Assessing Predictive Potential Associated with the MJO during the Boreal Winter

ARUN KUMAR, a JIESHUN ZHU, a,b AND WANQIU WANG a

a Climate Prediction Center, NOAA/NWS/NCEP, College Park, Maryland
b Earth System Science Interdisciplinary Center, University of Maryland, College Park, College Park, Maryland

(Manuscript received 23 April 2020, in final form 2 October 2020)

ABSTRACT: In this paper, the question of potential predictability in meteorological variables associated with skillful prediction of the Madden–Julian oscillation (MJO) during boreal winter is analyzed. The analysis is motivated by the fact that dynamical prediction systems are now capable of predicting MJO up to 30 days or earlier (measured in terms of anomaly correlation for RMM indices). Translating recent gains in MJO prediction skill and relating them back to potential for predicting meteorological variables—for example, precipitation and surface temperature—is not straightforward because of a chain of steps that go into the computation and evaluation of RMM indices. This paper assesses potential predictability in meteorological variables that could be attributed to skillful prediction of the MJO. The analysis is based on the observational data alone and assesses the upper limit of MJO-associated predictability that could be achieved.

KEYWORDS: Madden-Julian oscillation; Forecasting

1. Introduction

The Madden–Julian oscillation (MJO) is the most prominent mode of variability in the equatorial latitudes on subseasonal time scales (Madden and Julian 1971). MJO variability has been documented to modulate monsoon variability (Lavender and Matthews 2009), hurricane activity in tropical latitudes (Maloney and Hartmann 2000), extratropical weather regimes (Lin and Brunet 2009; Moore et al. 2010; Zhou et al. 2012) and plays an important role in the development of ENSO events (Kessler and Hartmann 2000; Zhang and Gottschalck 2002; McPhaden et al. 2006; Wang et al. 2011) (see also reviews by Zhang 2005; Lau and Waliser 2012).

Apart from controlling large-scale environmental factors conducive for the evolution of various meteorological phenomena (e.g., potential for the development of hurricanes in the Atlantic Ocean), skillful prediction of the MJO also holds promise for imparting prediction skill in meteorological variables such as precipitation and surface temperature, which are of importance in the context of decision making. Empirical forecast tools have been developed that exploit this link and utilize MJO information for predictions (Zhou et al. 2012; Riddle et al. 2013; Johnson et al. 2014).

In the last decade, advances have been made in the prediction of MJO using dynamical models (e.g., Vitart 2017). These are due to improvements in the observations and data assimilation systems, improvements in the physical parameterization schemes used in the models, as well as due to an increase in horizontal and vertical model resolutions (e.g., National Research Council 2010; National Academies of Sciences, Engineering, and Medicine 2016). Further improvements in the prediction of MJO variability may rely on advances in coupled data assimilation and prediction systems and resolving issues such as the successfully modeling of MJO propagation across the Maritime Continent (Slingo et al. 1996; Lin et al. 2006; Jiang et al. 2015; Zhu et al. 2017) and the representation of diurnal and intraseasonal sea surface temperature (SST) variations (Ge et al. 2017).

The quantification of improvements in the MJO skill is generally based on the assessment of prediction skill of real-time multivariate MJO (RMM; Wheeler and Hendon 2004) indices that are derived based on (i) meridional means of daily outgoing longwave radiation (OLR) and zonal wind at 850 and 200 hPa (U850 and U200, respectively) (after the removal of interannual variability related to ENSO) and (ii) computing the leading empirical orthogonal functions (EOFs) that typically account for ~25% of variability in the input variables. Basic consideration of loss of predictability in initialized predictions with lead time also dictates that the skill in predicting RMM indices will decrease with the forecast lead time as has been documented (e.g., Vitart 2017). Even if in an idealized scenario that the RMM indices are predicted perfectly, it remains to be quantified how much the knowledge of RMM indices alone can contribute to the prediction skill in meteorological variables, for example, precipitation and surface temperature, the variables that are of critical importance in the context of decision making. A straightforward assessment of this question gets obscured because of the multiplicity of steps that are involved in the calculation of RMM indices.

The pragmatic goal of this analysis is to quantify predictive potential for the meteorological variables that can be associated with the skillful prediction of the MJO. Using a simple approach, we quantify to what extent daily variability in meteorological variables can be associated with the knowledge of RMM indices, and further, how the predictability may vary for different time averages. The knowledge of variance explained by the RMM indices allows us to estimate the predictability (and prediction skill) in meteorological variables that can be achieved based on the knowledge of the MJO variability. In this paper our focus is on the illustration of the approach that can be used to quantify predictive skill associated with the MJO.

DOI: 10.1175/MWR-D-20-0128.1

For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).
2. Data and analysis procedure

The observational data analyzed in this study include OLR from the NOAA Advanced Very High Resolution Radiometer (Liebmann and Smith 1996), U850 and U200, and 2-m temperature (T2m) from the Climate Forecast System Reanalysis (Saha et al. 2010), and rainfall estimate from the Climate Prediction Center morphing technique (CMORPH) based on satellite radiances using several channels (Joyce et al. 2004). The analysis period is the boreal winter (November–April) during 1988–2017 (1998–2017 for rainfall). Analyses with different periods of data (e.g., 1982–2010) indicate the robustness of our results. The raw daily mean anomalies are computed relative to a daily climatology that is calculated as annual mean plus the first four harmonics of the 30-yr average (20-yr average for rainfall).

The MJO is defined following the commonly used real-time multivariate MJO index (RMM; Wheeler and Hendon 2004), which defines MJO evolution as an eight-phase cycle. The RMM indices are represented by the first two normalized principal components of the multivariate EOFs of combined equatorial U850, U200, and OLR. Daily RMM index values, including the phase and magnitude of the MJO, are obtained online (http://www.bom.gov.au/climate/mjo/). The RMM index has been widely used for the MJO monitoring and prediction. We note that there are other MJO indices (e.g., Kiladis et al. 2014; Kikuchi et al. 2012), which usually are highly correlated with the RMM indices [particularly during the boreal winter the correlation is around 0.75 (Kiladis et al. 2014; Kikuchi et al. 2012)]; however, if a prediction of such indices is widely made, such indices can also be used to assess potential predictability following the approach used in this analysis. Thus, all our calculations remain similar when applying the alternative MJO indices.

From the time series of RMM indices over the analysis period (1998–2017 for rainfall and 1988–2017 for others), the corresponding spatial patterns in other fields are obtained on the basis of regressions of unfiltered daily anomalies of respective fields against daily values of RMM indices. From the spatial regression patterns and RMM indices, daily values of MJO-related components are linearly reconstructed for the fields of interest. Variance percentage associated with MJO variability is then obtained with the daily reconstructed and raw fields. The ratio of variance between MJO-reconstructed fields and raw fields quantifies the estimate of the potential for predictive skill in the variable that is linked with the knowledge of the MJO. Note that the estimate thus obtained is based on a linear approach and, therefore, may underestimate the predictive skill.

3. Results

a. Analysis of variables included in the calculation of RMM index

The regression patterns of OLR and U850 with each of the two RMM indices are shown in Fig. 1. The spatial pattern replicates the familiar MJO-associated spatial patterns in the daily variability of the respective fields that are used in the computation of the RMM indices (Kiladis et al. 2014). The spatial pattern of OLR associated with the RMM1 has a minimum (implying enhanced convection) located over the Maritime Continent. This minimum in the OLR is also accompanied by low-level convergence as evident in the regression pattern for the U850.

The spatial pattern of OLR associated with the RMM2 has a maximum (implying suppressed convection) centered over the equatorial Indian Ocean and is in quadrature with the spatial pattern in the OLR associated with the RMM1. Consistent with suppressed convection, the associated U850 is indicative of low-level divergence.

The spatial patterns of regression patterns in OLR and zonal wind can be linearly combined to give the traditional composites associated with the MJO phases 1–8 (Wheeler and Hendon 2004). This is illustrated in Fig. 2 where linear combinations based on different sine and cosine weighting corresponding to the different phases of the MJO are shown. The
sine and cosine weightings (see the figure caption) are selected on the basis of the quadrant in which a particular phase of the MJO is defined. The OLR and U850 spatial patterns that are based on the linear combination have a good match with the traditional composites (e.g., Wheeler and Hendon 2004), and show a clear eastward propagation from phase 1 to phase 8.

We next compute the daily explained variance for the fields used in the computation of RMM indices on the basis of linear reconstruction. For each day, the OLR, U850, and U200 associated with the MJO are computed by (i) multiplying the regression patterns in Fig. 1 and the value of the RMM indices and (ii) subsequently summing up the individual contributions from RMM1 and RMM2. Once the reconstructed values for each day are obtained, the ratio of variance associated with the MJO versus variance of total field—that is, the fraction of daily variability that can be linearly associated with the MJO—is computed, shown in Fig. 3.

For the OLR, a maximum of approximately 15%–20% of daily variability is linearly associated with the knowledge of the MJO and the maximum is in the equatorial eastern Indian Ocean. The region of the maximum is collocated with the largest amplitude in the regression patterns for the OLR in Fig. 1. Of the three variables that go in the computation of MJO indices, the smallest explained variance is for the OLR.
This characteristic has been noted earlier (Straub 2013) and has led to effort in developing alternate MJO indices that have better link with the local precipitation (e.g., Ventrice et al. 2013).

The explained variance for U850 is larger than for the OLR with maximum value of 25%–30% over the “Maritime Continent.” Relative to the location in the maximum for the variance of OLR, the variance maximum for U850 are located east and west of it and a local minimum is collocated with the maximum in the OLR variance. This is consistent with the fact that the low-level convergence is collocated with precipitation with wind maxima to the east and west of it.

The RMM indices explain the largest amount of variance for U200. The regions of maximum explained variance, however, is shifted farther eastward and westward of the maximum variance explained in the OLR. A large shift in the maximum for U200 variance is consistent with eastward-propagating Rossby waves and westward-propagating Kelvin waves that are associated with a maximum in precipitation variability and OLR variability, respectively, over the eastern Indian Ocean. The difference in the spatial structure of explained variance between U850 and U200 is because U850 is also largely affected by boundary layer processes, which may influence the signature of Rossby and Kelvin waves.

The fraction of variances in Fig. 3 is an estimate of predictability for these variables given a perfect knowledge of MJO (as depicted by RMM indices). Note that the estimate is based on a linear approach, and hence, may represent an underestimate of predictability if the local response to the knowledge of MJO has appreciable nonlinearity. Besides this limitation, the fact that the explained variance of OLR (and hence, precipitation) is the smallest indicates that the predictive potential associated with RMM indices for precipitation may also be small and is localized in the vicinity of the Indian Ocean. It should also be recognized that for initialized predictions, the skill in forecasting MJO indices decreases with lead time, and therefore, a further decrease in the explained variance is to be expected during forecasts. As the fraction of explained variance (which is also a measure of signal-to-noise ratio) can be related to the prediction skill (Kumar and Hoerling 2000; Kumar et al. 2001), it also provides an estimate of skill in predicting OLR based on the prediction of the MJO.

Figure 4, by comparing the daily values with the reconstructed values, provides additional visual assessment of predictive potential of MJO. Time series and scatterplots of OLR and U200 are for averages over selected areas of maximum MJO-explained variance in Fig. 3. For the sake of clarity, the time series plots are shown only over 2007–10 whereas the scatterplots cover the entire analysis period.

In general, there is good correspondence between the daily raw time series of OLR and U200 and the corresponding
linearly reconstructed values. As expected, the reconstructed values have smoother temporal variations but capture the daily variability. For negative OLR anomalies (enhanced convection), comparisons indicate that reconstruction underestimates extreme negative values. This is confirmed in the scatterplot (Fig. 4c) where there is a systematic underestimation in the extreme negative OLR values. This is anticipated, as a linear approach like EOF followed by linear regression-based reconstruction of the fields may not be able to capture extreme values. Note also that the distribution of OLR anomalies itself is skewed with larger amplitude of negative anomalies compared to that of positive ones. This is congruent with the fact that the observed distribution of daily precipitation itself is skewed toward larger positive values (Thom 1951), and further, is consistent with the fact that precipitation is a positive definite variable and generally follows a gamma distribution. For U200, on the other hand, not only the distribution of daily U200 anomalies is more symmetric, the reconstructed U200 anomalies are able to capture extreme values. This may be because a larger fraction of daily U200 variability itself is captured by the MJO (Fig. 3).

One could argue that the forecasts that are based on the MJO will be conditional (often referred to as the forecast of opportunity) and the explained variance in Fig. 3 is an underestimation of predictive potential associated with the MJO when it is in an active phase. To address this, we estimate the explained variance only for the days when the absolute value of the amplitude of the RMM index is larger than 2. This procedure computes the variance of daily raw and reconstructed values only over a subset of days that satisfy the criteria that the amplitude of the MJO index is greater than 2. The results suggest that indeed a larger fraction of variance (shown in Fig. 5) is explained by stronger MJOs (note the difference in color scale relative to that in Fig. 3). The spatial patterns of explained variance, however, have the same spatial structure as that for in the unconditional case (Fig. 3), that is, computed over all days. Similar conclusions are also achieved when the analysis is done for days with

---

**Fig. 4.** (a),(b),(d),(e) Time series of original (black curves) and reconstructed (red curves) anomalous fields for (top) 2007–08 and (middle) 2009–10 and (c),(f) scatterplots of original (x axis) vs reconstructed (y axis) anomalous fields during winter 1988–2017 for (left) OLR (W m\(^{-2}\)) averaged over (10°S–10°N, 70°–100°E) and (right) U200 (m s\(^{-1}\)) averaged over (0°–20°N, 50°–90°E).
MJO amplitude larger than 1 (not shown). Note that the increase in variances comes at the expense of not assessing MJO’s predictive potential each day.

In Fig. 6, the explained variance (averaged between 15°S and 15°N) associated with different phases of the MJO is shown. Appreciable variance for the OLR remains confined between 60° and 150°E and propagates eastward with the MJO phase progression. The largest explained variance is associated with the MJO phases 3 and 7 in the eastern Indian Ocean. The largest amplitude of explained variance for U850 has a greater zonal extension and the maximum centers are located to the east and west of the OLR and move eastward with the phases of the MJO. Last, the spatial extent of explained variance has the widest spatial extent for U200, and as mentioned earlier, is due to the propagation of Rossby and Kelvin waves in the upper troposphere associated with the MJO convection over the eastern Indian Ocean.

We also estimated how the explained variance changes with time-averaging length. This helps quantify the time-scale dependence of MJO impacts. With an increase in time averaging, one would expect that the high-lead-frequency variability will get filtered out, leading to a reduction in the amplitude of the variability, while the variability associated with the slower MJO time scale may retain its amplitude. As the length of time averaging increases and spans across different phases of MJO (e.g., time averaging is done across opposite phases of the MJO), the amplitude of the component associated with the MJO may also decline. This might suggest that variances explained by the MJO may peak at a certain time-averaging length, and hence, for this averaging time length the knowledge of MJO will have the largest potential for predictive skill.

To illustrate this, the OLR variances of daily, 11-day, and 51-day time averaging for raw and reconstructed data are shown in Fig. 7. Going from daily (Fig. 7a) to 11-day average (Fig. 7b), the variance of raw data has a large decrease. A similar change in the amplitude of the reconstructed data is not seen (Fig. 7e vs Fig. 7d), indicating that the reconstructed anomalies are already filtered on MJO time scale and variations at synoptic time scales are largely removed from the reconstructed daily values. Increasing the averaging time to 51 days (Fig. 7c), the variance of raw data decreases further, and the largest OLR variance moves farther eastward to the central Pacific Ocean where the ENSO-related convection activity dominates on seasonal-to-interannual time scale. In the MJO-related reconstructions (Fig. 7f), the variance decreases as well, but it remains in the eastern tropical Indian and far western Pacific Oceans in association with the MJO activity.

The change in fraction of variance explained by the MJO against the averaging time is shown in Fig. 8. For short averaging times, the fraction of explained variance increases with
FIG. 6. Phase dependence of variances (%) explained by the RMM indices during the boreal winter (November–April), 1988–2017, for (a) OLR, (b) U850, and (c) U200, averaged over 15°S–15°N. Phase 1 is repeated as phase 9 for continuity of the display.

Explained Variances (%) by RMM index

FIG. 7. Variance of (a)–(c) original and (d)–(f) reconstructed OLR anomalies (W² m⁻⁴) during the boreal winter (November–April), 1988–2017, for (top) daily averaging, (middle) 11-day averaging, and (bottom) 51-day averaging.

Variance of Original & Reconstructed OLR
averaging time for all three variables, indicating that the process of time-averaging decreases the variability of raw data faster than it does for the MJO-associated reconstructed data. The ratio reaches a peak value at the averaging time of about 12–16 days for all three variables. Beyond that, further increases in averaging time sacrifices the amplitude of reconstructed signal, leading to a reduction in the fractional variance explained by the MJO. It is noted that the time-

FIG. 8. Explained variances (%) as a function of averaging time in days during the boreal winter (November–April), 1988–2017, for OLR (black curves), U850 (red curves), and U200 (green curves), averaged over the tropics (15°S–15°N).

FIG. 9. As in Fig. 3, but for (a) T2m and (b) precipitation during the boreal winter (November–April), 1998–2017.
The averaging period for which the fraction of explained variance maximizes may be location and variable dependent.

b. Analysis of surface temperature and precipitation

The above section analyzed the variables that themselves went in the derivation of RMM indices (Wheeler and Hendon 2004); in this section we will provide additional analyses about surface temperature and precipitation, which are of more societal relevance.

Figure 9 presents the fraction of daily T2m and precipitation that can be linearly reconstructed by the RMM indices. When compared with the variables used in the computation of RMM indices (Fig. 3), the explained variance is much smaller for T2m and precipitation (Fig. 9), generally less than 10%. For the T2m, the maximum is in the eastern tropical Africa and the two off-equatorial regions straddling the Maritime Continent, and the latter two regions also extend northeast into the extratropical Pacific oceans. The large explained variance might be due to atmospheric Rossby wave activity associated with the MJO-related convections. Over the land, except for the tropical Africa there are also local regions of larger variance explained in the middle Eurasian continent, Arabian Peninsula, part of Australia and western coastal and northeastern regions of the south American continent. For the north American continent, relatively larger value of variance is explained in the western coastal region and the Great Lake region, but the explained variance is clearly less than 5%.

For the precipitation, the explained variance is even smaller. A maximum of 5% daily precipitation variability can be linearly explained by the MJO over the region near the Maritime Continent. The region is collocated with the largest explained variance in OLR, but the explained variance is much smaller for precipitation (Fig. 9b vs Fig. 3a). Over the other regions, the MJO generally explains less than 1% of the variance of daily precipitation.

To explore the explained variance for stronger MJOs, Fig. 10 repeats the Fig. 9 calculation but based on the days when the amplitude of the RMM index is larger than 2. As for the variables included in the computation of RMM indices (Fig. 5), a larger fraction of variance is explained by stronger MJOs for T2m and precipitation as well (Fig. 10; note the difference in color scale relative to that in Fig. 9), but the spatial patterns of explained variance is similar to those for the unconditional case (Fig. 9).

Since the MJO teleconnection pattern is associated with the longitudinal positioning of tropical convections (see the review by Stan et al. 2017), the patterns of explained variance in T2m and precipitation could vary with different phases of the MJO. Figures 11 and 12 show the explained variances during eight phases of the MJO for T2m and precipitation, respectively. It
can be seen that the phase dependency of explained variance is evident for both T2m and precipitation, even though the explained variance itself is small. For T2m, in particular, the overall higher variance is explained during phases 3, 4, 7, and 8 than during the other four phases. For example, whereas for the MJO phases 3, 4, 7, and 8 5%–8% variance of daily T2m over tropical Africa is explained by the MJO, it only explains 3%–4% during other four phases; a similar difference also appears over the Eurasian continent. Over the continental United States, the phase dependency in T2m is slightly different; a larger region with relatively high explained variance is seen in the MJO phases of 2 and 6, and higher explained variance for the Great Lake region is seen in phases 1 and 5, in some consistency with the composite analyses by Zhou et al. (2012, their Fig. 3).

For the precipitation (Fig. 12), while the explained variance is generally maximized over the region near the Maritime Continent, the spatial details and amplitudes of explained variance vary with different phases of the MJO. In phases 2 and 6, there are generally two detached local maxima located in the central equatorial Indian Ocean and the Maritime Continent, respectively; in phases 1 and 5, there are also two local maxima but the one in the central Indian Ocean is much weaker; in other phases, there is basically one local maximum. Furthermore, the local maximum also presents an eastward propagation in association with the migration of MJO-related
convection. All these features are in good consistency with the OLR features, but the explained variance is much smaller for precipitation (Fig. 12 vs Fig. 6a).

4. Summary and future work

Continual skill improvements in predictions of the MJO have occurred in recent decades. For the ECMWF, sub-seasonal prediction system, for example, the MJO skill as measured by correlation between the observed and the predicted bivariate RMM index now exceeds 0.5 beyond 30 days (Vitart 2017). As further improvements in skill in predicting the MJO are likely to occur, what improvements in skill of meteorological variables of societal relevance should be expected is a question of practical relevance. However, with the preponderance of quantification of MJO skill based on the assessment of MJO indices themselves, it is not straightforward to assess how those improvements will translate into skill in predicting other variables. To provide an assessment, one can take an idealized approach by assuming that the future information about the MJO is perfectly known and explore what fraction of variability in the meteorological variables can be associated with the MJO, and further, how this fraction changes with the time-averaging. These questions are addressed for the variables that go into the computation of the RMM indices. Further, fraction of variability in the T2m and
precipitation associated with the MJO was also calculated. Note that calculating explained variances cannot give an estimate about how many days of skillful prediction can be made, but it gives an estimate about how much variance of a variable is potentially predictable. In addition, as compared with predictability studies that are based on model forecasts or reforecasts, our observational approach provides a much more objective measure of predictability that is independent of model biases (and their influence on skill).

In future, we will follow the analysis strategy for other variables: for example, vertical wind shear and tropospheric moisture content. Using lead–lag analysis, it may be possible to quantify the propagation characteristics of tropical to extratropical linkages associated with the MJO. Such an analysis is valuable in managing our expectations with regard to the upper limits of potential predictability in the meteorological variables associated with the MJO. Assessment of potential predictability, however, needs to be adjusted by skill in predicting the MJO itself, which undoubtedly will decay with increasing lead time, and therefore predict skill will be smaller than the upper limits of predictability that we assessed in this analysis.

Acknowledgments. We thank NOAA’s Climate Program Office for their support. Author J. Zhu is partially supported by the NASA Ocean Salinity Science Team Grant NNX17AK09G.

REFERENCES


Kessler, W. S., and R. Kleeman, 2000: Rectification of the MJO itself, which undoubtedly will decay with increasing lead time, and therefore predict skill will be smaller than the upper limits of predictability that we assessed in this analysis.

Acknowledgments. We thank NOAA’s Climate Program Office for their support. Author J. Zhu is partially supported by the NASA Ocean Salinity Science Team Grant NNX17AK09G.

REFERENCES


