

Improving the prediction of the Madden-Julian Oscillation of the ECMWF model by post-processing

Riccardo Silini¹, Sebastian Lerch², Nikolaos Mastrantonas³, Holger Kantz⁴, Marcelo Barreiro⁵, and Cristina Masoller⁶

¹UPC Universitat Politècnica de Catalunya

²Karlsruhe Institute of Technology

³European Centre for Medium Range Weather Forecasts

⁴Max Planck Institute for the Physics of Complex Systems, Dresden, Germany

⁵Facultad de Ciencias, Universidad de la Republica, Uruguay

⁶Universitat Politecnica de Catalunya

November 24, 2022

Abstract

The Madden-Julian Oscillation (MJO) is a major source of predictability on the sub-seasonal (10- to 90-days) time scale. An improved forecast of the MJO, may have important socioeconomic impacts due to the influence of MJO on both, tropical and extratropical weather extremes. Although in the last decades state-of-the-art climate models have proved their capability for forecasting the MJO exceeding the 5 weeks prediction skill, there is still room for improving the prediction. In this study we use Multiple Linear Regression and an Artificial Neural Network as post-processing methods to improve one of the currently best dynamical models developed by the European Centre for Medium-Range Weather Forecast (ECMWF). We show that the post-processing with the machine learning algorithm employed leads to an improvement of the MJO prediction. The largest improvement is in the prediction of the MJO geographical location and intensity.

Improving the prediction of the Madden-Julian Oscillation of the ECMWF model by post-processing

Riccardo Silini¹, Sebastian Lerch², Nikolaos Mastrantonas^{3,4}, Holger Kantz⁵,
Marcelo Barreiro⁶, Cristina Masoller¹

¹Universitat Politècnica de Catalunya, Edifici Gaia, Rambla Sant Nebridi 22, 08222 Terrassa, Barcelona, Spain.

²Institute of Economics, Karlsruhe Institute of Technology, Blücherstr. 17, 76185 Karlsruhe, Germany.

³European Centre for Medium-Range Weather Forecast (ECMWF), Reading, UK.

⁴Technische Universität Bergakademie Freiberg (TUBAF), Freiberg, Germany.

⁵Max-Planck Institute for the Physics of Complex Systems, 01187 Dresden, Germany.

⁶Departamento de Ciencias de la atmósfera, Facultad de Ciencias, Universidad de la República, Iguazú 4225, 11400 Montevideo, Uruguay.

Key Points:

- We use a machine learning algorithm to correct the ECMWF prediction of the Madden-Julian Oscillation.
- We obtain improved predictions of both MJO amplitude and phase exceeding those obtained using a multiple linear regression method.
- Results suggest a tendency of machine learning to overcome the Maritime Continent barrier.

Corresponding author: Riccardo Silini, riccardo.silini@outlook.com

Abstract

The Madden-Julian Oscillation (MJO) is a major source of predictability on the sub-seasonal (10- to 90-days) time scale. An improved forecast of the MJO, may have important socioeconomic impacts due to the influence of MJO on both, tropical and extratropical weather extremes. Although in the last decades state-of-the-art climate models have proved their capability for forecasting the MJO exceeding the 5 weeks prediction skill, there is still room for improving the prediction. In this study we use Multiple Linear Regression and an Artificial Neural Network as post-processing methods to improve one of the currently best dynamical models developed by the European Centre for Medium-Range Weather Forecast (ECMWF). We show that the post-processing with the machine learning algorithm employed leads to an improvement of the MJO prediction. The largest improvement is in the prediction of the MJO geographical location and intensity.

Plain Language Summary

The Madden-Julian Oscillation (MJO) has important socioeconomic impacts due to its influence on both, tropical and extratropical weather extremes. Although in the last decades state-of-the-art climate models have proved their capability for forecasting the MJO exceeding the 5 weeks prediction skill, there is still room for improvement. In this study we use artificial intelligence to correct the predictions provided by one of the currently best climate models, developed by the European Centre for Medium-Range Weather Forecast (ECMWF). We show that artificial intelligence leads to an improved prediction of the MJO geographical location and intensity.

1 Introduction

The Madden-Julian Oscillation (MJO) with its 30- to 60-day oscillation is the major sub-seasonal fluctuation in tropical weather (Madden & Julian, 1971, 1972; Vitart, 2009; Lau & Waliser, 2011; Zhang et al., 2013; Ferranti et al., 2018). It is the main source of intra-seasonal fluctuations in the Indian monsoon (Taraphdar et al., 2018; Díaz et al., 2020), and is also known to modulate the tropical cyclogenesis (Camargo et al., 2009; Klotzbach, 2010; Fowler & Pritchard, 2020), to have a two-way interaction with El Niño-Southern Oscillation (ENSO) (Bergman et al., 2001), to influence the Asian-Australian monsoon (Wheeler et al., 2009), and be influenced by the quasi-biennial oscillation (Wang et al., 2019; Martin, Son, et al., 2021). Moreover, the MJO not only affects the tropical weather, but also the extra-tropical weather through teleconnections (Alvarez et al., 2017; Ungerovich et al., 2021). Therefore, MJO has a large impact on the economy, society, and agriculture, motivating the wide interest in its prediction.

Many efforts have been made in this direction in the last decades, with dynamical models leading to the current best forecasts (Jiang et al., 2020), but despite the continuous progress of the dynamical models, there is still room for improvement in the MJO prediction (Zhang et al., 2013; Jiang et al., 2020).

In particular, an improvement of the prediction skill when MJO crosses the Maritime Continent (MC) barrier (C.-H. Wu & Hsu, 2009; H.-M. Kim et al., 2016; Barrett et al., 2021) will be of practical importance due to the influence of MJO on ENSO, as an improved MJO prediction may contribute to improving the prediction of ENSO.

Machine Learning (ML) algorithms have been extensively used in many fields, and they are gaining a foothold in weather and climate modeling (O’Gorman & Dwyer, 2018; Dijkstra et al., 2019; Tseng et al., 2020; Gagne II et al., 2020; Silini et al., 2021) among others. Although MJO predictions obtained using ML models do not outperform dynamical models (Silini et al., 2021; Martin, Barnes, & Maloney, 2021), a hybrid approach, combining dynamical models and ML techniques, may improve the results. In this way,

68 it is possible to use dynamical models that have been developed across decades, based
 69 on physical phenomena, in combination with data-driven ML techniques, an approach
 70 that has shown its ability to reduce the gap between observations and dynamical mod-
 71 els' forecasts (Rasp & Lerch, 2018; McGovern et al., 2019; Scheuerer et al., 2020; Haupt
 72 et al., 2021; Vannitsem et al., 2021).

73 Recently, it has been shown that bias correction in linear dynamics (J. Wu & Jin, 2021),
 74 and using Long Short-Term Memory (LSTM) networks on multi-model means (H. Kim
 75 et al., 2021), improved the MJO prediction.

76 Currently, the best forecast dynamical model in terms of MJO prediction skill is the one
 77 developed by the European Centre for Medium-Range Weather Forecast (ECMWF) (Jiang
 78 et al., 2020). Therefore, in this study we attempt to improve ECMWF forecasts by us-
 79 ing Multiple Linear Regression (MLR) and Artificial Neural Networks (ANNs), as post-
 80 processing tools.

81 To quantify the forecast skill we use four metrics, namely the bivariate correlation co-
 82 efficient (COR), the bivariate root-mean-square error (RMSE), with threshold values COR=0.5
 83 and RMSE=1.4, as well as the amplitude error and the phase error (Rashid et al., 2011).

84 We apply the post-processing ML techniques to the ensemble mean of ECMWF, and we
 85 show that ANNs is able to correct the dynamical model forecasts leading to an improved
 86 MJO prediction. In particular, it improves the prediction of the MJO over the MC and
 87 its amplitude, while the phase errors obtained with the two post-processing techniques
 88 are similar.

89 2 Data, Methods and Models

90 2.1 RMM Data

91 For this study, we use the Real-time Multivariate MJO (RMM) index (Wheeler & Hen-
 92 don, 2004) as labels for the supervised learning method, which is used to characterize
 93 the MJO geographical position and intensity. The first two principal components of the
 94 combined empirical orthogonal functions (EOFs) of outgoing longwave radiation (OLR),
 95 zonal wind at 200 and 850 hPa averaged between 15°N and 15°S are labeled RMM1 and
 96 RMM2. With a polar transformation, it is possible to define the MJO phase and am-
 97 plitude. The phase is divided into 8 classes, each corresponding to a different sector of
 98 the phase diagram defining the observed MJO life cycle. The amplitude, describing the
 99 MJO intensity, when smaller than 1 defines a non-active MJO. The ERA5 RMM1 and
 100 RMM2 from 13th June 1999 to 29th June 2019 were downloaded from (*ECMWF RMM*
 101 *reforecasts data*, 2021). This time window is selected to match the ECMWF reforecasts,
 102 presented in the previous section.

103 2.2 ECMWF RMM reforecasts

104 The samples used as input for the ANN and to assess the model performance, are ECMWF
 105 reforecasts with Cyrcl 46r1 freely available from (*ECMWF RMM reforecasts data*, 2021).
 106 This dataset is composed of 110 initial dates per year for 20 years, between the 13th June
 107 1999 and the 29th June 2019. In total there are 2200 starting dates, from which a 46-
 108 lead-days prediction is available. The dataset provides the prediction of four variables:
 109 the first two principal components of the RMM index, and their polar transformation.
 110 For each starting day and variable there are 12 time series of 46 points. One is the con-
 111 trolled forecast (cf) corresponding to a forecast without any perturbations, then there
 112 are 10 perturbed forecasts members (pf) which have slightly different initial conditions
 113 from the cf to take into consideration errors in observations and the chaotic nature of
 114 weather. Finally there is the ensemble mean (em), which corresponds to the mean of the
 115 11 members (cf + 10 pf). In this particular study, we made use solely of the em data

116

2.3 Prediction skill

117
118
119
120

To know how good a model is in predicting, we present here the metrics that will be used. For sake of comparison, we use the same metrics adopted in (H. Kim et al., 2018), which are adapted from (Lin et al., 2008; Rashid et al., 2011), where they define the COR and RMSE as follows:

$$\text{COR}(\tau) = \frac{\sum_{t=1}^N [a_1(t)b_1(t, \tau) + a_2(t)b_2(t, \tau)]}{\sqrt{\sum_{t=1}^N [a_1^2(t) + a_2^2(t)]} \sqrt{\sum_{t=1}^N [b_1^2(t, \tau) + b_2^2(t, \tau)]}}, \quad (1)$$

$$\text{RMSE}(\tau) = \sqrt{\frac{1}{N} \sum_{t=1}^N [|a_1(t) - b_1(t, \tau)|^2 + |a_2(t) - b_2(t, \tau)|^2]}, \quad (2)$$

121
122
123
124
125
126

where $a_1(t)$ and $a_2(t)$ correspond to the observed RMM1 and RMM2 at time t , while $b_1(t, \tau)$ and $b_2(t, \tau)$ will be the respective forecasts for time step t for a lead time of τ days, and N is the number of forecasts. The bivariate correlation coefficient expresses the strength of the linear relationship between the forecasts and the observations, while the root-mean-square error compares the difference between the values of the forecasts and the observations.

127
128
129
130

In this study we use $\text{COR}=0.5$ and $\text{RMSE}=1.4$ as prediction skill thresholds (Rashid et al., 2011). The RMM prediction skill is defined as the time in which the COR takes a value below 0.5 and RMSE gets above 1.4. For a given lead time, the COR and RMSE are the average value up to that lead time.

131

2.4 Amplitude and phase error

132
133
134

To characterize the MJO it is convenient to perform a change of coordinates from cartesian to polar (RMM1, RMM2) \rightarrow (A, φ). The MJO amplitude $A(t)$, describing its intensity, can be written as:

$$A(t) = \sqrt{RMM1^2(t) + RMM2^2(t)}, \quad (3)$$

135
136

while the MJO phase $\varphi(t)$, describing the geographical position of the enhanced rainfall region center, can be written as:

$$\varphi(t) = \tan^{-1} \left(\frac{RMM2(t)}{RMM1(t)} \right). \quad (4)$$

137
138

By definition (Rashid et al., 2011), the amplitude error for a given lead time $E_A(\tau)$ can be expressed as

$$E_A(\tau) = \frac{1}{N} \sum_{t=1}^N (A_{pred}(t, \tau) - A_{obs}(t)), \quad (5)$$

139
140
141

where N represents the number of predicted days, $A_{obs}(t)$ is the observations amplitude at time t and $A_{pred}(t, \tau)$ is the predictions amplitude at time t with a lead time of τ days. The phase error $E_\varphi(\tau)$ is defined by

$$E_\varphi(\tau) = \frac{1}{N} \sum_{t=1}^N \tan^{-1} \left(\frac{a_1(t)b_2(t, \tau) - a_2(t)b_1(t, \tau)}{a_1(t)b_1(t, \tau)} \right), \quad (6)$$

142
143

where $a_1(t)$ and $a_2(t)$ correspond to the observed RMM1 and RMM2 at time t , while $b_1(t, \tau)$ and $b_2(t, \tau)$ correspond to the predicted RMM1 and RMM2 at time t with a lead

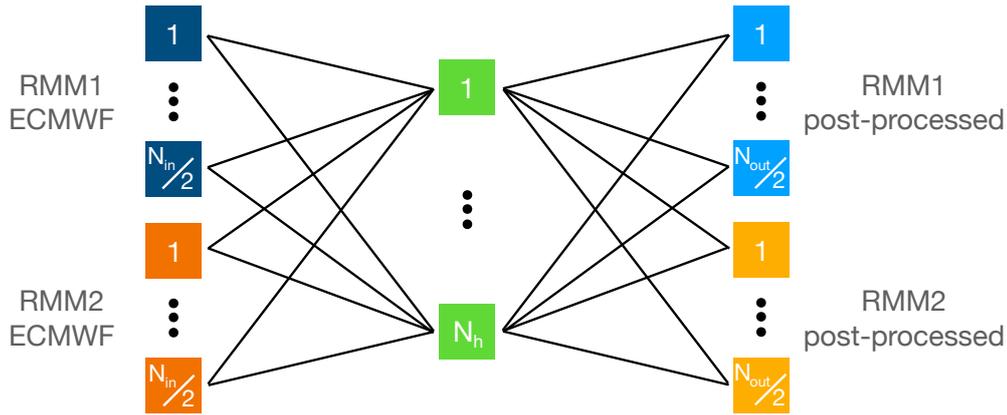


Figure 1. ANN architecture employed for this study.

144 time of τ days. These two metrics allow us to analyze in more detail the model perfor-
 145 mance to predict the MJO, in conjunction with COR and RMSE.

146 2.5 Post-processing methods

147 The post-processing machine learning tool built for this study is a fully connected feed-
 148 forward neural network (FFNN) composed of an input layer containing N_{in} neurons, a
 149 single hidden layer with N_h neurons, and an output layer with N_{out} neurons, as shown
 150 in Fig 1. The activation function used is the Rectified Linear Unit (ReLU), which trans-
 151 forms the weighted sum of the input values by returning 0 in case of a negative-sum, and
 152 the result of the sum otherwise. Dealing with a supervised regression problem, the mean-
 153 squared error (MSE) is extensively employed as loss function, and it is used in the frame-
 154 work of this study to compare the neural network output with the observations (labels).
 155 An adaptive optimizer (Adam) is selected to automatically manage the learning rate dur-
 156 ing the training phase.

157 We use an adaptive number of neurons depending on the number of days we want to fore-
 158 cast. The ECMWF reforecasts provide predictions up to a lead time of 46 days for both
 159 RMM1 and RMM2, and we build a different network for each lead time. This means that
 160 the number of output neurons N_{out} can fall between 2 and 92 because we use both RMM1
 161 and RMM2.

162 After selecting the number of output neurons (which is even and in fact defines our lead
 163 time, $\tau = N_h/2$), we adapt the number of input N_{in} and hidden neurons N_h as follows.
 164 As input, the networks receive the ECMWF reforecasts, which also limit the number of
 165 input neurons N_{in} in the range between 2 and 92. After training the networks with dif-
 166 ferent N_{in} , we found the best result is obtained with $N_{in} = N_{out} + 6$ with an upper
 167 limit of 92. Using all 92 inputs, the prediction skill for short lead times slightly decreases.
 168 For simplicity, a fixed number of 92 inputs could also be used. An interpretation of the
 169 reason behind this result is that to correct the prediction values for a given day, the fu-
 170 ture predicted values can help up to some extent. To correct the prediction of a given
 171 day, for each RMM we use the predicted values of up to 3 days after that particular day.
 172 To avoid overfitting, we want the number of hidden neurons to be relatively small, for
 173 this reason after some tests, we select $N_h = N_{in}/2$. The training has been performed
 174 over 100 epochs which allows to not overfit the model, and the model performance is tested
 175 using a walk-forward validation, where we found the best minimum number of samples,
 176 out of 2200 available, to be 1700. From the minimum number of samples, the train set

177 is then extended by 100 samples (~ 1 year) for each run and tested on the subsequent
 178 200 samples (~ 2 years). Other methods to avoid overfitting could also be used, such
 179 as early stopping or drop-out.

180 The MLR corrects the RMMs separately, using the ECMWF predictions as input and
 181 the ERA5 observations as output, with the same walk-forward validation used for the
 182 ANNs.

183 **3 Results and Analysis**

184 The first part of this section will be devoted to the results obtained for the MJO am-
 185 plitude and phase. In the second part we present the prediction skill assessment using
 186 the COR 0.5 level, and RMSE 1.4 level as metrics, while in the last part of the section
 187 we show how the different forecast methods perform for different MJO initial phases.

188 The results are obtained training the ANN from 13th of June 1999 using a walk-forward
 189 validation, and averaging the error obtained by testing over different unseen time win-
 190 dows from 5th December 2014 to 29th June 2019. The size of the windows is defined by
 191 the selected number of initial days from which the ECMWF forecast starts. Due to the
 192 bi-weekly acquisition of ECMWF, this means that each window of 200 points corresponds
 193 to 2 years approximately. Each member of the ensemble over which the average is per-
 194 formed, corresponds to a test set used for the walk-forward validation. Different sizes
 195 of the test set between 100 and 500 samples have been tested, leading to prediction skills
 196 that vary sensibly. For this reason, it is important to take into account that results may
 197 vary depending on the test set and its size, albeit preserving the same general result: the
 198 post-processing corrections improve the ECMWF forecasts.

199 In Fig. 2, we show the error on the MJO amplitude for events starting with an ampli-
 200 tude larger than 1. We can notice an underestimation of the amplitude as expected (Jiang
 201 et al., 2020). Nevertheless, the post-processed amplitudes are closer to the observed ones,
 202 with respect to the raw ECMWF forecast. The maximum improvement occurs for a lead
 203 time of 28 days when the ECMWF-ANN model has a RMSE similar to the RMSE of
 204 the uncorrected ECMWF at a lead time of 20 days.

205 By the definition of the amplitude error, errors of opposite sign could potentially can-
 206 cel out resulting in misleading conclusions. For this reason in Fig. 2 we also provide the
 207 RMSE of the amplitude error, which shows a similar behavior as before. Both post-processing
 208 techniques improve the results, with the ANN bringing the highest benefits in terms of
 209 the magnitude of error reduction, and the forecasting horizon of the improvement.

210 In Fig. 3, we present the MJO phase error. The post-processing techniques provide an
 211 improved prediction, during which all three models predict a negative phase. A positive
 212 phase error indicates a faster propagating MJO, while a negative error represents a slower
 213 propagation. The ECMWF forecast shows an overall slower propagation of the MJO with
 214 respect to the observations, and both post-processing corrections provides an increment
 215 of the MJO speed prediction. In particular, at the 18 days lead time we can notice an
 216 increment of the ECMWF phase error, which MLR and ML tend to correct.

217 Fig. 4, shows the COR and RMSE of the ECMWF ensemble mean forecasts, the MLR,
 218 and ANN post-processing. A COR of 0.5 is taken here as baseline for useful prediction
 219 skill. It is possible to notice that the improvement provided by both post-processing meth-
 220 ods slightly increases with the lead time. ML post-processing is the overall best of the
 221 three methods. All models COR are overlapped up to 2 weeks, from which the ML post-
 222 processing diverges from the other two up to 3 weeks, when the MLR diverges from the
 223 ECMWF prediction, joining the ML prediction up to 5 weeks. Using the previously se-
 224 lected parameters used to build and train the ANN architecture (number of neurons per
 225 layer, minimum sample size, test set size, etc.), Fig. 4 displays a prediction skill improve-

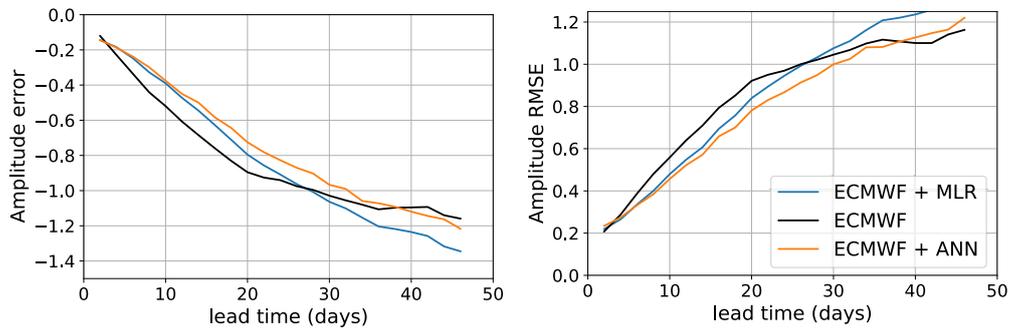


Figure 2. MJO amplitude error (left) and amplitude RMSE (right) as a function of the lead time for events starting with an amplitude larger than 1. The color indicates the forecast model, the black line corresponds to the ECMWF forecast, the blue line corresponds to the MLR correction of the ECMWF forecast, while the orange line corresponds to the post-processed ECMWF forecast with an ANN.

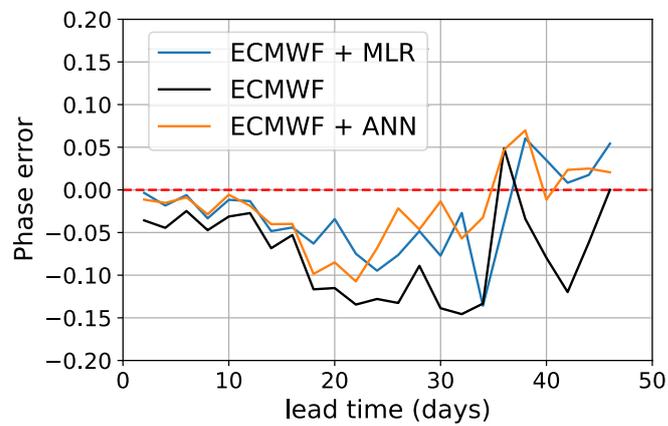


Figure 3. MJO phase error for events starting with an amplitude larger than 1. The color indicates the forecast model, the black line corresponds to the ECMWF forecast, the blue line corresponds to the MLR correction of the ECMWF forecast, while the orange line corresponds to the post-processed ECMWF forecast with an ANN.

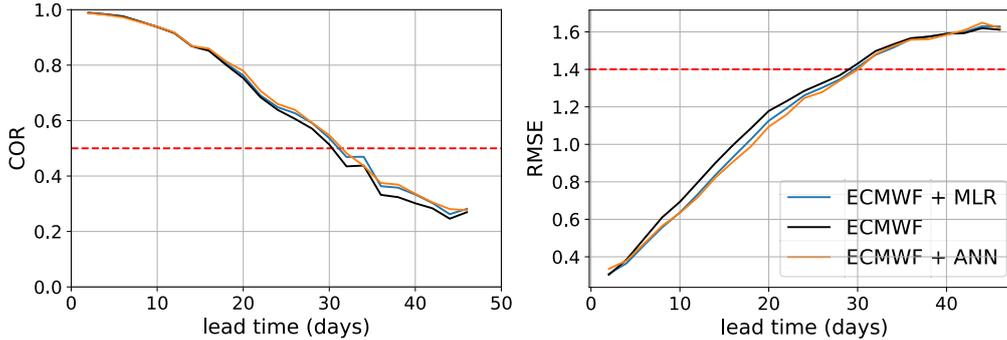


Figure 4. COR (left) and RMSE (right) as a function of the forecast lead time for events starting with an amplitude larger than 1. The color indicates the forecast model and the red dashed line indicates the prediction skill threshold of $\text{COR}=0.5$ and $\text{RMSE}=1.4$. The black line corresponds to the ECMWF forecast, the blue line corresponds to the post-processed ECMWF forecast with MLR, while in orange it is shown the post-processed ECMWF forecast with an ANN.

226 ment at the $\text{COR}=0.5$ level of about 1 day. In terms of RMSE, up to a lead time of 4
 227 weeks, neither post-processing technique crosses the RMSE-threshold of 1.4, and they
 228 both improve the prediction skill with respect to the ECMWF model alone.

229 In Fig. 5, we display the comparison between the observations, the ECMWF forecast,
 230 and its corrections, in a Wheeler-Hendon phase diagram for two different starting dates
 231 of the same MJO event. The dots are marked every 7 days to identify the weeks. In the
 232 left panel, the 3 weeks prediction starts on the 21st November 2018 and displays its pro-
 233 gression from the Western Hemisphere over the Indian Ocean. It is possible to notice that
 234 both post-processing techniques display very similar prediction, with a slightly larger am-
 235 plitude than ECMWF, closer to the observations for all lead times. In the right panel,
 236 the 3 weeks prediction starts on the 5th December 2018 in the Indian Ocean. We can
 237 see a drop of accuracy in the ECMWF prediction, and the MLR post-processing, approach-
 238 ing the MC. The ML correction instead preserves a larger amplitude, closer to the ob-
 239 servations.

240 It is also possible to notice that while the speed of the MJO event is well predicted in
 241 the left panel, in the right one there is a drop of the MJO speed forecast over the Indian
 242 Ocean and MC.

243 Here we presented an example of a strongly active MJO event, where the corrections clearly
 244 improve the ECMWF prediction and it is among the best found. All predictions from
 245 the 12th of December 2014 to the 18th of June 2019, can be found in (Silini, 2021b). Look-
 246 ing at these results it is possible to appreciate the general improvement provided by the
 247 post-processing corrections.

248 Finally we study the amplitude error, the phase error, the COR, and RMSE, as a func-
 249 tion of the different initial phases of MJO. As displayed in Fig. 6, applying post-processing
 250 methods improves the amplitude error for all initial phases. The MLR provides an im-
 251 provement with respect to the ECMWF model, but the ML correction leads to the low-
 252 est error. Concerning the initial phases, we find the lowest amplitude error when an MJO
 253 event starts over the MC, while the largest is found in phase 2, over the Indian Ocean.
 254 With the MJO propagating at an average speed of 5 ms^{-1} , events starting in phase 2
 255 will cross the MC in 2-3 weeks time (H.-M. Kim et al., 2014). The phase displays

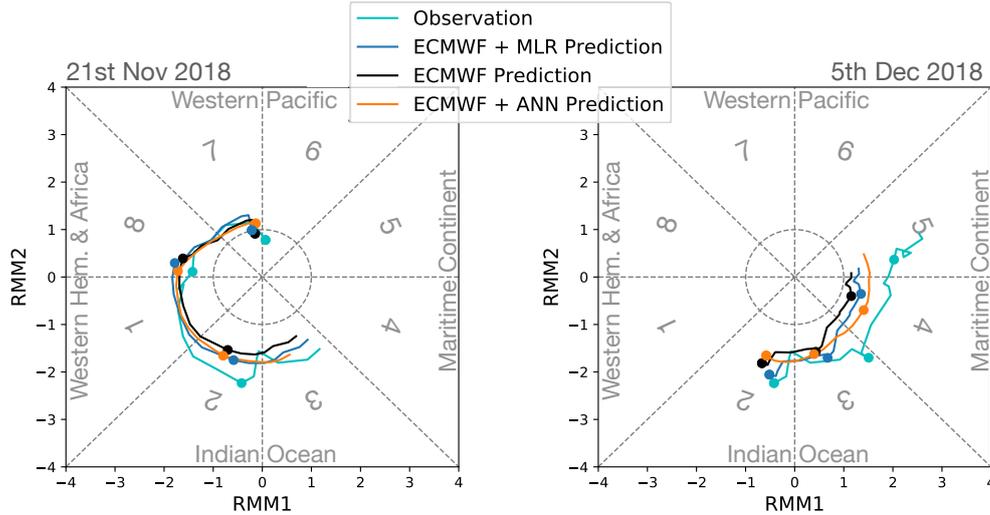


Figure 5. Wheeler-Hendon phase diagram for two different starting dates of the same MJO event, and a 3 weeks prediction. The left panel starting date is the 21st November 2018. The MJO enhanced rainfall region travels across the western Hemisphere and Indian Ocean. The right panel starting date is the 5th of December 2018, and represents a 3 weeks prediction approaching and traveling over the MC. The rotation of the event in the phase diagram is counter-clockwise, and the dots are included every 7 days, marking the different weeks.

256 a large worsening of the MJO localization prediction, when the forecast starts between
 257 the MC and Western Pacific (phase 6-8). This observation is consistent with Fig. 5, where
 258 we noticed a drop in the accuracy of the MJO speed prediction over the Indian Ocean
 259 and MC. The COR finds its maximum when starting over the MC continent, consistently
 260 with the amplitude error. The ML correction has the highest COR except for phase 8,
 261 where MLR leads to the highest one. The RMSE is very consistent with the COR, in
 262 which we find the the minimum in phase 4, and the ML correction having the lowest er-
 263 ror, except for phase 8. Overall, we can conclude that the ML post-processing is worth
 264 applying especially to reduce the error on the amplitude prediction, while MLR could
 265 be useful for a better prediction of the MJO location.

266 4 Conclusions

267 We employed a MLR and a ML algorithm to perform a post-processing correction of the
 268 prediction of the dynamical model that currently holds the highest MJO prediction skill (Jiang
 269 et al., 2020), developed by ECMWF.

270 The largest improvement is found in the MJO amplitude and phase individually, which
 271 decreases the underestimation of the amplitude, providing a more accurate predicted ge-
 272 ographical location of the MJO. The amplitude and phase estimation are improved for
 273 all lead times up to 5 weeks.

274 We obtained an improved prediction skill of about 1 day for a COR of 0.5.

275 Plotting the forecasts in a Wheeler-Hendon phase diagram we found an improvement pre-
 276 dicting the MJO propagation, notably across the MC, which helps overcome the MC bar-
 277 rier.

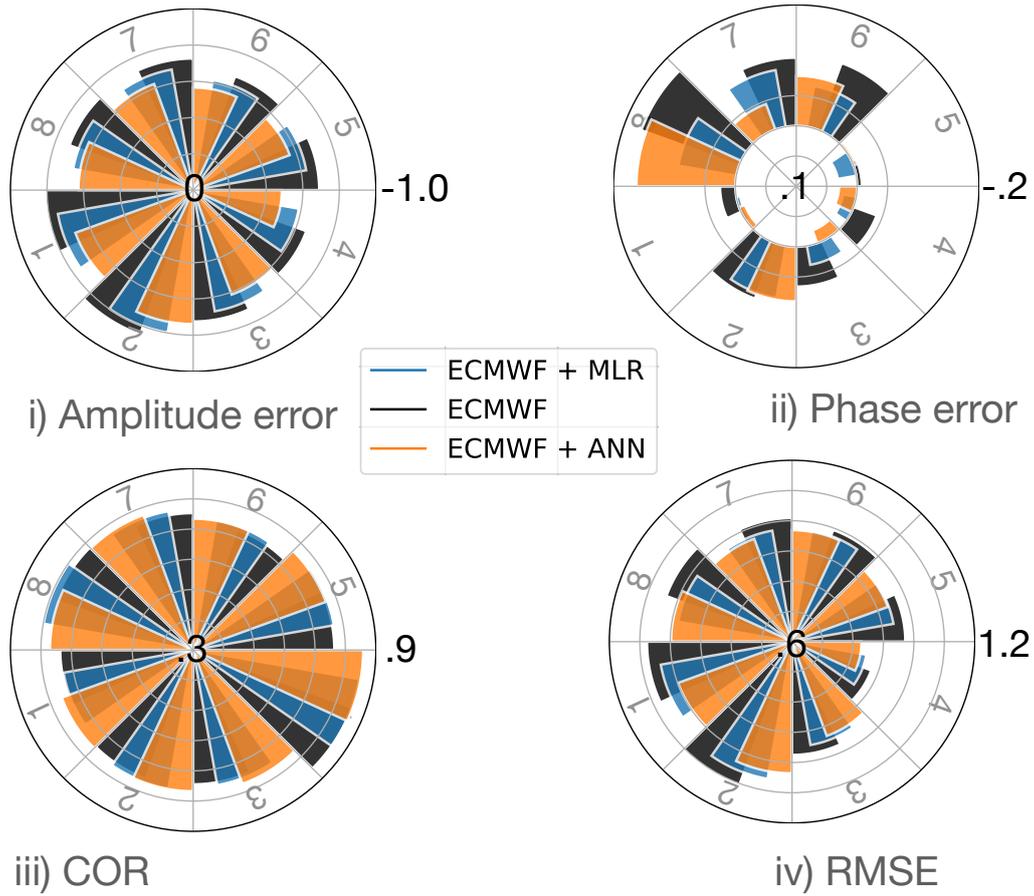


Figure 6. Amplitude error, phase error, COR, and RMSE, for the different MJO initial phases, for events starting with an amplitude larger than 1. The plots show the mean for lead times up to 5 weeks. The different colors represent the different prediction methods.

278 Considering the results obtained for each initial MJO phase, we found that both post-
 279 processing tools improve the prediction, with the ML correction being the best.

280 The ML technique provides an improvement over MLR for all initial phases except phase
 281 8. In the case of phase forecast it might be also sufficient to use MLR instead of ML. This
 282 suggests a predominance of linear corrections to improve the MJO phase forecast.

283 This study confirms the potential of post-processing techniques to reduce the knowledge
 284 and bias gap between dynamical models forecasts and observations, providing advance-
 285 ment in MJO prediction.

286 As future work, it would be interesting to test a stochastic approach to post-processing
 287 (as in (Rasp & Lerch, 2018), which would allow to obtain a probabilistic forecast

288 Although the improvement provided by the MLR and ML techniques, a post-processing
 289 method will always strongly rely on the accuracy of the dynamical model's forecasts. For
 290 this reason, it is crucial to work on both dynamical models and machine learning meth-
 291 ods to progress.

292 **5 Data availability**

293 The RMM data, and the ECMWF reforecasts can be freely downloaded from (*ECMWF*
 294 *RMM reforecasts data*, 2021).

295 **6 Code availability**

296 The Keras TensorFlow (Abadi et al., 2015) trained FFNN can be found in (Silini, 2021a).

297 **7 Acknowledgments**

298 The authors thank Linus Magnusson for providing access to the data. This work received
 299 funding from the European Union's Horizon 2020 research and innovation programme
 300 under the Marie Skłodowska Curie Grant Agreement No 813844. C.M. also acknowledges
 301 funding by the Spanish Ministerio de Ciencia, Innovacion y Universidades (PGC2018-
 302 099443-B-I00) and the ICREA ACADEMIA program of Generalitat de Catalunya. S.L.
 303 acknowledges support by the Vector Stiftung through the Young Investigator Group "Ar-
 304 tificial Intelligence for Probabilistic Weather Forecasting

305 **8 Author contributions**

306 R. S. performed the analysis and prepared the figures; R. S. and S. L. and N. M. designed
 307 the study; all authors discussed the results, wrote and reviewed the manuscript

308 **9 Competing interests**

309 The authors declare no competing interests.

310 **References**

- 311 Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., . . . Zheng,
 312 X. (2015). *TensorFlow: Large-scale machine learning on heterogeneous sys-*
 313 *tems*. Retrieved from <http://tensorflow.org/> (Software available from
 314 tensorflow.org)
 315 Alvarez, M. S., Vera, C. S., & Kiladis, G. N. (2017). Mjo modulating the activity
 316 of the leading mode of intraseasonal variability in south america. *Atmosphere,*

- 8(12). doi: <https://doi.org/10.3390/atmos8120232>
- Barrett, B. S., Densmore, C. R., Ray, P., & Sanabia, E. R. (2021). Active and weakening mjo events in the maritime continent. *Climate Dynamics*. doi: <https://doi.org/10.1007/s00382-021-05699-8>
- Bergman, J. W., Hendon, H. H., & Weickmann, K. M. (2001). Intraseasonal air–sea interactions at the onset of el niño. *Journal of Climate*, *14*(8), 1702–1719.
- Camargo, S. J., Wheeler, M. C., & Sobel, A. H. (2009). Diagnosis of the mjo modulation of tropical cyclogenesis using an empirical index. *Journal of the Atmospheric Sciences*, *66*(10), 3061 - 3074. doi: <https://doi.org/10.1175/2009JAS3101.1>
- Díaz, N., Barreiro, M., & Rubido, N. (2020). Intraseasonal predictions for the south american rainfall dipole. *Geophysical Research Letters*, *47*(21). doi: <https://doi.org/10.1029/2020GL089985>
- Dijkstra, H. A., Petersik, P., Hernández-García, E., & López, C. (2019). The application of machine learning techniques to improve el niño prediction skill. *Frontiers in Physics*, *7*, 153. doi: <https://doi.org/10.3389/fphy.2019.00153>
- ECMWF RMM reforecasts data*. (2021). <https://acquisition.ecmwf.int/ecpds/data/list/RMMS/ecmwf/reforecasts/>. (Accessed: 2021-02)
- Ferranti, L., Magnusson, L., Vitart, F., & Richardson, D. S. (2018). How far in advance can we predict changes in large-scale flow leading to severe cold conditions over europe? *Quarterly Journal of the Royal Meteorological Society*, *144*(715), 1788-1802. doi: <https://doi.org/10.1002/qj.3341>
- Fowler, M. D., & Pritchard, M. S. (2020). Regional mjo modulation of northwest pacific tropical cyclones driven by multiple transient controls. *Geophysical Research Letters*, *47*(11), e2020GL087148. doi: <https://doi.org/10.1029/2020GL087148>
- Gagne II, D. J., Christensen, H. M., Subramanian, A. C., & Monahan, A. H. (2020). Machine learning for stochastic parameterization: Generative adversarial networks in the lorenz '96 model. *Journal of Advances in Modeling Earth Systems*, *12*(3), e2019MS001896. doi: <https://doi.org/10.1029/2019MS001896>
- Haupt, S. E., Chapman, W., Adams, S. V., Kirkwood, C., Hosking, J. S., Robinson, N. H., ... Subramanian, A. C. (2021). Towards implementing artificial intelligence post-processing in weather and climate: proposed actions from the oxford 2019 workshop. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *379*(2194), 20200091. doi: <https://doi.org/10.1098/rsta.2020.0091>
- Jiang, X., Adames, A. F., Kim, D., Maloney, E. D., Lin, H., Kim, H., ... Klingaman, N. P. (2020). Fifty years of research on the madden–julian oscillation: Recent progress, challenges, and perspectives. *Journal of Geophysical Research: Atmospheres*, *125*(17), e2019JD030911. doi: <https://doi.org/10.1029/2019JD030911>
- Kim, H., Ham, Y. G., Joo, Y. S., & Son, S. W. (2021). Deep learning for bias correction of mjo prediction. *Nature Communications*, *12*(3087). doi: <https://doi.org/10.1038/s41467-021-23406-3>
- Kim, H., Vitart, F., & Waliser, D. E. (2018). Prediction of the madden–julian oscillation: A review. *Journal of Climate*, *31*(23), 9425 - 9443. doi: <https://doi.org/10.1175/JCLI-D-18-0210.1>
- Kim, H.-M., Kim, D., Vitart, F., Toma, V. E., Kug, J.-S., & Webster, P. J. (2016). Mjo propagation across the maritime continent in the ecmwf ensemble prediction system. *Journal of Climate*, *29*(11), 3973-3988. doi: <https://doi.org/10.1175/JCLI-D-15-0862.1>
- Kim, H.-M., Webster, P. J., Toma, V. E., & Kim, D. (2014). Predictability and prediction skill of the mjo in two operational forecasting systems. *Journal of Climate*, *27*(14), 5364 - 5378. Retrieved from <https://journals.ametsoc.org/view/journals/clim/27/14/jcli-d-13-00480.1.xml> doi: 10.1175/JCLI-D

- 372 -13-00480.1
- 373 Klotzbach, P. J. (2010). On the madden–julian oscillation–atlantic hurricane rela-
374 tionship. *Journal of Climate*, *23*(2), 282–293. doi: [https://doi.org/10.1175/](https://doi.org/10.1175/2009JCLI2978.1)
375 [2009JCLI2978.1](https://doi.org/10.1175/2009JCLI2978.1)
- 376 Lau, W. K. M., & Waliser, D. E. (2011). Predictability and forecasting. In *In-*
377 *traseasonal variability in the atmosphere-ocean climate system*. Berlin, Heidel-
378 berg: Springer Berlin Heidelberg. doi: [https://doi.org/10.1007/978-3-642-13914](https://doi.org/10.1007/978-3-642-13914-7_12)
379 [-7_12](https://doi.org/10.1007/978-3-642-13914-7_12)
- 380 Lin, H., Brunet, G., & Derome, J. (2008). Forecast skill of the madden–julian oscil-
381 lation in two canadian atmospheric models. *Monthly Weather Review*, *136*(11),
382 4130 - 4149. doi: <https://doi.org/10.1175/2008MWR2459.1>
- 383 Madden, R. A., & Julian, P. R. (1971). Detection of a 40–50 day oscillation in the
384 zonal wind in the tropical pacific. *Journal of Atmospheric Sciences*, *28*(5), 702
385 - 708. doi: [https://doi.org/10.1175/1520-0469\(1971\)028%3C0702:DOADOI%](https://doi.org/10.1175/1520-0469(1971)028%3C0702:DOADOI%3E2.0.CO;2)
386 [3E2.0.CO;2](https://doi.org/10.1175/1520-0469(1971)028%3C0702:DOADOI%3E2.0.CO;2)
- 387 Madden, R. A., & Julian, P. R. (1972). Description of global-scale circulation cells
388 in the tropics with a 40–50 day period. *Journal of Atmospheric Sciences*,
389 *29*(6), 1109 - 1123. doi: [https://doi.org/10.1175/1520-0469\(1972\)029%3C1109:](https://doi.org/10.1175/1520-0469(1972)029%3C1109:DOGSCC%3E2.0.CO;2)
390 [DOGSCC%3E2.0.CO;2](https://doi.org/10.1175/1520-0469(1972)029%3C1109:DOGSCC%3E2.0.CO;2)
- 391 Martin, Z. K., Barnes, E. A., & Maloney, E. D. (2021). Using simple, explainable
392 neural networks to predict the madden–julian oscillation. *Earth and Space Sci-*
393 *ence Open Archive*. doi: <https://doi.org/10.1002/essoar.10507439.1>
- 394 Martin, Z. K., Son, S.-W., Butler, A., Hendon, H., Kim, H., Sobel, A., . . . Zhang,
395 C. (2021). The influence of the quasi-biennial oscillation on the madden–
396 julian oscillation. *Nature Reviews Earth and Environment*(2), 477–489. doi:
397 <https://doi.org/10.1038/s43017-021-00173-9>
- 398 McGovern, A., Lagerquist, R., II, D. J. G., Jergensen, G. E., Elmore, K. L., Home-
399 yer, C. R., & Smith, T. (2019). Making the black box more transpar-
400 ent: Understanding the physical implications of machine learning. *Bul-*
401 *letin of the American Meteorological Society*, *100*(11), 2175–2199. doi:
402 <https://doi.org/10.1175/BAMS-D-18-0195.1>
- 403 O’Gorman, P. A., & Dwyer, J. G. (2018). Using machine learning to parameter-
404 ize moist convection: Potential for modeling of climate, climate change, and
405 extreme events. *Journal of Advances in Modeling Earth Systems*, *10*(10),
406 2548–2563. doi: <https://doi.org/10.1029/2018MS001351>
- 407 Rashid, H. A., Hendon, H. H., Wheeler, M. C., & Alves, O. (2011). Prediction of the
408 madden–julian oscillation with the poama dynamical prediction system. *Cli-*
409 *mate Dynamics*, *36*, 649–661. doi: <https://doi.org/10.1007/s00382-010-0754-x>
- 410 Rasp, S., & Lerch, S. (2018). Neural networks for postprocessing ensemble weather
411 forecasts. *Monthly Weather Review*, *146*(11), 3885–3900. doi: [https://doi.org/](https://doi.org/10.1175/MWR-D-18-0187.1)
412 [10.1175/MWR-D-18-0187.1](https://doi.org/10.1175/MWR-D-18-0187.1)
- 413 Scheuerer, M., Switanek, M. B., Worsnop, R. P., & Hamill, T. M. (2020). Using
414 artificial neural networks for generating probabilistic subseasonal precipitation
415 forecasts over california. *Monthly Weather Review*, *148*(8), 3489–3506. doi:
416 <https://doi.org/10.1175/MWR-D-20-0096.1>
- 417 Silini, R. (2021a). *Feed-forward neural networks*. Retrieved from `\url{https://`
418 `doi.org/10.5281/zenodo.5801453}` doi: [https://doi.org/10.5281/](https://doi.org/10.5281/zenodo.5801453)
419 [zenodo.5801453](https://doi.org/10.5281/zenodo.5801453)
- 420 Silini, R. (2021b). *Wheeler-hendon phase diagrams*. Retrieved from `\url{https://`
421 `doi.org/10.5281/zenodo.5801415}` doi: [https://doi.org/10.5281/](https://doi.org/10.5281/zenodo.5801415)
422 [zenodo.5801415](https://doi.org/10.5281/zenodo.5801415)
- 423 Silini, R., Barreiro, M., & Masoller, C. (2021). Machine learning prediction of the
424 madden–julian oscillation. *npj Climate and Atmospheric Science*, *4*(57). doi:
425 <https://doi.org/10.1038/s41612-021-00214-6>

- 426 Taraphdar, S., Zhang, F., Leung, L. R., Chen, X., & Pauluis, O. M. (2018). Mjo
 427 affects the monsoon onset timing over the indian region. *Geophysical Research*
 428 *Letters*, *45*(18). doi: <https://doi.org/10.1029/2018GL078804>
- 429 Tseng, K.-C., Barnes, E. A., & Maloney, E. (2020). The importance of past mjo ac-
 430 tivity in determining the future state of the midlatitude circulation. *Journal of*
 431 *Climate*, *33*(6), 2131 - 2147. doi: <https://doi.org/10.1175/JCLI-D-19-0512.1>
- 432 Ungerovich, M., Barreiro, M., & Masoller, C. (2021). Influence of madden–julian os-
 433 cillation on extreme rainfall events in spring in southern uruguay. *International*
 434 *Journal of Climatology*, 1-13. doi: <https://doi.org/10.1002/joc.7022>
- 435 Vannitsem, S., Bremnes, J. B., Demaeyer, J., Evans, G. R., Flowerdew, J., Hemri,
 436 S., ... Ylhaisi, J. (2021). Statistical postprocessing for weather forecasts:
 437 Review, challenges, and avenues in a big data world. *Bulletin of the Ameri-*
 438 *can Meteorological Society*, *102*(3), E681-E699. doi: [https://doi.org/10.1175/](https://doi.org/10.1175/BAMS-D-19-0308.1)
 439 [BAMS-D-19-0308.1](https://doi.org/10.1175/BAMS-D-19-0308.1)
- 440 Vitart, F. (2009). Impact of the madden julian oscillation on tropical storms and
 441 risk of landfall in the ecmwf forecast system. *Geophysical Research Letters*, *36*.
 442 doi: <https://doi.org/10.1029/2009GL039089>
- 443 Wang, S., Tippett, M. K., Sobel, A. H., Martin, Z. K., & Vitart, F. (2019). Im-
 444 pact of the qbo on prediction and predictability of the mjo convection. *Journal*
 445 *of Geophysical Research: Atmospheres*, *124*(22), 11766-11782. doi: [https://doi](https://doi.org/10.1029/2019JD030575)
 446 [.org/10.1029/2019JD030575](https://doi.org/10.1029/2019JD030575)
- 447 Wheeler, M. C., & Hendon, H. H. (2004). An all-season real-time multivari-
 448 ate mjo index: Development of an index for monitoring and prediction.
 449 *Monthly Weather Review*, *132*(8), 1917 - 1932. doi: [https://doi.org/10.1175/](https://doi.org/10.1175/1520-0493(2004)132(1917:AARMMI)2.0.CO;2)
 450 [1520-0493\(2004\)132\(1917:AARMMI\)2.0.CO;2](https://doi.org/10.1175/1520-0493(2004)132(1917:AARMMI)2.0.CO;2)
- 451 Wheeler, M. C., Hendon, H. H., Cleland, S., Meinke, H., & Donald, A. (2009).
 452 Impacts of the madden-julian oscillation on australian rainfall and circula-
 453 tion. *Journal of Climate*, *22*(6), 1482-1498. doi: [https://doi.org/10.1175/](https://doi.org/10.1175/2008JCLI2595.1)
 454 [2008JCLI2595.1](https://doi.org/10.1175/2008JCLI2595.1)
- 455 Wu, C.-H., & Hsu, H.-H. (2009). Topographic influence on the mjo in hte maritime
 456 continent. *Journal of Climate*, *22*(20), 5433-5448. doi: [https://doi.org/10](https://doi.org/10.1175/2009JCLI2825.1)
 457 [.1175/2009JCLI2825.1](https://doi.org/10.1175/2009JCLI2825.1)
- 458 Wu, J., & Jin, F.-F. (2021). Improving the mjo forecast of s2s operation models by
 459 correcting their biases in linear dynamics. *Geophysical Research Letters*, *48*(6),
 460 e2020GL091930. doi: <https://doi.org/10.1029/2020GL091930>
- 461 Zhang, C., Gottschalck, J., Maloney, E. D., Moncrieff, M. W., Vitart, F., Waliser,
 462 D. E., ... Wheeler, M. C. (2013). Cracking the mjo nut. *Geophysical Research*
 463 *Letters*, *40*(6), 1223-1230. doi: <https://doi.org/10.1002/grl.50244>