Article ID: 1006-8775(2011) 04-0317-009

FORECASTING THE MJO INDEX BASED ON SSA-AR

JIANG Zhi-hong (江志红)[1], ZHANG Qin (张　勤)[1, 2], ZHU Hong-rui (朱红蕊)[1], WU Li-guang (吴立广)[1]

(1. Key Laboratory of Meteorological Disaster of Ministry of Education, Nanjing University of Information Science & Technology, Nanjing, 210044 China; 2. Wyle Information System, CPC/NOAA,USA)

Abstract: Experiments of forecasting daily bi-variate index of the tropical atmospheric Madden–Julian Oscillation (MJO) are performed in the context of adaptive filtering prediction models by combining the singular spectrum analysis (SSA) with the autoregressive (AR) methods. the MJO index, a pair of empirical orthogonal function (EOF) time series, called RMM1 and RMM2, predicts by the combined statistical SSA and AR models: firstly, according to the index of historic observation decomposed by SSA and then reconstructed by selecting the first several components based on prominent variance contributions; after that, established an AR prediction model from the composite (scheme A) or built the forecast models for each of these selected reconstructed components, separately (Scheme B). Several experimental MJO index forecasts are performed based on the models. The results show that both models have useful skills of the MJO index forecast beyond two weeks. In some cases, the correlation coefficient between the observed and predicted index series stays above 0.5 in 20 leading days. The SSA-AR model, based on the reconstructed composite series, has better performance on MJO forecast than the AR model, especially for the leading time longer than 5 days. Therefore, if we build a real-time forecast system by the SSA-AR model, it might provide an applicable tool for the operational prediction of the MJO index.

Key words: Madden-Julian Oscillation; singular spectrum analysis; autoregressive model

CLC number: P435 **Document code:** A **doi:** 10.3969/j.issn.1006-8775.2011.04.001

1 INTRODUCTION

Tropical atmospheric intraseasonal oscillation (ISO), i.e. the atmospheric 30-60 day low frequency oscillation, also called Madden-Julian oscillation (MJO), is the most prominent oscillation signal on the intraseasonal scale over the global tropical regions. Its activities are significantly related to the tropical climate systems, such as the evolution (onset, strengthening, break, and withdrawal) of monsoon systems (Madden and Julian[1, 2]; Li[3, 4]; Jia and Li[5]; Lin et al.[6]). Since Madden and Julian presented the MJO in the early 1970s, it has been an interesting topic for meteorological researchers both in China and overseas.

It is a well-known fact that prediction on a time scale from two weeks to a season is a crux in long-range weather forecasts and short-range climatic predictions. The MJO, as an intrinsic variation on a time scale between the short range (2-7 days) and the seasonal range (90 days), provides a direct background for high frequency weather disturbances, and is also a principal component of monthly and seasonal mean variations. Meanwhile, it modulates the activities of monsoon systems and affects the middle latitude weather through jets. Therefore, prediction of the MJO has drawn much attention from researchers in China and overseas in the recent years. Von Storch and Xu[7] and Lo and Hendon[8] used the principal oscillation pattern (POP) technique and the empirical orthogonal function (EOF) method, respectively, to identify the MJO and made prediction, yielding a correlation skill of 0.4 for 15-day forecasts for all seasons; the latter reached 0.48 for the correlation coefficient between the observations in the validation set and the forecasts. Maharaj and Wheeler[9] made 14-15 days ahead the prediction of a daily bi-variate index of the MJO, defined by Wheeler and Hendon[10] (WH04 hereafter), using the autoregression (AR) method. Therefore, considerable predictability of the MJO may lie on a time scale from a week to a month. At present, some institutions, such

Received 2010-07-12; **Revised** 2011-08-18; **Accepted** 2011-10-15
Foundation item: National Key Technologies R & D Program (2009BAC51B01); Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD); Natural Science Foundation of China (40875058)
Biography: JIANG Zhi-hong, Ph.D., professor, doctorate tutor, primarily undertaking research on climatology.
Corresponding author: JIANG Zhi-hong, e-mail: zhjiang@nuist.edu.cn
Note: Beginning from V.14(1), 2008, *Journal of Tropical Meteorology* is indexed and abstracted in Science Citation Index Expanded and Journal Citation Reports/Science Edition.

as Climate Prediction Center in U.S. and Australian Bureau of Meteorology Research Center, have employed both dynamical models (GEFS/NCEP, GFS/NCEP, etc) and statistical methods (lag regression and linear autoregression) to forecast the MJO, with emphasis on operational forecasts of the MJO[11, 12]. Obviously, increasing the ability of models to forecast the MJO using various methods will be beneficial to the ability of predicting the short range climate on the time scale of two weeks to a month.

The singular spectrum analysis (SSA), as a digital signal processing technique, was introduced to climatic diagnosis and prediction research by Vaulard et al.[13] and Ghil and Mo[14] in the early 1990s, and it was mainly used to both analyze the quasi-periodic signals and extract related signals. Ding et al.[15] conducted forecast experiments for Niño-region averaged monthly sea surface temperature (SST) anomaly series by reconstruction of the adaptive filtered signals of SSA and then based on the AR prediction modeling. The results show that the SSA-AR model based on the reconstructed SST anomaly series are superior in forecast accuracy over the corresponding AR model based on the original SST anomaly series. Furthermore, the predictions of the 1997/98 strong ENSO event by the SSA-AR model were very successful, as well as predicting the other three historic strong ENSO events (1974–1976, 1980–1983 and 1984–1986) several months ahead with a high confidence level.

Two SSA-AR prediction schemes based on the reconstructed index series and an AR prediction model based on the original index series (following WH04) were designed in this paper, four prediction cases were selected, and a series of prediction experiments were performed and the experimental results were compared in order to select the best model among them for applying to the real-time operational forecast of the MJO.

2 MJO INDEX AND ITS PREDICTION MODELS

2.1 *MJO index vector*

Wheeler and Hendon[10] presented a MJO index comprising a pair of projection time series of the leading empirical orthogonal functions (EOFs) of the combined fields: 850-hPa zonal wind, 200-hPa zonal wind, and satellite-observed outgoing longwave radiation (OLR) averaged near-equatorially (15° S–15° N). Two components of the MJO index vary mostly on the 30- to 80-day timescale, and are called the Real-time Multivariate MJO series 1 (RMM1) and 2 (RMM2). During times of identifiable MJO activities, RMM1 leads RMM2 by about a quarter cycle, indicating the eastward propagation of the MJO along the equator. Fig. 1 displays the 2-D phase diagram from 1st November to 31st December 2007. It can be seen from the figure that RMM1 and RMM2 in the phase diagram could define eight different regions of the phase space, corresponding to the location of the MJO. On one hand, the MJO index vector quantitatively describes the intensity of the MJO, such as the strong activity of the MJO in Fig. 1 where the amplitude of the oscillations, ($\sqrt{RMM1^2 + RMM2^2}$), is larger than one (normalized standard variation), and there are weak activity regions inside of the circle. On the other hand, the index vector has directivity and is able to delineate the location as well as the eastward propagation direction of the MJO along the equator. In WH04, the relations of the MJO with the onsets of the Australian monsoon and the Indian monsoon and the modulation of the MJO on Australian precipitation are investigated by using the MJO index. Jeong et al.[16] recently explored the influence of the MJO on East Asian winter precipitation in terms of the MJO index. These studies suggest that the pair of the RMM1 and RMM2 series is a good indicator of the MJO's intensity and variation. At present, the RMM1 and RMM2 are computed daily for real-time MJO monitoring, and forecasts from dynamic models (GEFS/NCEP, GFS/NCEP, etc) in Climate Prediction Center/NOAA, U.S. and Australian Bureau of Meteorology Research Center, which provide predicted values of the index two weeks in advance. The forecast MJO index has become a very useful tool to apply to global tropics hazards assessment and short-term climate prediction.

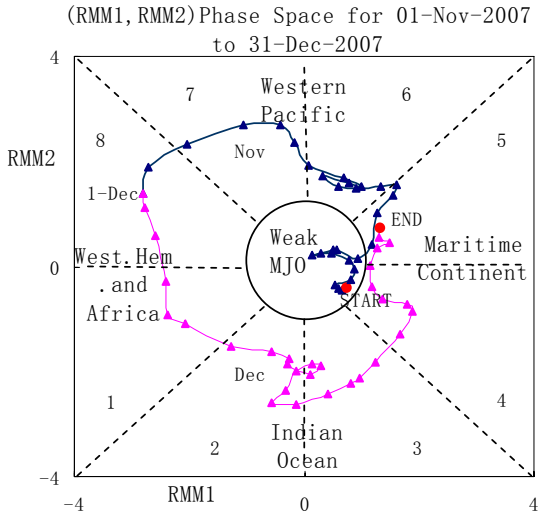

Fig. 1. (RMM1, RMM2) phase space points for all days in November to December 2007. Eight defined regions of the phase space are labeled, as is the region considered to signify weak MJO activity. Also labeled are the approximate locations of the enhanced convective signal of the MJO for that location of the phase space.

2.2 *SSA-AR(p) random vector prediction models*

2.2.1 SSA

The SSA is an anamorphosis of principal component analysis (PCA) or EOFs. Arranging a time series $\{Xt\}$, $t = 1,\ldots, N$ successively in one time unit lag in the next row. For m intervals it yields an $m \times n$ matrix.

$$X = \begin{pmatrix} x_1 & x_2 & \cdots & x_{N-m+1} \\ x_2 & x_3 & \cdots & x_{N-m+2} \\ \cdots & \cdots & \cdots & \cdots \\ x_m & x_{m+1} & \cdots & x_N \end{pmatrix} = (x_{it})_{m\times(N-m+1)} . \quad (1)$$

The EOF expansion of matrix X is called time-domain EOF expansion (denoted as TEOF).

Eigenvectors obtained from the matrix are called canonical waveform signals, denoted as TEOF1, TEOF2,…, and their corresponding principal components are denoted as TPC1, TPC2, …, etc. Then the component series of various frequencies are reconstructed with corresponding TPCs and TEOFs, respectively, and are denoted as RCs. The calculation of the kth canonical wave for the ith phase component is based on

$$X_t^{(k)} = \begin{cases} \dfrac{1}{m}\sum_{i=1}^{m} L_{ki} a_{k,t-i+1}, & m \le t \le N-m+1 \\ \dfrac{1}{t}\sum_{i=1}^{t} L_{ki}, & 1 \le t \le m-1 \\ \dfrac{1}{N-t+1}\sum_{i=t-N+m}^{m} L_{ki} a_{k,t-i+1}, & N-m+2 \le t \le N \end{cases}, \quad (2)$$

where L_{ki} is the ith phase component of the kth canonical wave, $a_{k,t-i+1}$ is the value of the kth principal component at the time of t-i+1. If we cut off the high frequency, p represents the series number of first p components. The first p prominent canonical waveforms or periodical components in Eq. (2) are used to reconstruct the original series,

$$X_t \approx \sum_{k=1}^{p} X_t^{(k)}, \quad (3)$$

where $X_t^{(k)}$ is the reconstructed series of the kth component of X_t, generally denoted as RCs-K (Ding and Jiang[17]).

2.2.2 AR(p) MODEL OF RANDOM VECTORS

AR prediction modeling is very similar to multivariate linear prediction modeling. The autoregressive model of a time series takes some values of the time series at past times as predictors. Assuming that time series $X_t = [x_{1t}, x_{2t},\ldots, x_{nt}]^{\mathrm{T}}$ is an n-dimension stationary random series, then the following random vector AR(p) model can be established.

$$X_t = A_1 X_{t-1} + A_2 X_{t-2} + \ldots + A_p X_{t-p} + \mu_t, \quad (4)$$

where A_1, A_2, …, A_p are the regression coefficients, each one is a matrix, such as $A_1 = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$. And $\mu_t = [\mu_{1t}, \mu_{2t},\ldots, \mu_{nt}]^{\mathrm{T}}$ is an n-dimension white noise. Regression coefficient matrices A_1, A_2, …, A_p can be obtained by estimation of the minimum square root of variance, and then the AR model in a form of random vectors, such as Eq. (4), can be established.

2.2.3 SSA-AR(p) MODELS

The following two schemes were used in this paper to build up SSA-AR(p) models for forecasting the MJO index vector:

Scheme A: (i) decompose the MJO index yielding TEOFs and TPCs; (ii) compute RCs component using Eq. (2); (iii) select the first several components according to prominent variance contributions, and then calculate the composite of the RCs according to Eq. (3); (iv) establish the SSA-AR(p) prediction model for the composite using the AR(p). The model established on Scheme A is denoted as SSA-AR(A).

The reconstructed component prediction model based on Scheme B uses the same steps (i) and (ii) as in Scheme A. Then, step (iii) is used to select the first several components with prominent variance contributions, and establish a random vector autoregressive prediction model in the form of Eq. (4) for each of these selected reconstructed component, separately. After that, (iv) prediction of the original series is realized by adding up the predicted values of these selected reconstructed components. The model

with Scheme B is denoted as the SSA-AR(B).

To assess the performance of Schemes A and B, the MJO index series were also directly used to build the AR(*p*) prediction model, which is the same as that used by Maharaj and Wheeler (MW), and therefore denoted as MW-AR model. For details of its modeling method, refer to Maharaj and Wheeler[9].

3 ANALYSIS OF RESULTS FROM SSA-AR MODELS

3.1 *Analysis of the results of SSA*

The results of SSA performance with RMM1 and RMM2 series from 1st January 1979 through 31st December 2000 (*N*=8036) are listed in Table 1. The TPCs1-6 / TPCs1-8 of the first six/eight canonical waveform vectors for the RMM1 series account for 69.83% / 80.23% of the total variance, respectively, and the TPCs1-10 accounts for 86.18% of the total variance. For the RMM2 series, the TPCs1-6, TPCs1-8 and TPCs1-10 account for 72.33%, 82.57% and 87.83% of the total variance, respectively. The reconstructed component series, RCs, for RMM1 and RMM2, are calculated from Eq. (2) by their selected TPCs and TEOFs, and summing up the RCs1-*K* by Eq. (3) yields their composite series, respectively. The *Xt* obtained in such a way is virtually equivalent to a low-pass filtered series of the original series. Fig. 2 shows the original series and the RCs1-6 composite series of RMM1 from 1st January 1979 to 31st December 1982. It is shown that the RCs1-6 composite series is RMM1 with high-frequency noises filtered. The correlation coefficient between the original and the RCs1-6 composite of RMM1 is 0.902. Applying SSA decomposition to RMM2 is similar to that of RMM1 (Figure omitted) and the correlation coefficient of RMM2 is 0.95.

Table 1. Variance contributions (VC) and accumulative VCs (AVC) of the TPCs1-10 of the first ten canonical waveforms for the RMM1 and RMM2 series from 1st January 1979 through 31st December 2000.

RMM1	VC(%)	AVC(%)	RMM2	VC(%)	AVC(%)
TPC1	17.00	17.00	TPC1	17.26	17.26
TPC2	16.24	33.24	TPC2	16.87	34.13
TPC3	11.54	44.78	TPC3	12.85	46.98
TPC4	10.17	54.95	TPC4	11.01	57.99
TPC5	7.470	62.42	TPC5	7.540	65.53
TPC6	7.410	69.83	TPC6	6.800	72.33
TPC7	6.190	76.02	TPC7	6.230	78.56
TPC8	4.210	80.23	TPC8	4.010	82.57
TPC9	3.390	83.62	TPC9	3.020	85.59
TPC10	2.560	86.18	TPC10	2.240	87.83

3.2 *SSA-AR(p) modeling and analysis of prediction results*

For the convenience of comparison with MW's results, data from 1st January 1979 through 31st December 2000 (*N*=8036) were used to fit AR models. The final prediction error (FPE) criterion was used to select the optimal order *p* for the RCs1-*K* composite series and component series of RMM1 and RMM2. The results show that most RCs series are consistent with AR(1) models.

Similarly with MW's, data from 1st January 2001 to 10th November 2003 were used for the validation of SSA-AR(A), SSA-AR(B) and AR models. The correlation coefficients between the observed and predicted values are used to assess prediction abilities of the three models.

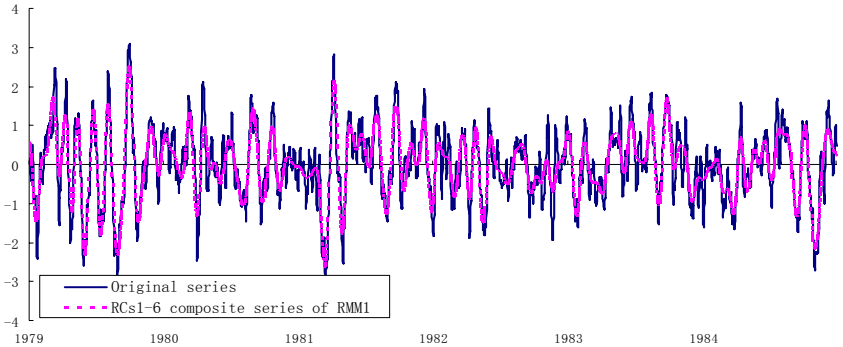

Fig. 2. The original series and RCs1-6 composite series of RMM1 from 1st January 1979 through 31st December 1984.

In SSA-AR prediction modeling, the prediction effectiveness is directly related to the number of component series with prominent variance contributions used. Table 2 exhibits the correlation skills of 6, 8 and 10 RCs SSA-AR(A) models; for short-range (1-7 days) forecasts, the skills of the RCs1-6 SSA-AR models for RMM1 and RMM2 are lower than those of 8 and 10 RCs models. However, for the extended-range (8-20 days) forecasts the opposite is true, and the RCs1-6 SSA-AR models yield a correlation skill of 0.5 for 20-day forecasts in the validation experiment, and the corresponding skills of the 8 and 10 RCs models are only 0.466 and 0.453, respectively. The objective of this paper is to select a model for operationally making the longer lead time forecasts of the MJO, therefore the RCs1-6 SSA-AR models are appropriate. Similar prediction experiments were also performed for the SSA-AR(B) models, yielding similar results. Therefore, the comparison of prediction modeling for SSA-AR(A) and SSA-AR(B) below is all based on the first six reconstructed component series RCs1-6.

Table 2. RMM1 and RMM2 prediction abilities (correlation coefficient between the predicted and observed values) of SSA-AR(A), SSA-AR(B) and AR(p) models in the validation experiment from 1st January 2001 to 10th November 2003.

SSA-AR prediction modeling									MW-AR(P) prediction modeling		
SSA-AR(A)							SSA-AR(B)				
RCs1-6 compos. series			RCs1-8 compos. series		RCs1-10 compos. series		RCs1-6 component series				
Lead time (days)	RMM 1	RMM 2	RMM1	RMM 2	RMM1	RMM2	RMM 1	RMM2	Lead time (days)	RMM1	RMM2
1	0.885	0.942	0.927	0.963	0.946	0.974	0.884	0.940	1	0.977	0.985
2	0.876	0.939	0.914	0.957	0.929	0.967	0.870	0.934	2	0.931	0.957
3	0.859	0.934	0.895	0.947	0.903	0.954	0.848	0.926	3	0.880	0.924
4	0.837	0.926	0.871	0.933	0.870	0.937	0.82	0.915	4	0.829	0.889
5	0.811	0.916	0.842	0.916	0.834	0.915	0.788	0.899	5	0.782	0.853
6	0.782	0.903	0.812	0.894	0.796	0.888	0.756	0.878	6	0.741	0.815
7	0.753	0.887	0.781	0.868	0.760	0.857	0.727	0.852	7	0.708	0.775
8	0.725	0.867	0.751	0.839	0.728	0.823	0.704	0.820	8	0.682	0.733
9	0.698	0.844	0.723	0.807	0.700	0.785	0.687	0.782	9	0.662	0.690
10	0.675	0.816	0.698	0.772	0.677	0.745	0.677	0.739	10	0.645	0.650
11	0.655	0.786	0.676	0.735	0.658	0.704	0.675	0.692	11	0.629	0.613
12	0.638	0.752	0.655	0.697	0.641	0.663	0.674	0.644	12	0.606	0.580
13	0.624	0.717	0.636	0.659	0.626	0.623	0.673	0.599	13	0.589	0.549
14	0.611	0.680	0.618	0.621	0.610	0.585	0.671	0.569	14	0.571	0.518
15	0.600	0.643	0.600	0.585	0.593	0.549	0.666	0.550	15	0.552	0.490
16	0.588	0.606	0.581	0.551	0.574	0.518	0.647	0.540	16	0.534	0.465
17	0.576	0.571	0.561	0.519	0.552	0.490	0.618	0.529	17	0.517	0.443
18	0.562	0.537	0.539	0.491	0.528	0.466	0.58	0.518	18	0.498	0.424
19	0.546	0.507	0.515	0.466	0.501	0.447	0.533	0.511	19	0.478	0.409
20	0.526	0.480	0.488	0.444	0.473	0.432	0.494	0.507	20	0.456	0.396

Figures 3 and 4 show the 10-day forecasts of RMM1 series, respectively, by the RCs1-6 SSA-AR(A) and RCs1-6 SSA-AR(B) models as well as the MW-AR model. For the sake of brief description, only the predicted RMM1 series from 1st January 2001 to 23rd August 2002 are given. The correlation coefficients between the original and predicted series for the three prediction models are 0.675, 0.677 and 0.645, respectively, and therefore the correlation skill of the AR model based on the original

series for 10-day forecasts is relatively low compared to those of the SSA-AR(A) and SSA-AR(B) models. The reason is that the SSA decomposition of the original series is equivalent to performing a low-pass filtering to the original series, thus removing the influence of background noises and improving the prediction effectiveness of SSA-AR models.

Figure 5 shows the RCs1-6 SSA-AR(A), RCs1-6 SSA-AR(B) and MW-AR model's prediction skill, as determined by the correlations, up to day 20. It can be seen from the figure that the skills of the 1-3 day for RMM1 forecasts by MW-AR model is better than those by SSA-AR(A) and SSA-AR(B) models. However, its more-than-3-day forecasts are obviously worse than those by the SSA-AR models. If we assume that a correlation coefficient of 0.5 is a useful skill criterion, the lead time of the skillful forecasts of SSA-AR(A) and SSA-AR(B) models is 20 and 19 days, respectively, both longer than the 17 days of the MW-AR model. The results for RMM2 forecasts are similar to RMM1. The skill of the 1-2-day RMM2 forecasts of MW-AR model is better than those of the SSA-AR models; after the first three days, however, forecasts are obviously worse than those by the SSA-AR models. The lead time of the skillful forecasts of SSA-AR(A) and SSA-AR(B) models is also 20 and 19 days, respectively, both longer than the 14 days of MW-AR model. For forecasts of both RMM1 and RMM2, the average skill of the 1-3 day forecasts of MW-AR model is better than those of the SSA-AR models; but after 3 days, the forecast skill rapidly decreases and the lead time of skillful forecasts is 5 days shorter than those of both SSA-AR(A) and SSA-AR(B) models, respectively.

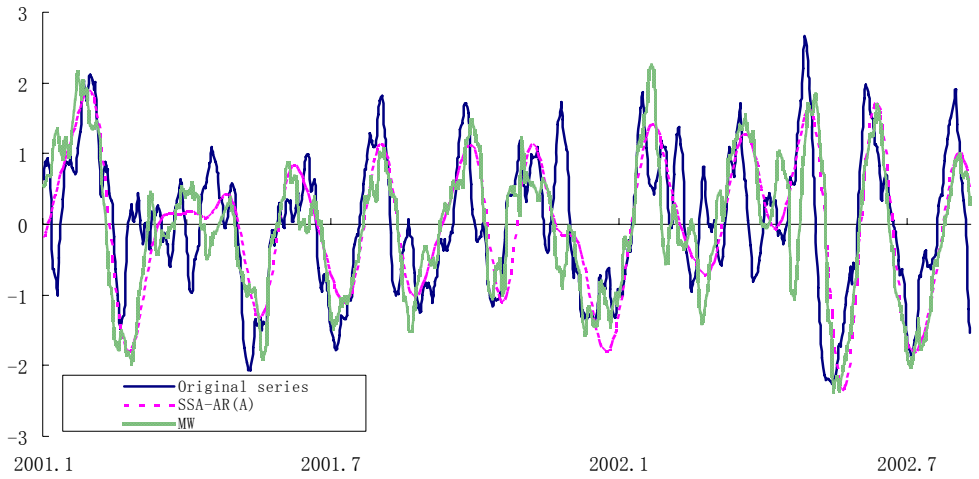

Fig. 3. The original RMM1 series from 1st January 2001 to 23rd August 2002, and its 10-day predicted series by the RCs1-6 SSA-AR(A) model and MW-AR model.

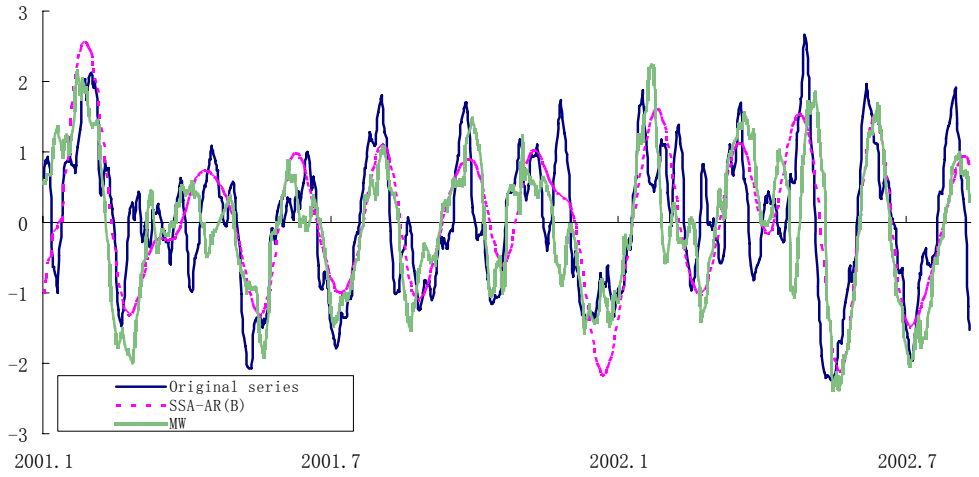

Fig. 4. Same as Fig.3 except by the RCs1-6 SSA-AR(B) model.

Comparing SSA-AR(A) with SSA-AR(B) schemes reveals that their skills are close to each other and the lead time of the 0.5 correlation is both 20 days. However, the computation of prediction modeling with SSA-AR(A) scheme is easier than that of SSA-AR(B) scheme.

Put in a brief summary from the above analysis, the forecast skills of the two SSA-AR schemes in extended-range forecasts are better than that of the MW-AR scheme because the SSA-AR(A) and SSA-AR(B) models are based on RCs1-6 composite series and component series, respectively, which reduce or remove the background noises from the reconstructed series by the decomposition-composition of SSA in prediction modeling and yield 5-day improvement of the MJO index forecasts.

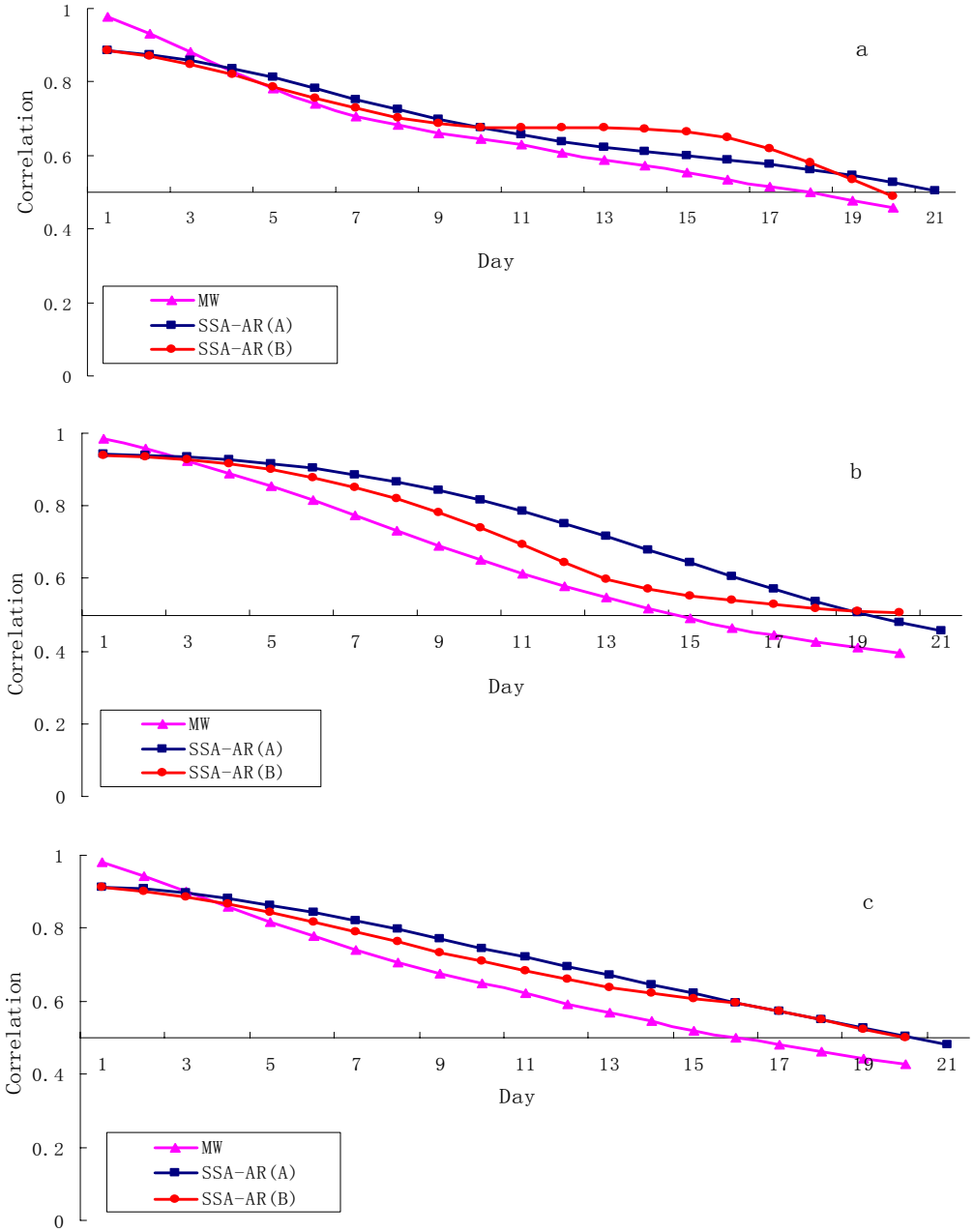

Fig. 5. Prediction skill of MW-AR models, SSA-AR(A) and SSA-AR(B) models for (a) RMM1 series, (b) RMM2 series, and (c) both RMM1 and RMM2 series (the abscissa: days ahead of time; the ordinate: correlation coefficient).

3.3 *Analysis of prediction cases of the MJO index*

To validate MJO index forecast of the SSA-AR models, the 20-day forecasts of RMM1 and RMM2 started from 16th November 2002, 22nd October 2003, 1st May, 2006 and 3rd January 2008 were plotted in Fig. 6. For convenience of comparison, the first two initializations were selected to fall on the same dates as that in MW's (Maharaj and Wheeler[9]).

As illustrated in Fig. 6, all tracks of predicted and observed values in the four cases rotate counterclockwise from the initial points, reflecting the eastward propagation characteristic of the MJO along the equator. For the 2002 case, the amplitude predicted by the 1-3 day forecasts of MW-AR models

is closest to the observation. However, after 3 days, the predicted amplitude rapidly decays, much different from the observation. The prediction skills of the amplitude of the two SSA-AR models are obviously superior over that of the MW-AR model, and the amplitude predicted by SSA-AR(A) models is closest to the observation. With regard to the eastward propagation, the 20-day ahead prediction skill of eastward propagation speed for MW-AR models is inferior to that of SSA-AR(B) models, whose predicted speed is closest to the observation. For the 2003 case, the amplitude of the MJO predicted by MW-AR models is closer to the observation than those by the other two models. However its predicted eastward propagation speed is obviously inferior to that by SSA-AR(B) models.

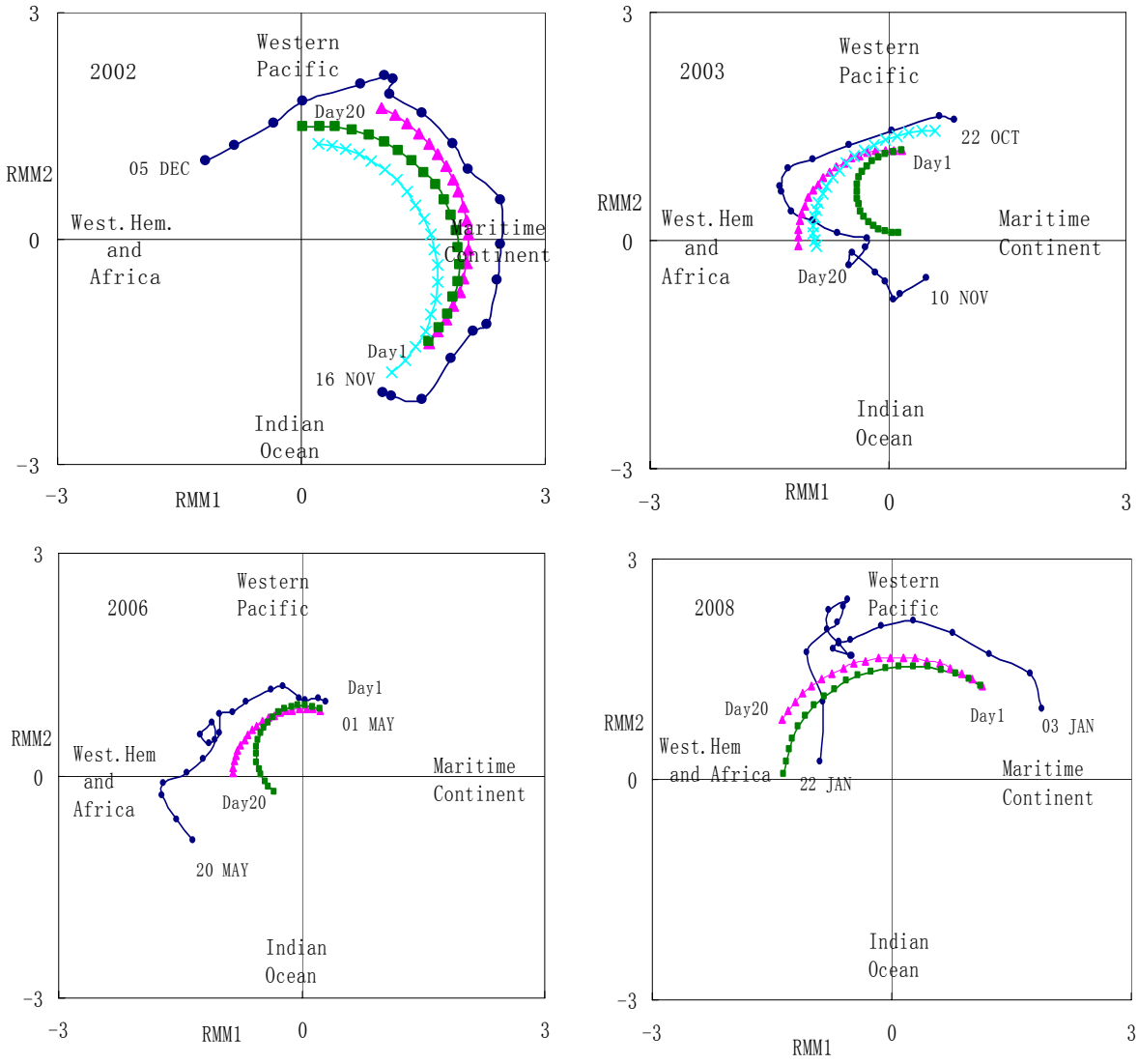

Fig. 6. RMM1 and RMM2 forecasts by SSA-AR(A) models (triangles), SSA-AR(B) models (squares) and MW-AR models (crosses), and their validating observations (circles) for four example periods (Initialization dates of the 1- to 20-day ahead forecasts are 15th November 2002, 21st October 2003, 30th April 2006 and 2nd January 2008), as represented in the two-dimensional phase space they define. Also labeled are the approximate locations around the earth where the enhanced convective signal of the MJO will be located for that part of the (RMM1, RMM2) phase space (e.g. 'Indian Ocean' for the MJO signal in convection located over the near-equatorial Indian Ocean).

A consistent conclusion can be drawn from the four prediction cases that the 1-5 day amplitude forecast skills for the two SSA-AR models are close to each other. After 5 days, the predicted amplitude by SSA-AR(B) models decays rapidly but the predicted amplitude by SSA-AR(A) is closer to the observation. The forecast skills of eastward propagation speed of SSA-AR(A) model in all four cases are obviously slower than that of SSA-AR(B) models.

The above forecast experiments have demonstrated that both SSA-AR(A) and SSA-AR(B) schemes are able to make 20-day lead useful forecasts of the MJO index. The phase evolution of the MJO described by both schemes is consistent with the

observations, except that there are still some differences in the eastward propagation speed, period and amplitude of the MJO between predicted and observed values in some individual cases. Generally speaking, the forecast skill of SSA-AR(A) models for the amplitude of the MJO is higher, while that of SSA-AR(B) models for the eastward propagation speed of the MJO is better.

4 SUMMARY

(1) SSA-AR(A) and SSA-AR(B) models for the MJO vector both achieve an average correlation skill of 0.5 for 20-day forecasts, which is a 5-day improvement in lead time relative to MW-AR models (Maharaj and Wheeler[9]). It derives from using SSA in prediction models because SSA is equivalent to performing a low-pass filtering to the original series of the MJO index, removing both high-frequency noises and non-periodic weak signals, and enhancing predictability of the MJO component by reconstructed series.

(2) In the forecast experiments of the four cases, the predicted values from two SSA-AR schemes on the whole fit well with the observation values, except with slightly slower eastward propagation speed and faster decay amplitude of the MJO in some individual cases.

(3) Relatively, the forecast skill of SSA-AR(A) models for the amplitude of the MJO is higher, while that of SSA-AR(B) models for the eastward propagation speed of the MJO is better. If SSA-AR(A) and SSA-AR(B) schemes are further improved from various aspects, they might provide effective operational tools to forecast the MJO index for longer than 20 days.

REFERENCES:

[1] MADDEN R A, JULIAN P R. Detection of a 40-50 day oscillation in the zonal wind in the tropical Pacific [J]. J. Atmos. Sci., 1971, 28: 702-708.
[2] MADDEN R A, JULIAN P R. Description of global scale circulation cells in the tropics with a 40-50 day period [J]. J. Atmos. Sci., 1972, 29: 1109-1123.
[3] LI Chong-yin. Atmospheric Low-Frequency Oscilations [M]. Beijing: China Meteorological Press, 1993: 19-43 (in Chinese).
[4] LI Chong-yin. Research progresses in atmospheric intraseasonal oscillations [J]. Prog. Nat. Sci., 2004, 14(7): 734-741 (in Chinese).
[5] JIA Xiao-long, LI Chong-yin. Seasonal characteristics of tropical atmospheric intraseasonal oscillations and its behavior in the SAMIL-$R_{42}L_9$ model [J]. J. Trop. Meteor., 2007, 23(3): 219-228 (in Chinese).
[6] LIN Ai-lan, LIANG Jian-yin, GU De-jun. Research progresses in influence of tropical atmospheric intraseasonal oscillations on the East Asian monsoon region and its different time scale changes [J]. J. Trop. Meteor., 2008. 24(1): 11-19 (in Chinese).
[7] VON STORCH H, XU J. Principal oscillation pattern analysis of the tropical 30–60 day oscillation. Part I: Definition of an index and its prediction [J]. Climate Dyn., 1990, 4: 179-190.
[8] LO F, HENDON H H. Empirical extended-range prediction of the Madden-Julian oscillation [J]. Mon. Wea. Rev., 2000, 128: 2528-2543.
[9] MAHARAJ E A, WHEELER M C. Forecasting an index of the madden-oscillation [J]. Int. J. Climatol., 2005, 25: 1611-1618.
[10] WHEELER M C, HENDON H H. An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction [J]. Mon. Wea. Rev., 2004, 132: 1917-1932.
[11] National Weather Service (USA). MJO Monitoring and Forecast [EB/OL]. http://www.cpc.ncep.noaa.gov/products/people/wd52qz/mjoindex/ MJO_INDEX.html. [2009-01-15]
[12] Bureau of Meteorology Research Center (Australia). An All-season Real-time Multivariate MJO Index [EB/OL]. http://www.bom.gov.au/bmrc/clfor/cfstaff/matw/maproom/RMM/ [2009-01-15]
[13] VAUTARD R, YIOU P, GHIL M. SSA: A toolkit for short, noisy chaotic signals [J]. Phys. D., 1992, 58: 95-126.
[14] GHIL M, MO K C. Interseasonal oscillations in the global atmosphere [J]. J. Atmos. Sci., 1991, 48: 752-779.
[15] DING Yu-guo, JIANG Zhi-hong, ZHU Yan-feng. Experiment on short-term climatic prediction to SSTA over the Nino oceanic region [J]. J. Trop. Meteor., 1998. 14(4): 289-296 (in Chinese).
[16] JEONG Jee-hoon, KIM Baek-min, HO Chang-hoi, et al. Systematic variation in wintertime precipitation in East Asia by MJO-induced extratropical vertical motion [J]. J. Climate, 2008, 21: 788-801.
[17] DING Yu-guo, JIANG Zhi-hong. Signal Processing of Meteorological Data Time Series [M]. Beijing: China Meteorological Press. 1998. 91-172 (in Chinese).

Citation: JIANG Zhi-hong, ZHANG Qin, ZHU Hong-rui et al. Forecasting the MJO index based on SSA-AR. *J. Trop. Meteor.*, 2011, 17(4): 317-325.