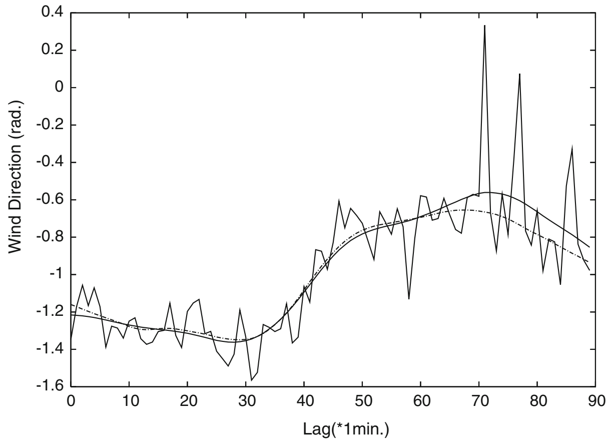

Fig. 4 An example on the estimation of trend

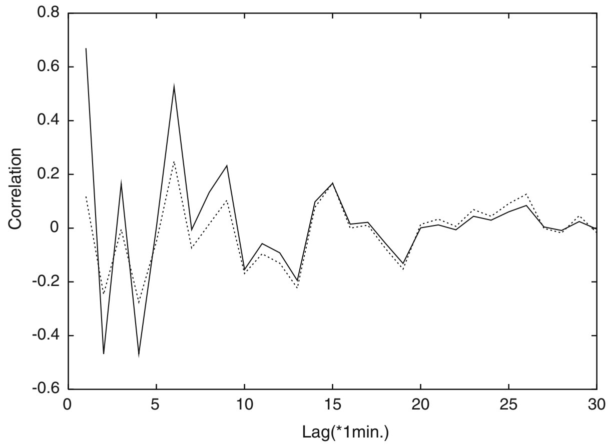

Fig. 5 Sample circular autocorrelation and sample autocorrelation

On the other hand, the sample autocorrelation function of the time series data WD_t^* is given by

$$\hat{\rho}^{**}(\tau) = \frac{\sum_{t=1}^{N-\tau}(\mathrm{WD}_{t+\tau}^* - \overline{\mathrm{WD}^*})(\mathrm{WD}_t^* - \overline{\mathrm{WD}^*})}{\sum_{t=1}^{N}(\mathrm{WD}_t^* - \overline{\mathrm{WD}^*})^2}, \quad \overline{\mathrm{WD}^*} = \frac{1}{N}\sum_{t=1}^{N}\mathrm{WD}_t^*.$$

Figure 5 displays the estimates of $\hat{\rho}^*(\tau)$ and $\hat{\rho}^{**}(\tau)$ $(0 \leq \tau \leq 30)$, where the vertical axis denotes the correlation, the horizontal axis denotes τ in minutes, and the bold and dotted lines correspond to $\hat{\rho}^*(\tau)$ and $\hat{\rho}^{**}(\tau)$, respectively. We observe that they change similarly with the same tendency, although $|\hat{\rho}^*(\tau)|$ takes slightly larger values than $|\hat{\rho}^{**}(\tau)|$ when τ is small. We can evaluate from this result that the sample

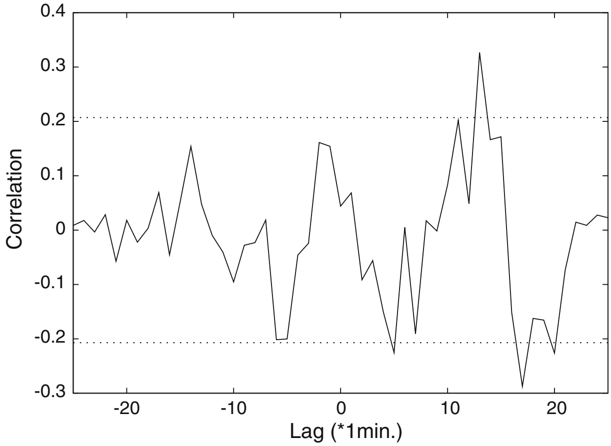

Fig. 6 Cross correlation between $\{WS_t\}$ and $\{WL_t\}$

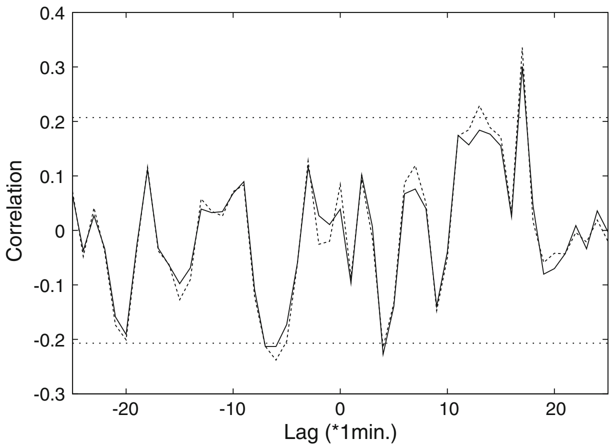

Fig. 7 Cross correlations between (1) $\{WD_t\}$ and $\{WL_t\}$, (2) $\{\sin(WD_t)\}$ and $\{WL_t\}$

circular autocorrelation coefficient can be approximated by the linear correlation to some extent.

2.2 Cross correlation among observed variables

Next, we investigate the features on the cross correlations among the wind direction WD_t, the wind speed WS_t, and the waveheight WL_t. We estimated the cross correlation function among the variables WL_t, WS_t, WD_t and $\sin(WD_t)$, based on the time series data after subtraction of their trends estimated by LOWESS method. Figure 6 shows the cross correlation between WS_t and WL_t and Fig. 7 displays the cross correlations between (1) WD_t and WL_t (dotted line) and (2) $\sin(WD_t)$ and WL_t (bold line), where the horizontal axis is the time lag in minutes and the two parallel lines in each figure

Table 1 Estimates of the cross correlations

Lag(s)	WD_t	$\sin(WD_t)$	$WS_t\ \sin(WD_t)$
0	0.038	0.084	0.063
1	−0.096	−0.085	−0.110
2	0.102	0.093	0.098
3	0.006	−0.014	−0.017
4	−0.227	−0.214	−0.194
5	−0.139	−0.133	−0.092
6	0.067	0.087	0.073
7	0.076	0.119	0.150
8	0.040	0.050	0.034
9	−0.139	−0.147	−0.153
10	−0.039	−0.050	−0.071
11	0.174	0.173	0.125
12	0.156	0.184	0.178
13	0.183	0.228	0.158
14	0.176	0.188	0.153
15	0.155	0.171	0.126
16	0.025	0.032	0.060
17	0.300	0.335	0.388
18	0.046	0.016	0.049
19	−0.080	−0.058	−0.024
20	−0.070	−0.042	0.014

denote Bartlett's bounds (i.e., $\pm 1.96N^{-1/2}$). It is noted that the positive time lag can be interpreted as the time which is necessary to give impact on the change of the waveheight. Figure 6 suggests the possibility that the change of wind speed affects the waveheight after 10–20 min. Also, Fig. 7 shows the possibilities that (1) the change of wind direction (and its sine transformation) affect the waveheight after 15–20 min and (2) the cross correlation between $\sin(WD_t)$ and WL_t tends to become higher than does that between WD_t and WL_t.

The interaction between the wind speed and the wind direction may be expected, too. Based on the above results, the cross correlations between the variables $\sin(WD_t)$ and $\cos(WD)$ and WL_t take higher values than that between WD_t and WL_t. Under the interaction, we may expect that the variables $WS_t\ \sin(WD_t)$ and $WS_t\ \cos(WD_t)$ and WL_t will exhibit higher cross correlation than that between WD_t and WL_t. Table 1 shows the estimation result on the cross correlations between WL_t and (1) WD_t, (2) $\sin(WD_t)$ and (3) $WS_t\ \sin(WD_t)$. We observe that $WS_t\ \sin(WD_t)$ gives higher cross correlation than $\sin(WD_t)$ when the time lag is 17, although we cannot find strong correlation when the time lag is small up to about 15 min. As a result, it is also expected to improve the prediction accuracy by using the variables transformed by sine and cosine functions such as $\sin(WD_t)$ and $WS_t\ \sin(WD_t)$ as explanatory variables, instead of using the original WD_t.

3 An angular–linear time series model for waveheight prediction

In this section, we develop a time series model taking account of angular–linear variables of wind direction. Our preliminary analysis in Sect. 2 has suggested that there is no clear difference between the structures of the circular autocorrelation and the standard linear autocorrelation, and that we may expect to obtain larger correlation by using the wind speed transformed by sine and cosine functions as variables. Based on these results, we develop the model in the class of linear time series.

3.1 The linear regression including angular change as explanatory variables

We describe our theoretical motivation on how we construct a general linear regression model for the above situation, based on the result of Johnson and Wehrly (1978). Let $\Theta = (\Theta_1, \ldots, \Theta_p)'$ be a p-dimensional angular variable and $X = (X_1, \ldots, X_q)'$ be a q-dimensional linear variable (throughout this paper, the symbol $'$ indicates the transposition of vectors or matrices). Also put

$$H(\Theta) = \begin{pmatrix} \cos\Theta_1 & \cdots & \cos n\Theta_1 & \sin\Theta_1 & \cdots & \sin n\Theta_1 \\ \vdots & & & & & \\ \cos\Theta_p & \cdots & \cos n\Theta_p & \sin\Theta_p & \cdots & \sin n\Theta_p \end{pmatrix}.$$

Johnson and Wehrly (1978) gave the following result on a multivariate angular–linear distribution.

Theorem (Johnson and Wehrly, 1978)

Suppose (Θ, X) has the joint density function

$$f(\theta, x) = C \exp\left[-\frac{1}{2}x'\Sigma^{-1}x + \lambda'\Sigma^{-1}x + a(\theta)'\Sigma^{-1}x\right],$$

$$\theta = (\theta_1, \ldots, \theta_p)' \in [0, 2\pi)^p, x = (x_1, \ldots, x_q)' \in R^q$$

where C is a normalizing constant, Σ is a $q \times q$ positive definite dispersion matrix of X, $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_q)'$ is a constant vector and $a(\boldsymbol{\theta}) = (a_1(\boldsymbol{\theta}), \ldots, a_q(\boldsymbol{\theta}))'$ such that

$$a_i(\theta) = \sum_{j=1}^{p}\sum_{k=1}^{n} a_{ijk}\cos(k(\theta_j - \mu_{ijk}))$$

$$= \sum_{j=1}^{p}\sum_{k=1}^{n} \left[\alpha_{ijk}\cos(k\theta_j) + \beta_{ijk}\sin(k\theta_j)\right], \quad i = 1, \ldots, q$$

where n is a positive integer, $(a_{ikj}, \alpha_{ijk}, \beta_{ijk})$ are constants, and $\mu_{ijk} \in [0, 2\pi)$. Then it maximizes theentropy of the multivariate angular–linear distribution under the condition that the values of $E(XX')$, $E(X)$, $E(X \otimes H(\Theta))$ are given, where $\otimes$ is the Kronecker product.

Let (Z, X, Θ) be random variables to express the change of the waveheight, wind speed and wind direction, respectively. Our interest here is on the distribution which Z follows. Based on this theorem, the conditional distribution of Z, given $(X, \Theta) = (x, \theta)$, is

$$N\left(\beta_0 + \beta_1 x + \sum_{k=1}^{n} [\beta_{2k}\cos(k\theta) + \beta_{3k}\sin(k\theta)]\,,\, \sigma_1^2(1-\rho^2)\right)$$

where $(\beta_{\cdot}, \beta_{\cdot\cdot})$ are constants and n is a positive integer. This suggests that we can use a linear regression model for waveheight,

$$Z = \beta_0 + \beta_1 X + \sum_{k=1}^{K} [\beta_{2k}\cos(k\Theta) + \beta_{3k}\sin(k\Theta)] + \delta, \tag{3}$$

where δ is a random variable which follows Normal distribution. In our analysis, however, the variables Z, X and Θ have time series structures. Therefore, the model (3) should be extended so that the time series structures of (Z, X, Θ) have been reflected.

3.2 Building a time series structure including angular–linear variables

Suppose that, based on the N samples $\{\mathrm{WL}_t, \mathrm{WS}_t, \mathrm{WD}_t\}$ $(t = 1, \ldots, N)$, we predict the future values of the waveheight $\{\mathrm{WL}_{N+l}; l = 1, \ldots, L\}$. We assume that the time series $\{\mathrm{WL}_t\}$, $\{\mathrm{WS}_t\}$ and $\{\mathrm{WD}_t\}$ are stationary, by applying a proper transformation (the detail is given in the next section). By analogy with (3), we write the model relating to WL_t as

$$\begin{aligned}\mathrm{WL}_t = m_L &+ \sum_{i=1}^{p} \beta_i^{(1)}\mathrm{WL}_{t-i} + \sum_{i=1}^{p}\sum_{k=1}^{K} \beta_{i,k}^{(3)} \cos(k \cdot \mathrm{WD}_{t-i}) \\ &+ \sum_{i=1}^{p}\sum_{k=1}^{K} \beta_{i,k}^{(4)} \sin(k \cdot \mathrm{WD}_{t-i}) \\ &+ \sum_{i=1}^{p} \beta_i^{(2)}\mathrm{WS}_{t-i} + \varepsilon_t^{(1)}, \quad \varepsilon_t^{(1)} \sim \mathrm{WN}(0, \sigma_{\mathrm{WL}}^2)\end{aligned}$$

where p and K are orders, m_L is the unknown mean, β's are unknown weights, and $\varepsilon_t^{(1)}$ is a random variable, which follows a white noise process. In the same manner, we write WS_t as

$$\begin{aligned}\mathrm{WS}_t &= m_S + \sum_{i=1}^{p} \gamma_i^{(1)} \mathrm{WL}_{t-i} + \sum_{i=1}^{p}\sum_{k=1}^{K} \gamma_{i,k}^{(3)} \cos(k \cdot \mathrm{WD}_{t-i}) \\ &\quad + \sum_{i=1}^{p}\sum_{k=1}^{K} \gamma_{i,k}^{(4)} \sin(k \cdot \mathrm{WD}_{t-i}) \\ &\quad + \sum_{i=1}^{p} \gamma_i^{(2)} \mathrm{WS}_{t-i} + \varepsilon_t^{(2)}, \quad \varepsilon_t^{(2)} \sim \mathrm{WN}(0, \sigma_{\mathrm{WN}}^2)\end{aligned}$$

and $\sin(h \cdot \mathrm{WD}_t)$ and $\cos(h \cdot \mathrm{WD}_t)$ $(h = 1, \ldots, K)$ as

$$\begin{aligned}\sin(h \cdot \mathrm{WD}_t) &= m_h + \sum_{i=1}^{p} \delta_i^{(1)} \mathrm{WL}_{t-i} + \sum_{i=1}^{p}\sum_{k=1}^{K} \delta_{i,k}^{(3)} \cos(k \cdot \mathrm{WD}_{t-i}) \\ &\quad + \sum_{i=1}^{p}\sum_{k=1}^{K} \delta_{i,k}^{(4)} \sin(k \cdot \mathrm{WD}_{t-i}) \\ &\quad + \sum_{i=1}^{p} \delta_i^{(2)} \mathrm{WS}_{t-i} + \delta_t^{(h)}, \quad \delta_t^{(h)} \sim \mathrm{WN}(0, \sigma_h^2),\end{aligned}$$

and so forth, where m_s, m_h, γ's and δ's are unknown weights. Now put the state vector at time t by

$$\boldsymbol{y}_t^{(K)} \equiv (\mathrm{WL}_t, \mathrm{WS}_t, \cos(\mathrm{WD}_t), \sin(\mathrm{WD}_t), \ldots, \cos(K \cdot \mathrm{WD}_t), \sin(K \cdot \mathrm{WD}_t))' \tag{4}$$

Then we can write

$$\boldsymbol{y}_t^{(K)} = \boldsymbol{m}^{(K)} + A_1^{(K)} \boldsymbol{y}_{t-1}^{(K)} + \cdots + A_p^{(K)} \boldsymbol{y}_{t-p}^{(K)} + \boldsymbol{\delta}_t^{(K)}, \boldsymbol{\delta}_t^{(K)} \sim \mathrm{WN}(\boldsymbol{0}, \Sigma^{(K)}) \tag{5}$$

where $\boldsymbol{m}^{(K)}$ is the unknown mean vector, $A_i^{(K)}$ $(i = 1, \ldots, p)$ is the unknown coefficient matrix. This forms a multivariate vector autoregressive model of the pth order, and the estimates for elements of unknown matrices $A_i^{(K)}$ can be obtained by using the least squares method (e.g., Brockwell and Davis 1996). Thus, we can construct an l-step $(l = 1, \ldots, L)$ ahead predictor based on (5) by

$$\hat{\boldsymbol{y}}_{N+l}^{(K)} = \hat{\boldsymbol{m}}^{(K)} + \hat{A}_1^{(K)} \boldsymbol{z}_{N+l-1}^{(K)} + \hat{A}_2^{(K)} \boldsymbol{z}_{N+l-2}^{(K)} + \cdots + \hat{A}_p^{(K)} \boldsymbol{z}_{N+l-p}^{(K)} \tag{6}$$

and $\boldsymbol{z}_{N+l-m}^{(K)} = \boldsymbol{y}_{N+l-p}^{(K)}$ $(l \le p)$, $\boldsymbol{z}_{N+l-m}^{(K)} = \hat{\boldsymbol{y}}_{N+l-p}^{(K)}$ $(l > p)$, where $\hat{A}_i$ is the least squares estimator of A_i, The predicted values of WL_{N+l} $(l = 1, \ldots, L)$ can be obtained from the prediction of $\hat{\boldsymbol{y}}_{N+l}^{(K)}$.

3.3 Principal component to include the multiple directional information

However, the model (5) with (4) has a drawback in computational aspect. Since (5) involves $(2+2K)+p(2+2K)^2$ unknown parameters to be estimated, when both K and p become large, it is probable that the accuracy of the estimates of parameter becomes worse. For improving the prediction accuracies, the dimension of $\hat{\boldsymbol{y}}_t^{(K)}$ should be small. In order to taking account of the multiple directional information with the small numbers of variables, we focus on the following linear sum

$$\begin{aligned}\widetilde{\mathrm{WD}}_t^{(K)} \equiv\ & \omega_1 \cos(\mathrm{WD}_t) + \omega_2 \sin(\mathrm{WD}_t) + \cdots + \omega_{2K-1} \cos(K \cdot \mathrm{WD}_t) \\ & + \omega_{2K} \sin(K \cdot \mathrm{WD}_t)\end{aligned}$$

where ω_i $(i = 1, \ldots, 2K)$ are unknown weights. We propose to use the model (5) with the state vector

$$\tilde{\boldsymbol{y}}_t^{(K)} \equiv \left(\mathrm{WL}_t, \mathrm{WS}_t, \widetilde{\mathrm{WD}}_t^{(K)}\right)'$$

instead of using (4).

Here, it is necessary to determine the optimum order K and the value of ω_i. For determining unknown ω_i, we introduce the concept of principal component analysis. $\widetilde{\mathrm{WD}}_t^{(K)}$ can be written as

$$\widetilde{\mathrm{WD}}_t^{(K)} = \Omega_K' D_t^{(K)},$$

where

$$\begin{aligned}\Omega_K &= (\omega_1, \ldots, \omega_{2K})', \\ D_t^{(K)} &= (\cos(\mathrm{WD}_t), \sin(\mathrm{WD}_t), \ldots, \cos(K \cdot \mathrm{WD}_t), \sin(K \cdot \mathrm{WD}_t))'.\end{aligned}$$

We select the values of Ω_K so that

$$V\left(\widetilde{\mathrm{WD}}_t^{(K)}\right) = \Omega_K' \Sigma_t^{(K)} \Omega_K$$

is maximized under the constraints $\Omega_K' \Omega_K = 1$, where $\Sigma_t^{(K)}$ is the dispersion matrix of $D_t^{(K)}$. Ω_K can be obtained as the eigenvector $\mathbf{b}^{(K)}$ of the eigen equation,

$$\Sigma_t^{(K)} \mathbf{b}^{(K)} = \lambda \mathbf{b}^{(K)}$$

Let $\lambda_1 \geq \cdots \geq \lambda_{2K}$ be $2K$ eigenvalues of the eigen equation. We choose the eigenvector which corresponds to λ_1 with unit norm, say $\tilde{\mathbf{b}}_M^{(K)}$, with K fixed. We estimate $\widetilde{\mathrm{WD}}_t^{(K)}$ by

$$\widehat{\widehat{\mathrm{WD}}}_t^{(K)} = \tilde{\mathbf{b}}_M^{(K)} D_t^{(K)} \tag{7}$$

As for the selection of the order K, we choose the value of K such that the squared sum of the prediction errors,

$$S_l(K) = \frac{1}{N - l - T^* + 1} \sum_{t=T^*}^{N-l} \left(\mathrm{WL}_{t+l} - \widehat{\mathrm{WL}}_{t+l}^{(K)} \right)^2 \tag{8}$$

is minimized for every l, where $\widehat{\mathrm{WL}}_{t+l}^{(K)}$ is the predicted value by (6) with $\tilde{\boldsymbol{y}}_t^{(K)}$, and T^* is a prefixed value. For selection of p in (5), we adopted Akaike Information Criterion (AIC), under the value of K is fixed.

4 Evaluating the prediction accuracy of the waveheight

In the following, we examine the effectiveness of the methodology presented in the last section through the evaluation of the prediction accuracy on the waveheight by means of numerical experiments. We evaluate the accuracy by extrapolation based on the observation shown in Fig. 1. The basic procedure is as follows. First, we fit the model (5) and then obtain the prediction values of WL_t up to five steps ahead (1 step ahead corresponds to 1-min later) by the predictor (6), based on the time series data $\{\mathrm{WL}_t, \mathrm{WS}_t, \mathrm{WD}_t\}$ from $t = 1$ to $t = 50$. And then, under the sample size N is fixed as 50, we obtain the prediction values in the same way from the time series data from $t = 2$ to $t = 51$. Based on the prediction values obtained by the repetition of this procedure, we evaluate the prediction accuracy. As criteria, we used

$$\mathrm{MAE}(l) \equiv \frac{1}{M} \sum_{i=1}^{M} \left| \mathrm{WL}_{N+l}^{(i)} - \widehat{\mathrm{WL}}_{N+l}^{(i)} \right|$$

$$\mathrm{MAPE}(l) \equiv \frac{1}{M} \sum_{i=1}^{M} \left| \frac{\mathrm{WL}_{N+l}^{(i)} - \widehat{\mathrm{WL}}_{N+l}^{(i)}}{\mathrm{WL}_{N+l}^{(i)}} \right|$$

$$\mathrm{MCORR}(l) \equiv \frac{\sum_{i=1}^{M} \left(\mathrm{WL}_{N+l}^{(i)} - \overline{\mathrm{WL}}^{(i)}(l) \right) \left(\widehat{\mathrm{WL}}_{N+l}^{(i)} - \overline{\widehat{\mathrm{WL}}}^{(i)}(l) \right)}{\sqrt{\sum_{i=1}^{M} (\mathrm{WL}_{N+l}^{(i)} - \overline{\mathrm{WL}}^{(i)}(l))^2} \sqrt{\sum_{i=1}^{M} (\widehat{\mathrm{WL}}_{N+l}^{(i)} - \overline{\widehat{\mathrm{WL}}}^{(i)}(l))^2}},$$

$$\overline{\mathrm{WL}}(l) = \frac{1}{M} \sum_{i=1}^{M} \mathrm{WL}_{N+l}^{(i)}, \quad \overline{\widehat{\mathrm{WL}}}(l) = \frac{1}{M} \sum_{i=1}^{M} \widehat{\mathrm{WL}}_{N+l}^{(i)}$$

where l is the prediction step, $\mathrm{WL}_t^{(i)}$ the observation at the ith experiment, $\widehat{\mathrm{WL}}_t^{(i)}$ the predicted value at the ith experiment, M the number of repetitions. The mean absolute error (MAE) and the mean absolute percentage error (MAPE) give better evaluation as the predicted value gets closer to the observation. It is noted that MAPE becomes more unstable when the observation gets closer to zero. However, the use of MAPE did not cause serious problems in our experiment, because the observation and the predicted values of the waveheight were sufficiently far from zero. The mean correlation (MCORR) is defined as the sample correlation between the observations and predicted values, in order to evaluate the degree of accordance to their trends.

Our prediction experiment was carried out on the basis of the vector Autoregressive Integrated (ARI) model. This is ARIMA model with no MA parts, which was used as a basic model to express the sea surface movement in the wave development process in Hokimoto et al. (2003). Since the time series data of WL_t, WS_t and WD_t in Fig. 1 exhibit nonstationarity, we follow the methodology of ARIMA model by Box and Jenkins (1976), and regard the differenced time series to be stationary. We fit a vector AR model of (5) to the differenced series of the original observation, and then obtain the prediction of the original time series by "integrating" the predicted values of the differenced time series.

4.1 The effect of angular–linear structure on the linear prediction

We first investigate whether the angular–linear structure give positive effect on the prediction accuracy of the waveheight. For this purpose, we analyze whether it is possible to improve the prediction accuracy by taking into account the $2K$ variables $\{\sin(k \cdot \mathrm{WD}_t), \cos(k \cdot \mathrm{WD}_t)\}$ $(k = 1, \ldots, K)$, instead of using the variable WD_t directly. In our analysis, we apply the multivariate AR model

$$\boldsymbol{x}_t = \boldsymbol{m}_x + B_1\boldsymbol{x}_{t-1} + \cdots + B_p\boldsymbol{x}_{t-p} + \epsilon_t, \quad \epsilon_t \sim \mathrm{WN}(\mathbf{0}, \Sigma_\epsilon) \tag{9}$$

where $\boldsymbol{m}_x$ is unknown mean vector and B_i $(i = 1, \ldots, p)$ is unknown coefficient matrix. Here, we consider the following two cases for the state vector $\boldsymbol{x}_t$. The first case is the vector consisted from the difference of WL_t, WS_t and WD_t,

$$\boldsymbol{x}_t \equiv (\nabla\mathrm{WL}_t, \nabla\mathrm{WS}_t, \nabla\mathrm{WD}_t)' \tag{10}$$

where ∇ is the back-shift operator such that $\nabla\mathrm{WD}_t = \mathrm{WD}_t - \mathrm{WD}_{t-1}$, and so on. And the second case is the difference of (4) with $K = 1$, viz.,

$$\boldsymbol{x}_t \equiv (\nabla\mathrm{WL}_t, \nabla\mathrm{WS}_t, \nabla\cos(\mathrm{WD}_t), \nabla\sin(\mathrm{WD}_t))'. \tag{11}$$

Our interest is whether (11) provides more accurate prediction than (10).

Table 2 shows the values of MAE, MAPE and MCORR of (10), and Table 3 shows the ones of (11), with $M = 35$. It is noted we carried out the numerical experiment under the condition that the order p was fixed in the range from 1 to 5. Overall, the result of (11) gives smaller MAEs and MAPEs than the result of (10). It suggests the possibility that taking into account the angular–linear structure is effective for improving the prediction accuracies by the predictor based on (10). The result of MCORR also shows the similar tendency. In fact, MCORRs of (11) becomes larger

Table 2 Prediction accuracy by ARI model based on (10)

p	MAE					MAPE					MCORR				
	L=1	L=2	L=3	L=4	L=5	L=1	L=2	L=3	L=4	L=5	L=1	L=2	L=3	L=4	L=5
1	0.385	0.446	0.517	0.490	0.585	0.213	0.247	0.265	0.238	0.291	0.501	0.335	0.200	0.360	0.205
2	0.394	0.454	0.464	0.490	0.582	0.215	0.248	0.240	0.240	0.290	0.450	0.292	0.258	0.366	0.201
3	0.415	0.500	0.482	0.474	0.601	0.223	0.268	0.248	0.231	0.297	0.346	0.195	0.159	0.403	0.155
4	0.428	0.518	0.505	0.487	0.613	0.229	0.274	0.257	0.235	0.300	0.322	0.123	0.081	0.357	0.152
5	0.421	0.528	0.494	0.493	0.604	0.226	0.278	0.253	0.236	0.297	0.333	0.093	0.063	0.320	0.156

Table 3 Prediction accuracy by ARI model based on (11)

p	MAE					MAPE					MCORR				
	L=1	L=2	L=3	L=4	L=5	L=1	L=2	L=3	L=4	L=5	L=1	L=2	L=3	L=4	L=5
1	0.360	0.480	0.506	0.552	0.589	0.196	0.247	0.249	0.254	0.271	0.492	0.291	0.208	0.389	0.221
2	0.357	0.473	0.502	0.544	0.573	0.195	0.244	0.249	0.252	0.267	0.483	0.281	0.188	0.333	0.154
3	0.362	0.477	0.508	0.550	0.598	0.198	0.248	0.256	0.258	0.285	0.465	0.255	0.145	0.258	−0.078
4	0.356	0.474	0.502	0.540	0.586	0.194	0.247	0.253	0.252	0.277	0.479	0.260	0.146	0.312	−0.007
5	0.372	0.492	0.506	0.539	0.590	0.203	0.256	0.255	0.252	0.278	0.442	0.207	0.130	0.285	0.015

than the ones of (10) overall, and the difference between (10) and (11) tends to become clear as p becomes larger. However, it is probable that, as the order p and the prediction step L are larger, the prediction accuracy by (11) becomes worse to take negative correlations.

4.2 The effect of the principal component structure on the linear prediction

Next, we examine whether the principal component structure in the proposed model can improve the prediction accuracy, effectively. We carried out the numerical experiment in the same way as shown in the previous subsection, by using the state vector,

$$\boldsymbol{x}_t^{(K)} \equiv \left(\nabla \mathrm{WL}_t, \nabla \mathrm{WS}_t, \nabla \widetilde{\widehat{\mathrm{WD}}}_t^{(K)} \right)'. \tag{12}$$

Tables 4, 5 and 6 show the values of MAE, MAPE and MCORR with $K = 1$, $K = 13$ and $K = 25$, respectively. We observe through the comparison between Tables 2 and 4 that the predictor based on (12) with $K = 1$ could improve about 65–75% of the whole results using (10). However, the comparison between Tables 3 and 4 gives the result that (12) improves only 40–60% of the whole results based on (11). As far as the principal component with $K = 1$ is concerned, it is not clear whether this is effective for the prediction accuracy. However, based on the results of Tables 4, 5 and 6, we observe the tendency that MAEs and MAPEs of (12) become smaller and

Table 4 Prediction accuracy by ARI model based on (12) ($K = 1$)

p	MAE					MAPE					MCORR				
	L=1	L=2	L=3	L=4	L=5	L=1	L=2	L=3	L=4	L=5	L=1	L=2	L=3	L=4	L=5
1	0.379	0.443	0.517	0.483	0.578	0.209	0.247	0.265	0.235	0.298	0.502	0.334	0.209	0.365	0.213
2	0.390	0.451	0.454	0.483	0.583	0.213	0.247	0.235	0.236	0.291	0.452	0.294	0.271	0.373	0.200
3	0.415	0.493	0.486	0.477	0.600	0.223	0.265	0.245	0.232	0.297	0.349	0.204	0.173	0.409	0.152
4	0.420	0.511	0.504	0.490	0.608	0.225	0.271	0.256	0.236	0.298	0.324	0.134	0.084	0.357	0.137
5	0.411	0.520	0.484	0.495	0.601	0.222	0.274	0.249	0.240	0.295	0.336	0.099	0.079	0.312	0.156

Table 5 Prediction accuracy by ARI model based on (12) ($K = 13$)

p	MAE					MAPE					MCORR				
	L=1	L=2	L=3	L=4	L=5	L=1	L=2	L=3	L=4	L=5	L=1	L=2	L=3	L=4	L=5
1	0.391	0.449	0.489	0.502	0.602	0.219	0.247	0.255	0.242	0.297	0.478	0.320	0.237	0.346	0.158
2	0.342	0.433	0.463	0.497	0.575	0.191	0.234	0.240	0.244	0.284	0.447	0.303	0.247	0.358	0.166
3	0.381	0.463	0.484	0.485	0.582	0.201	0.244	0.247	0.238	0.283	0.379	0.236	0.154	0.376	0.144
4	0.383	0.464	0.481	0.488	0.583	0.203	0.245	0.245	0.240	0.283	0.374	0.231	0.157	0.375	0.142
5	0.381	0.467	0.482	0.487	0.586	0.201	0.247	0.245	0.239	0.285	0.375	0.231	0.162	0.367	0.159

Table 6 Prediction accuracy by ARI model based on (12) ($K = 25$)

p	MAE					MAPE					MCORR				
	L=1	L=2	L=3	L=4	L=5	L=1	L=2	L=3	L=4	L=5	L=1	L=2	L=3	L=4	L=5
1	0.378	0.445	0.490	0.485	0.600	0.209	0.245	0.254	0.236	0.296	0.481	0.338	0.234	0.370	0.163
2	0.353	0.435	0.456	0.502	0.574	0.195	0.238	0.236	0.247	0.284	0.433	0.296	0.259	0.360	0.180
3	0.367	0.440	0.478	0.493	0.575	0.195	0.234	0.245	0.242	0.281	0.380	0.256	0.174	0.381	0.146
4	0.367	0.446	0.473	0.480	0.571	0.196	0.236	0.241	0.234	0.277	0.384	0.258	0.178	0.396	0.173
5	0.365	0.439	0.475	0.488	0.576	0.194	0.233	0.243	0.238	0.280	0.381	0.260	0.171	0.387	0.147

MCORRs tend to become larger. In fact, in the both cases of $K = 13$ and $K = 25$, we observe that (12) improves 65–75% of the whole results of the prediction using (11). This suggests the possibility that the principal component structure of (12) worked effectively to some extent, which contributed to the improvement of the prediction accuracy.

4.3 The interaction effect between the wind and the waveheight

As shown in Sect. 2, there exists significant cross correlation between ($\mathrm{WS}_t \sin(\mathrm{WD}_t)$, $\mathrm{WS}_t \cos(\mathrm{WD}_t)$) and WL_t. Therefore, it is expected that the ARI model with the state

Table 7 Prediction accuracy by ARI model based on (13)

p	MAE					MAPE					MCORR				
	L=1	L=2	L=3	L=4	L=5	L=1	L=2	L=3	L=4	L=5	L=1	L=2	L=3	L=4	L=5
1	0.401	0.467	0.531	0.500	0.584	0.223	0.258	0.272	0.242	0.291	0.476	0.302	0.172	0.353	0.224
2	0.439	0.469	0.459	0.513	0.611	0.238	0.252	0.237	0.251	0.302	0.401	0.299	0.279	0.356	0.193
3	0.474	0.534	0.509	0.504	0.644	0.252	0.282	0.260	0.247	0.317	0.309	0.194	0.155	0.385	0.157
4	0.477	0.531	0.510	0.487	0.639	0.253	0.281	0.261	0.238	0.315	0.307	0.173	0.155	0.406	0.164
5	0.478	0.544	0.520	0.490	0.643	0.254	0.286	0.266	0.240	0.316	0.305	0.143	0.123	0.389	0.159

Table 8 Prediction accuracy by ARI model based on (14)

p	MAE					MAPE					MCORR				
	L=1	L=2	L=3	L=4	L=5	L=1	L=2	L=3	L=4	L=5	L=1	L=2	L=3	L=4	L=5
1	0.379	0.453	0.482	0.486	0.596	0.212	0.249	0.252	0.236	0.294	0.479	0.331	0.247	0.367	0.186
2	0.358	0.449	0.467	0.496	0.577	0.200	0.241	0.241	0.241	0.284	0.427	0.273	0.257	0.361	0.202
3	0.369	0.436	0.471	0.495	0.575	0.202	0.231	0.242	0.243	0.282	0.403	0.262	0.199	0.363	0.180
4	0.371	0.441	0.480	0.492	0.583	0.204	0.235	0.248	0.242	0.287	0.402	0.252	0.187	0.372	0.172
5	0.371	0.441	0.481	0.493	0.582	0.203	0.235	0.249	0.242	0.286	0.402	0.252	0.186	0.372	0.179

vectors,

$$\boldsymbol{x}_t^{(K)} \equiv (\nabla \mathrm{WL}_t, \nabla(\mathrm{WS}_t \cos(\mathrm{WD}_t)), \nabla(\mathrm{WS}_t \sin(\mathrm{WD}_t)))' \tag{13}$$

or

$$\boldsymbol{x}_t^{(K)} \equiv (\nabla \mathrm{WL}_t, \nabla(\mathrm{WS}_t \cdot \widetilde{\widehat{\mathrm{WD}}}_t^{(K)}))' \tag{14}$$

can also construct a reasonable predictor. Especially, (14) has an advantage in giving more reasonable prediction accuracies than (12), because (14) requires fewer unknown parameters of the ARI model than (12). The numerical results are shown in Tables 7 and 8. Table 7 shows that the result of (13) becomes worse than that of (12). In this sense, the state vector based on the interaction between the wind speed and wind direction is not necessarily effective for the waveheight prediction. The similar tendency can be found in the comparison between (14) and (12).

Based on these results, we can evaluate that the covariation between the changes of the wind speed and the wind direction is not remarkable (i.e., the physical processes of the wind speed and the wind direction do not completely synchronize with each other).

4.4 Appropriateness on the criterion of order selection

Finally, we examine whether the criterion proposed for order selection is appropriate for the actual prediction. The numerical experiments so far have been carried out

Table 9 Comparisons on MAE and MAPE based on the order selection

Vector	MAE					MAPE				
	L=1	L=2	L=3	L=4	L=5	L=1	L=2	L=3	L=4	L=5
(10)	0.496	0.566	0.593	0.652	0.731	0.224	0.245	0.253	0.256	0.288
(12)	0.386	0.463	0.557	0.607	0.714	0.171	0.199	0.233	0.243	0.281
(14)	0.363	0.462	0.548	0.613	0.696	0.163	0.199	0.232	0.244	0.274

under the condition that the order of K was given. However, it is probable that the prediction accuracy becomes worse, when the order K is not selected properly. So we examine the prediction accuracy when we used the proposed criterion for order selection.

The results of MAEs and MAPEs when we used the state vectors (10), (12) and (14) are given in Table 9. The prediction experiment was carried out in the range from $t = 70$ to $t = 85$, and the value of T^* was fixed as 50. Also, we adopted AIC for determination of the order p in the identification of (10). The MAEs and MAPEs of (12) becomes smaller than those of (10), which suggests that the proposed order selection criterion is appropriate for the actual prediction. On the other hand, we could not find out clear difference in the comparison between (12) and (14). But we can admit that the proposed method for order selection is available for improvement of the prediction accuracy.

5 Conclusions

We have developed an angular–linear time series model for waveheight prediction, based on the observations of the waveheight, wind speed, and wind direction. The result of our numerical experiments in the last section suggested the possibility that the time series model taking into account the angular–linear structure, based on the sine and cosine transformations of wind direction data, can improve the prediction accuracy of the linear nonstaionary time series model, in which the original wind direction data is used as an explanatory variable. As for the principal component introduced in our methodology, it is expected to improve the prediction accuracy further, by the development of the methodology based on the principal component analysis for the nonstationary time series structure.

Acknowledgments The authors would like to appreciate two reviewers and Guest Editor for their comments and suggestions.

References

Athanassoulis, G. A., Stefanakos, C. N. (1995). A nonstationary stochastic model for long-term time series of significant wave height. *Journal of Geophysical Research*, *100(C8)*, 16149–16162.

Box, G. E. P., Jenkins, G. M. (1976). *Time series analysis, forecasting and control (revised edition)*. San Francisco: Holden-Day.

Brockwell, P. J., Davis, R. A. (1991). *Time series: theory and methods*. New York: Springer.

Brockwell, P. J., Davis, R. A. (1996). *Introduction to time series and forecasting*. New York: Springer.
Brown, B. G., Katz, R. W., Murphy, A. H. (1984). Time series models to simulate and forecast wind speed and wind power. *Journal of Climate and Applied Meteorology*, *23*, 1184–1195.
Cunha, C., Guedes, S. C. (1999). On the choice of data transformation for modelling time series of significant waveheight. *Ocean Engineering*, *26*, 489–506.
Dahlhaus, R. (1997). Fitting time series models to nonstationary process. *Annals of Statistics*, *25(1)*, 1–37.
Dahlhaus, R., Giraitis, L. (1998). On the optimal segment length for parameter estimates for locally stationary time series. *Journal of Time Series Analysis*, *19(6)*, 629–655.
Daniel, A. R., Chen, A. A. (1991). Stochastic simulation and forecasting of hourly average wind speed sequences in Jamaica. *Sol Energy*, *46(1)*, 1–11.
Fisher, N. I. (1993). *Statistical analysis of circular data*. Cambridge: Cambridge University Press.
Fisher, N. I., Lee, A. J. (1992). Regression models for an angular response. *Biometrics*, *48*, 665–677.
Hokimoto, T., Kimura, N., Iwamori, T., Amagai, K., Huzii, M. (2003). The effects of wind forcing on the dynamic spectrum in wave development: A statistical approach using a parametric model. *Journal of Geophysical Research*, *108(C10)*, 5-1–5-12.
Jammalamadaka, S. R., Lund, U. J. (2006). The effect of wind direction on ozone levels: A case study. *Environmental and Ecological Statistics*, *13*, 287–298.
Jammalamadaka, S. R., SenGupta, A. (2001). *Topics in circular statistics*. Singapore: World Scientific.
Johnson, R. A., Wehrly, T. E. (1978). Some angular–linear distributions and related regression models. *Journal of the American Statistical Association*, *73*, 602–606.
Kitagawa, G., Gersch, W. (1985). A smoothness priors time-varying AR coefficient modeling of nonstationary covariance time series. *IEEE Transactions on Automatic Control*, *30*, 48–56.
Makridakis, S., Andserson, A., Carbobe, R., Fildes, R., Hibon, M., Lewandowski, R., Newton, J., Parzen, E., Winkler, R. (1984). *The forecasting accuracy of major time series methods*. Wiley: New York.
Mardia, K. V., Jupp, P. E. (2000). *Directional statistics*. New York: Wiley.
SenGupta, A. (2004). On the constructions of probability distributions for directional data, *Bulletin of the Calcutta Mathematical Society*, *96(2)*, 139–154.
SenGupta, A., Ugwuowo, F.I. (2006). Asymmetric circular–linear mutilvariate regression models with application to environmental data. *Environmental and Ecological Statistics*, *13*, 299–309.
Stefanakos, C. N., Athanassoulis, G. A., Barstow, S. F. (2002). Mutivariate time series modelling of significant wave height. *Proceedings of International Society of Offshore and Polar Engineers Conference*, *III*, 66–73.
Toll, R.S.J. (1997). Autoregressive conditional heteroscedasticity in daily wind speed measurements. *Theoretical and Applied Climatology*, *56*, 113–122.
Walton, T. L., Borgman, L.E. (1990). Simulation of non-stationary, non-gaussian water levels on the great lakes. *Journal of the ASCE, Waterway Port, Coastal, and Ocean Engineering Division*, *116(6)*, 664–685.
Yim, J.Z., Chou, C., Ho, P. (2002). A study on simulating the time series of siginificant wave height near the keelung harbor, *Proceedings of International Society of Offshore and Polar Engineers Conference*, *III*, 92–96.