

## ARTICLE OPEN



## Machine learning prediction of the Madden-Julian oscillation

Riccardo Silini<sup>1</sup>✉, Marcelo Barreiro<sup>2</sup> and Cristina Masoller<sup>1</sup>

The socioeconomic impact of weather extremes draws the attention of researchers to the development of novel methodologies to make more accurate weather predictions. The Madden–Julian oscillation (MJO) is the dominant mode of variability in the tropical atmosphere on sub-seasonal time scales, and can promote or enhance extreme events in both, the tropics and the extratropics. Forecasting extreme events on the sub-seasonal time scale (from 10 days to about 3 months) is very challenging due to a poor understanding of the phenomena that can increase predictability on this time scale. Here we show that two artificial neural networks (ANNs), a feed-forward neural network and a recurrent neural network, allow a very competitive MJO prediction. While our average prediction skill is about 26–27 days (which competes with that obtained with most computationally demanding state-of-the-art climate models), for some initial phases and seasons the ANNs have a prediction skill of 60 days or longer. Furthermore, we show that the ANNs have a good ability to predict the MJO phase, but the amplitude is underestimated.

*npj Climate and Atmospheric Science* (2021)4:57; <https://doi.org/10.1038/s41612-021-00214-6>

## INTRODUCTION

The Madden–Julian oscillation (MJO)<sup>1,2</sup> is a major source of weather predictability on the sub-seasonal time scale<sup>3–5</sup> and has an important influence on the tropical weather<sup>6</sup>. The MJO is a major source of intraseasonal fluctuations in monsoon systems<sup>7,8</sup> and modulates the development of tropical cyclones<sup>9</sup>. In addition, the MJO influences the extratropical regions through atmospheric teleconnections (e.g., refs. <sup>10,11</sup>) and its activity may affect El Niño–Southern Oscillation (ENSO)<sup>12</sup>. For these reasons, many efforts have focused on forecasting the MJO<sup>3,13–20</sup>.

Significant advances in the understanding of the physics involved in the MJO and better dynamical forecasting systems have allowed improving the skill of MJO prediction. For the dynamical models, the prediction skill of MJO is sensitive to the physics of the model and the quality of the initial conditions. Of the dynamical models considered in 2014 by Neena and coworkers<sup>13</sup>, the ensemble-mean prediction skill is highest for the model of the European Centre for Medium-Range Weather Forecast (ECMWF, 28 days) and for the model of the Australian Bureau of Meteorology (ABOM2, 24 days), and it is in the range of 15–20 days for most other models. More recently, the prediction skill of ECMWF has improved to exceed 4 weeks, while most models have improved their skill to the range of 20–25 days<sup>18</sup>. The MJO prediction skill has also been shown to depend on the initial amplitude and phase, the season of the year, the background mean state, and the extratropical influence<sup>18</sup>. Boreal winter leads, for most models, to a higher prediction skill that reaches up to 25–26 days, except for the ECMWF model that approaches 5 weeks<sup>20</sup>.

Machine learning (ML) algorithms are nowadays widely used in science and technology. In climate science, a major problem where ML techniques can be useful is the representation of ocean mixing processes and atmospheric convection, which are poorly resolved in weather prediction models and global climate models<sup>21,22</sup>. ML techniques have also been used to forecast important climate phenomena, such as ENSO<sup>22</sup>, and to reconstruct the historical MJO index<sup>23</sup> among others; however, to the best of

our knowledge, ML algorithms have not yet been used to predict MJO, except for correcting the bias of dynamical models<sup>24</sup>.

To fill this gap, here we use ML techniques to predict the real-time multivariate MJO (RMM) index<sup>25</sup>, which is an index commonly used to describe the evolution of MJO. We consider the period between January 1, 1979 and December 31, 2020. We train two artificial neural networks (ANNs), a feed-forward neural network (FFNN) and an autoregressive recurrent neural network (AR-RNN). We show that these ANNs provide a mean prediction skill of about 26–27 days. We also show that they lead to a very good prediction of the MJO phase, but to an underestimation of the MJO amplitude. We also analyze the influence of the initial phase in the prediction skill and the seasonal dependence of the prediction skill, and we compare our results with those reported in the literature<sup>13,17–20</sup>.

This paper is organized as follows. In the next section, we present the results of the analysis of the RMM index using the two ANNs. To quantify the prediction skill we use two measures that are widely used in the literature, the bivariate correlation coefficient (COR) and the root-mean-squared error (RMSE). We present then the discussion of the results and our conclusions.

## RESULTS

## Prediction skill

We begin by computing COR and RMSE as a function of the forecast lead time,  $\tau$ , for the two ANNs (see Methods). Averaging over all seasons we obtain the results shown in Fig. 1, where we display COR and RMSE as a function of  $\tau = 5, 10, \dots, 60$  days, for an initial RMM amplitude larger than 1. In this figure, we see that both ANNs perform very similarly. The AR-RNN seems to perform slightly better than FFNN up to 10 days prediction, after which, the two curves overlap up to 50 days when the latter starts providing a better prediction. Using the standard value  $COR = 0.5$  to define the prediction skill, we find a prediction skill of about 26–27 days for both ANNs, which is comparable to the best-known prediction skills obtained from most models<sup>18</sup>, except ECMWF. Regarding the RMSE, using the standard value  $RMSE = 1.4$  to define the

<sup>1</sup>Departament de Física, Universitat Politècnica de Catalunya, Edifici Gaia, Rambla Sant Nebridi 22, 08222 Terrassa, Barcelona, Spain. <sup>2</sup>Departamento de Ciencias de la Atmósfera, Facultad de Ciencias, Universidad de la República, Iguá 4225, 11400 Montevideo, Uruguay. ✉email: [riccardo.silini@outlook.com](mailto:riccardo.silini@outlook.com)

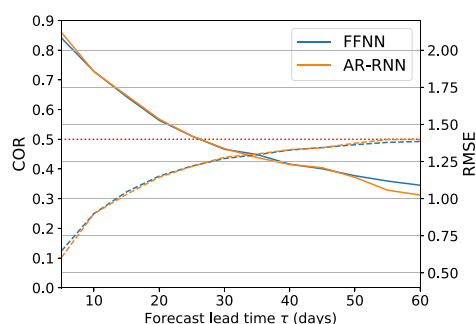
prediction skill, we see that the prediction skill is longer than 60 days, as, for both ANNs, RMSE never crosses this value for  $\tau$  values up to 60 days. A video<sup>26</sup> showing the real and the predicted MJO evolution in the Wheeler–Hendon phase diagram clearly visualizes the very good prediction ability.

We then compute the error of the predictions for the MJO amplitude and phase (see Methods). The results are presented in Fig. 2, where we notice that, for both ANNs, the phase is well predicted but the amplitude is underestimated, and its absolute error grows as the lead time increases.

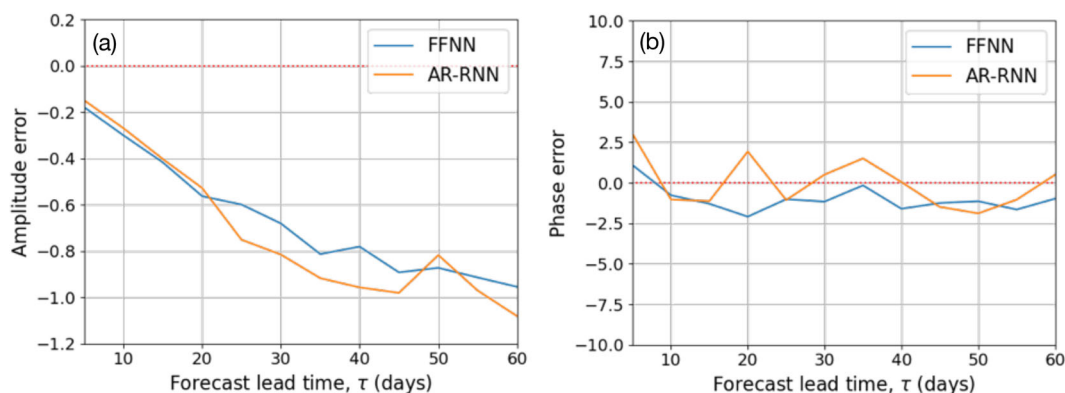
### Seasonally resolved prediction skill

We now perform the same analysis for the dataset restricted to each season using the FFNN, which is the fastest and simplest of the two ANNs. The results are presented in Figs. 3 and 4.

In Fig. 3, we see a large difference in the prediction skill in different seasons. Boreal spring (March–May, MAM) and fall (September–November, SON), the transition seasons, are the least predictable with COR prediction skills of 23–24 days and 16–17 days, respectively. In boreal summer (June–August, JJA) the prediction skill is around 31 days, while in boreal winter December–February (DJF) it is around 45 days. We also note that DJF has the largest RMSE, which means that the prediction correlates well with the observations, but the predicted and actual values are quite different. On the contrary, JJA has a very low RMSE, which means that even if JJA has a lower COR than DJF, the



**Fig. 1 COR and RMSE averaged over all seasons.** Bivariate correlation coefficient (COR) (solid) and root-mean-squared error RMSE (dashed) averaged over all seasons in the test set, as a function of the forecast lead time  $\tau$ . The color indicates the artificial neural network (FFNN feed-forward neural network, AR-RNN autoregressive recurrent neural network) and the dotted line indicates the threshold that defines the prediction skill. While the RMSE threshold (1.4) is never crossed, the COR value falls below 0.5 around 26–27 days.



**Fig. 2 MJO amplitude- and phase errors averaged over all seasons.** MJO amplitude (a) and phase (b) error averaged over all seasons in the test set, as a function of the lead time.

prediction is more accurate. The transition seasons are in the middle, with SON showing larger RMSE than MAM, as found for COR. The highest COR and RMSE are for DJF, which is likely due to the fact that MJO is most active during the extended boreal winter (DJFM), which would also partially explain the large (yet smaller than DJF), RMSE of MAM.

Figure 4 displays the amplitude and phase errors as a function of the lead time (as in Fig. 2, but here for the individual seasons). We notice that boreal winter (DJF) has the largest amplitude error, while boreal summer (JJA) has the lowest one. Regarding the phase error, we note that in JJA the predicted MJO propagation is faster than the real one, while in the other three seasons, the predicted propagation is slower.

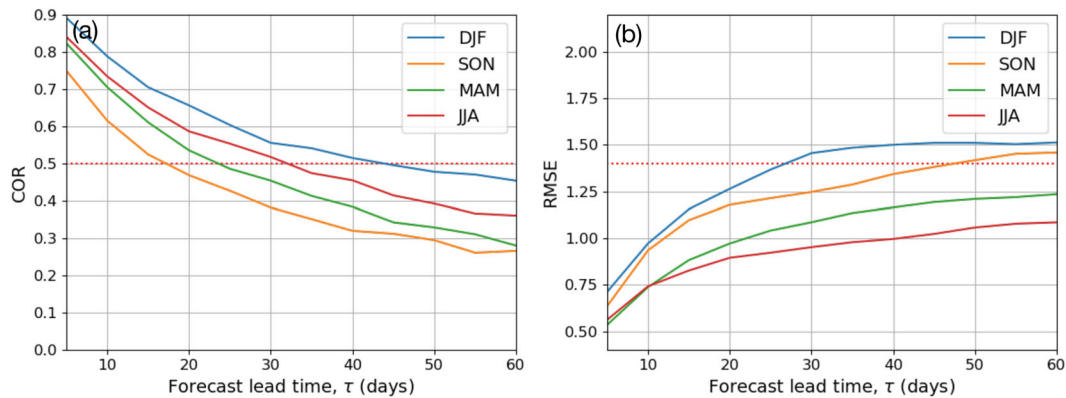
Finally, we study the dependence of the COR and RMSE prediction skill as a function of the MJO initial phase and the season. The results are presented in Figs. 5 (COR) and 6 (RMSE). In boreal winter (DJF in blue), we can notice that starting from phases 1, 2, 5, and 8 the prediction skill using COR is very high, in fact, it has a skill for up to 60 days or longer, while it falls below 20 days for phase 7. Nevertheless, Fig. 6 shows that for phases 5 and 8 the threshold is crossed below 30 days. By combining the information presented in the two figures, we can infer a prediction skill of about 60 days for phases 1 and 2.

For boreal fall (SON, orange) we also see a strong dependence of the skill on the initial phase: it is around 50 days for phases 4 and 7, while all other initial phases lead to prediction skills lower than 20 days. The skill in boreal spring (MAM, green) and summer (JJA, red) is more uniform across different initial phases, but the highest prediction skill achieved (given by COR) is around 40 days, and the lowest (below 20 days) are in phases 1, 3, 8 and 1, 5, 8, respectively. Overall, we can notice that the initial phase 1 provides a very high prediction skill in boreal winter, while it is low in all other seasons. Starting from phase 2, the prediction skill is larger than 35 days from December to May, while for initial phase 3 the highest prediction skill (around 40 days) is found in winter and summer. The initial phase 4 provides high skill (more than 40 days) in the transition seasons. Starting from initial phase 6, provides high skill from March to August, while starting from phase 7 gives a prediction skill above 40 days from June to November. Lastly, starting from phase 8 the prediction skill is always below 20 days.

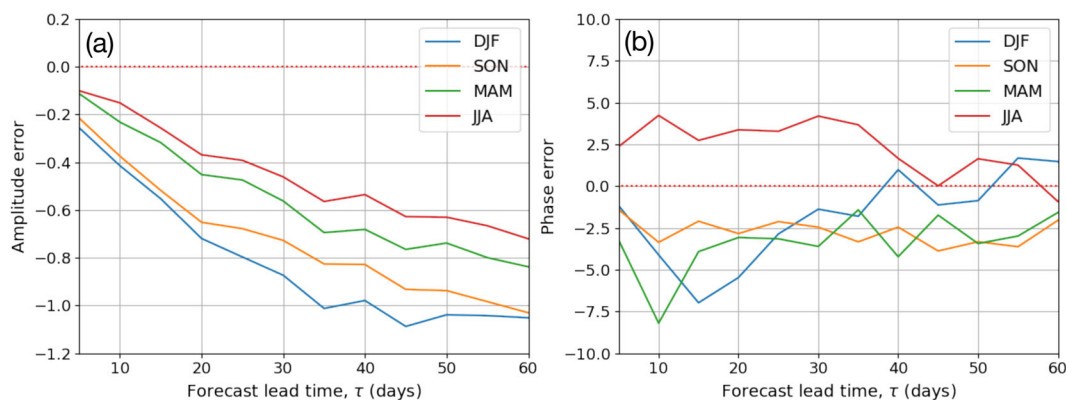
In Fig. 6 we also notice that the RMSE for MAM and JJA never crosses the 1.4 threshold, for up to 100 days.

## DISCUSSION

We have used two types of ANNs to predict the MJO. We have used a feed-forward neural network (FFNN) and an autoregressive recurrent neural network (AR-RNN) to predict the daily Real-time



**Fig. 3** Seasonally resolved COR and RMSE. COR (a) and RMSE (b) as a function of the leading time in days, obtained with the feed-forward neural network (FFNN). The different colors represent different seasons.



**Fig. 4** Seasonally resolved MJO amplitude- and phase errors. Amplitude (a) and phase (b) errors for the different seasons (represented with different colors), obtained with the FFNN.

Multivariate MJO indices, RMM1 and RMM2, analyzing the period between January 1, 1979 and December 31, 2020. First, we considered the whole dataset, and in the second step, we considered individual seasons (boreal winter, DJF, spring, MAM, summer, JJA, and fall, SON). We have quantified the prediction skill as a function of the leading time,  $\tau$ , using standard magnitudes and thresholds (COR and RMSE with thresholds 0.5 and 1.4, respectively<sup>27</sup>).

For the full dataset, using COR we have found a prediction skill of 26–27 days, which is comparable to most dynamical models. Using the RMSE, the prediction skill we have obtained is up to 60 days.

We have obtained a very good prediction of the RMM phase, but a poorer prediction of the RMM amplitude, which was systematically underestimated. Comparing these results with those reported in ref. <sup>28</sup>, we notice that the two ANNs used here lead to a worse prediction of the amplitude, but to a better prediction of the phase, in comparison with the predictions obtained from most dynamical models. The larger amplitude error is due to the systematic underestimation, as the error adds up. In contrast, dynamical models sometimes overestimate and sometimes underestimate, which leads to a lower amplitude error, due to partial compensation of positive and negative errors.

Consistent with previous studies<sup>27,29–32</sup> we have found significant differences among seasons.

We found that boreal fall and spring have the lowest prediction skill, being 16–17 and 23–24 days, respectively. In accordance with refs. <sup>27,29–31</sup>, we found the highest prediction skill in boreal winter, which in our case is of around 45 days. Another study<sup>32</sup> found the

highest prediction skill in boreal fall. In boreal summer we have found a prediction skill of about 31 days.

We have also studied the dependence of the prediction skill as a function of the initial MJO phase. We have found large variability in prediction skills in boreal winter and fall. In the best conditions, in boreal winter with an initial MJO phase of 1 and 2, the ANN has a prediction skill for up to 60 days or more. Our results indicate that the most difficult conditions to predict MJO is in boreal fall when the initial MJO phase is phase 1.

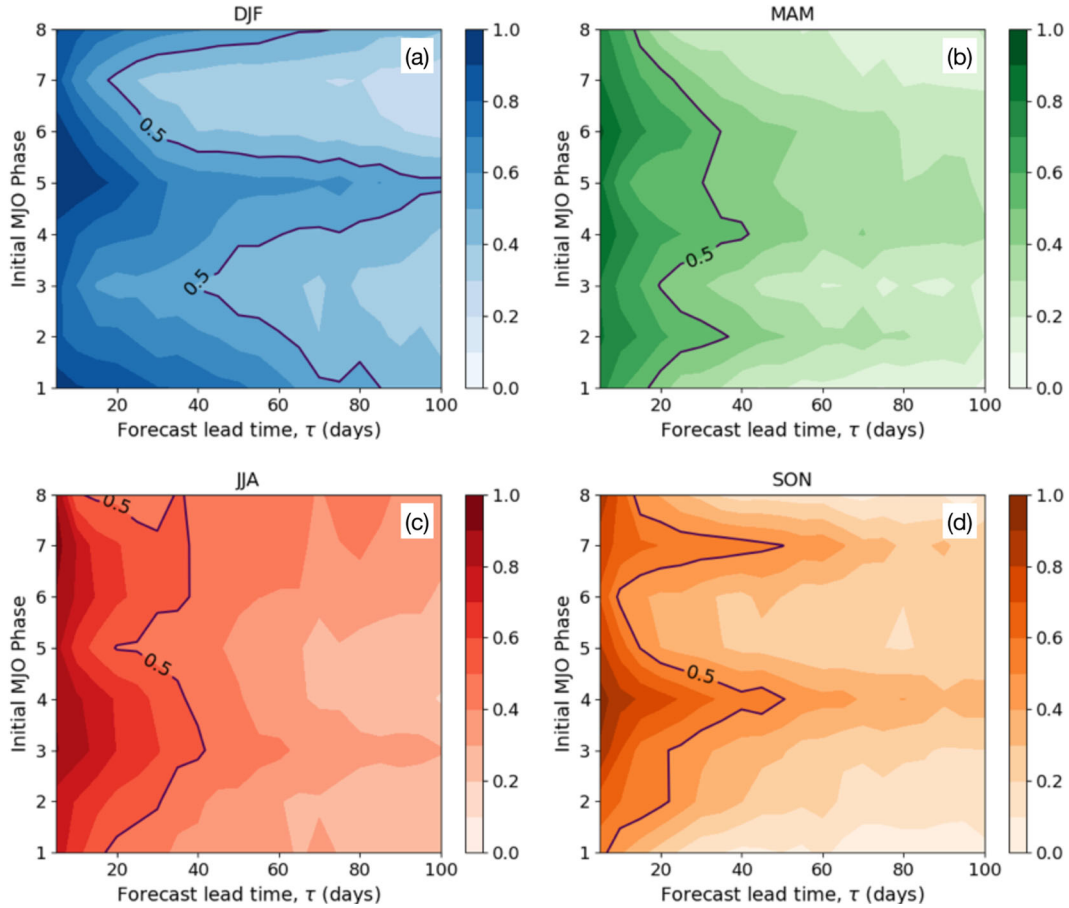
A major advantage of the ANNs considered is that they are computationally low-cost, and they do not have the limitations of dynamical models, where the MJO prediction skill depends strongly on the model's physics, initialization, and ocean–atmosphere coupling processes. On the other hand, the very own nature of ANNs preclude the understanding of the physical processes involved and thus they represent a complementary approach that, according to our results, is worth pursuing.

For future work, the MJO prediction skill could potentially be improved by training the ANNs independently for each season (for simplicity, here we have trained them on all seasons and tested them on individual seasons). A study of the predictability barrier of the RMM index from different seasons and phases could also shed light on the results obtained with machine learning methods<sup>33</sup>.

## METHODS

### Dataset

We use the daily Real-time Multivariate MJO indices, RMM1 and RMM2<sup>25</sup>, which are the first and second principal components of the combined empirical orthogonal functions (EOFs) of outgoing longwave radiation



**Fig. 5** Seasonally resolved COR as a function of the initial MJO phase. COR as a function of the initial MJO phase and forecast lead time  $\tau$ . Each plot corresponds to a different season: boreal winter (**a**; blue), spring (**b**; green), summer (**c**; red), and fall (**d**; orange).

(OLR), zonal wind at 200 and 850 hPa averaged between 15°N and 15°S. Using these two variables in a phase diagram it is possible to define the MJO phase and amplitude. The phase is classified in one of eight sectors of the phase diagram defining the observed MJO life cycle, while an amplitude smaller than 1 corresponds to a non-active MJO. RMM1 and RMM2, as well as the phase and amplitude since June 1, 1974 were downloaded from ref. <sup>34</sup>. The same tools used in this study could also be applied to other MJO indices, such as the OLR MJO index (OMI), the original OLR MJO index (OOMI), the real-time OLR MJO index (ROMI), and the filtered OLR MJO index (FMO), which can be downloaded from ref. <sup>35</sup>.

Due to missing data in the first years we limit the study to the period between January 1, 1979 and December 31, 2020, which is L2-normalized.

### Prediction skill quantifiers and errors

We use the same quantifiers of the prediction skill as in ref. <sup>18</sup>, which are adapted from refs. <sup>27,36</sup>. The bivariate correlation coefficient (COR) and the root-mean-squared error (RMSE) are defined as:

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

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

where  $a_1(t)$  and  $a_2(t)$  are the observed RMM1 and RMM2 at time  $t$ , and  $b_1(t, \tau)$  and  $b_2(t, \tau)$  are the respective forecasts for time  $t$  with a lead time of  $\tau$  days, and  $N$  is the number of predictions. COR expresses the strength of co-occurrence between the forecast and the observations, while RMSE does a term-by-term comparison of the actual difference between the forecast and the observations. The values  $\text{COR} = 0.5$  and  $\text{RMSE} = 1.4$  are usually used as skill thresholds<sup>27</sup>: the prediction skill refers to the time when the COR falls below 0.5 and RMSE grows above 1.4.

Through a change of coordinates from Cartesian to polar, we calculate the amplitude and phase, (RMM1, RMM2)  $\rightarrow$  ( $A, \varphi$ )<sup>27</sup>, and define their errors as

$$E_A(\tau) = \frac{1}{N} \sum_{t=1}^N [A_{\text{pred}}(t, \tau) - A_{\text{obs}}(t)], \quad (3)$$

$$E_\varphi(t, \tau) = \frac{1}{N} \sum_{t=1}^N \tan^{-1} \left( \frac{a_1(t)b_2(t, \tau) - a_2(t)b_1(t, \tau)}{a_1(t)b_1(t, \tau)} \right), \quad (4)$$

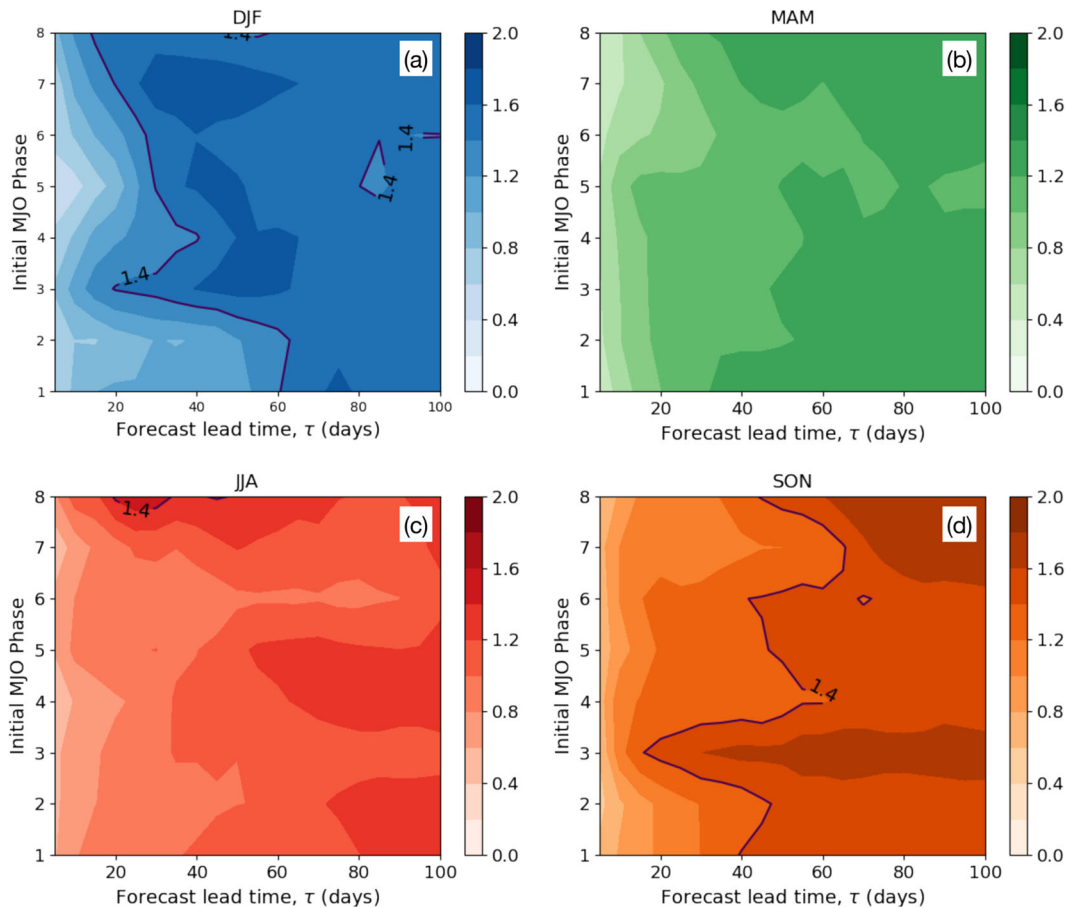
where  $A_{\text{obs}}(t)$  is the observed amplitude at time  $t$  and  $A_{\text{pred}}(t, \tau)$  is the predicted amplitude at time  $t$  with a lead time of  $\tau$  days.

### Artificial neural networks (ANNs)

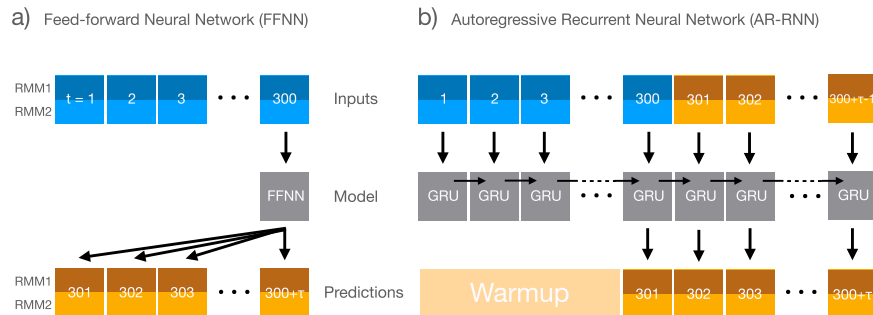
In this study, we use two well-known ANNs, schematically shown in Fig. 7: a feed-forward neural network (FFNN) and an autoregressive recurrent neural network (AR-RNN), both having an input layer of 300 units.

The FFNN uses the last point of the input layer and links it to one hidden layer composed of 64 units, itself linked to an output layer of  $\tau$  units fully connected, where  $\tau = 5, 10, \dots, 100$  is the forecast lead time. Each input and output is composed of two values, corresponding to RMM1 and RMM2, as shown in Fig. 7a.

The AR-RNN is a single gated recurrent unit (GRU)<sup>37</sup> layer composed of 64 units, displayed in Fig. 7b. Instead of predicting the entire output sequence in a single step, with this recurrent neural network, we decompose the prediction into individual time steps that are fed back into the network after a warm-up, which updates the internal state of the network and discards the outputs considering them poor predictions. GRU is chosen over a classical RNN to prevent the vanishing gradient problem, which corresponds to the potential tendency of the loss function gradients to approach zero, making the backpropagation of the error to



**Fig. 6** Seasonally resolved RMSE as a function of the initial MJO phase. RMSE as a function of the initial MJO phase and forecast lead time  $\tau$ . Each plot corresponds to a different season: boreal winter (a; blue), spring (b; green), summer (c; red), and fall (d; orange).



**Fig. 7** ANN architecture. Diagram of the ANNs employed in this study. **a** represents the FFNN, while **b** the AR-RNN.

not affect the first layer neurons of a multi-layer network. It is also preferred over a long short-term memory ANN due to the lower computational time required. Since we don't have several hidden layers, the vanishing gradient problem is not an issue, and in this way, we leave open the possibility of increasing the number of layers for achieving a better prediction skill.

For the FFNN the activation function is a rectified linear activation function (ReLU), which is responsible for transforming the summed weighted input from the node into the activation of the node or output for that input. Sigmoid functions generally work better in the case of classifiers, and just like *tanh* functions might be avoided due to the vanishing gradient problem. If by increasing the number of hidden neurons one might encounter multiple dead neurons, i.e., non-active neurons, we suggest using the leaky version of ReLU, or its parameterized version.

The mean squared error (MSE) is used as a loss function, which is the default loss used for regression problems and the RMM values are not widely spread and do not have outliers, which motivates this choice instead of using mean squared logarithmic error (MSLE) or mean absolute error (MAE).

Finally, the Adam optimizer is used for training, with a maximum of ten epochs. We selected patience of 1, used for the early stopping of the training to avoid overtraining, which corresponds to the delay in stopping. Adam optimizer is chosen being the best common method among adaptive optimizers, which doesn't require a tuning of the learning rate value. The maximum number of epochs is never reached as the learning is stopped if the validation error starts growing. We could increase the patience to account for possible local minima of the validation error, but that would require more computational time, and

we preferred to use fast and simple ANNs for a demonstration of their ability for MJO prediction.

To perform the backtesting or hindcast, we selected a train-validation-test splitting that preserves the temporal order of observations. Other methods like multiple train-test splits or the walk-forward validation could be applied and would result in a more robust estimation of the model performance on out-of-sample data. The drawback of such methods is the cost of creating multiple models, which would sensibly slow down the training.

The dataset is divided into three sets: the train set contains data from 1 January 1979 to 30 November 2006, the validation set, from 1 December 2006 to 30 November 2015, and the test set, from 1 December 2015 to 31 December 2020.

The ANNs are trained on the train set, and the model's internal parameters are updated every 16 (batch size) exposure of different training samples. After the training, the ANN is evaluated using the validation set to fine-tune the hyperparameters. This training and validation process is repeated a maximum of ten times. Then, a single evaluation is performed using the test set, which was not previously seen by the ANNs.

## DATA AVAILABILITY

The RMM data is freely available in ref. <sup>34</sup>.

## CODE AVAILABILITY

The Keras TensorFlow<sup>38</sup> trained FFNN can be found in ref. <sup>39</sup>.

Received: 4 May 2021; Accepted: 29 October 2021;

Published online: 25 November 2021

## REFERENCES

- Madden, R. A. & Julian, P. R. Detection of a 40–50 day oscillation in the zonal wind in the tropical pacific. *J. Atmos. Sci.* **28**, 702–708 (1971).
- Madden, R. A. & Julian, P. R. Description of global-scale circulation cells in the tropics with a 40–50 day period. *J. Atmos. Sci.* **29**, 1109–1123 (1972).
- Waliser, D. E. in *Intraseasonal Variability in the Atmosphere–Ocean Climate System* (eds Lau, W. K. M. & Waliser, D. E.) Ch. 12 (Springer, 2011).
- Zhang, C. et al. Cracking the MJO nut. *Geophys. Res. Lett.* **40**, 1223–1230 (2013).
- Ferranti, L., Magnusson, L., Vitart, F. & Richardson, D. S. How far in advance can we predict changes in large-scale flow leading to severe cold conditions over Europe? *Q. J. R. Meteorol. Soc.* **144**, 1788–1802 (2018).
- Vitart, F. Impact of the Madden Julian oscillation on tropical storms and risk of landfall in the ECMFW forecast system. *Geophys. Res. Lett.* **36**, L15802 (2009).
- Taraphdar, S., Zhang, F., Leung, L. R., Chen, X. & Pauluis, O. M. Mjo affects the monsoon onset timing over the indian region. *Geophys. Res. Lett.* **45**, 10011–10018 (2018).
- Díaz, N., Barreiro, M. & Rubido, N. Intraseasonal predictions for the south american rainfall dipole. *Geophys. Res. Lett.* **47**, e2020GL089985 (2020).
- Camargo, S. J., Wheeler, M. C. & Sobel, A. H. Diagnosis of the MJO modulation of tropical cyclogenesis using an empirical index. *J. Atmos. Sci.* **66**, 3061–3074 (2009).
- Alvarez, M. S., Vera, C. S. & Kiladis, G. N. Mjo modulating the activity of the leading mode of intraseasonal variability in south America. *Atmosphere* **8**, 232 (2017).
- Ungerovich, M., Barreiro, M. & Masoller, C. Influence of Madden-Julian oscillation on extreme rainfall events in spring in southern Uruguay. *Int. J. Climatol.* **41**, 1–13 (2021).
- Bergman, J. W., Hendon, H. H. & Weickmann, K. M. Intraseasonal air–sea interactions at the onset of el niño. *J. Clim.* **14**, 1702–1719 (2001).
- Neena, J. M., Lee, J. Y., Waliser, D., Wang, B. & Jiang, X. Predictability of the Madden–Julian oscillation in the intraseasonal variability hindcast experiment (ISVHE). *J. Clim.* **27**, 4531–4543 (2014).
- Kim, D. & Maloney, E. D. In *The Global Monsoon System* (eds Chang, C. P. et al.) Ch. 9 (World Scientific, 2017).
- Waliser, D. E. In *The Asian Monsoon* (ed. Wang, B.) Ch. 5 (Springer, 2006).
- Waliser, D. E. In *Predictability of Weather and Climate* (eds Palmer, T. & Hagedorn, R.) Ch. 11 (Cambridge Univ. Press, 2006).
- Lee, J. Y., Fu, X. & Wang, B. In *The Global Monsoon System* (eds Chang, C.-P. et al.) Ch. 12 (World Scientific, 2017).
- Kim, H., Vitart, F. & Waliser, D. E. Prediction of the Madden–Julian oscillation: a review. *J. Clim.* **31**, 9425–9443 (2018).
- Lim, Y., Son, S.-W. & Kim, D. Mjo prediction skill of the subseasonal-to-seasonal prediction models. *J. Clim.* **31**, 4075–4094 (2018).
- Jiang, X. et al. Fifty years of research on the Madden–Julian oscillation: recent progress, challenges, and perspectives. *J. Geophys. Res.* **125**, e2019JD030911 (2020).
- O’Gorman, P. A. & Dwyer, J. G. Using machine learning to parameterize moist convection: potential for modeling of climate, climate change, and extreme events. *J. Adv. Mod. Earth Syst.* **10**, 2548–2563 (2018).
- Dijkstra, H. A., Petersik, P., Hernández-García, E. & López, C. The application of machine learning techniques to improve el niño prediction skill. *Front. Phys.* **7**, 1–13 (2019).
- Tseng, K.-C., Barnes, E. A. & Maloney, E. The importance of past MJO activity in determining the future state of the midlatitude circulation. *J. Clim.* **33**, 2131–2147 (2020).
- Kim, H., Ham, Y. G., Joo, Y. S. & Son, S. W. Deep learning for bias correction of MJO prediction. *Nat. Commun.* **12**, 3087 (2021).
- Wheeler, M. C. & Hendon, H. H. An all-season real-time multivariate MJO index: development of an index for monitoring and prediction. *Mon. Weather Rev.* **132**, 1917–1932 (2004).
- Silini, R. Wheeler–Hendon phase diagram MJO forecast DJF 2015–2020, FFNN – 20 lead days (1.0). *Zenodo* <https://doi.org/10.5281/zenodo.4733942> (2021).
- Rashid, H. A., Hendon, H. H., Wheeler, M. C. & Alves, O. Prediction of the Madden–Julian oscillation with the POAMA dynamical prediction system. *Clim. Dyn.* **36**, 649–661 (2011).
- Vitart, F. Madden–Julian oscillation prediction and teleconnections in the s2s database. *Q. J. R. Meteorol. Soc.* **143**, 2210–2220 (2017).
- Wheeler, M. & Weickmann, K. M. Real-time monitoring and prediction of modes of coherent synoptic to intraseasonal tropical variability. *Mon. Weather Rev.* **129**, 2677–2694 (2001).
- Lin, J.-L. et al. Tropical intraseasonal variability in 14 IPCC AR4 climate models. Part I: convective signals. *J. Clim.* **19**, 2665–2690 (2006).
- Seo, K.-H. Statistical–dynamical prediction of the Madden–Julian oscillation using NCEP climate forecast system (CFS). *Int. J. Climatol.* **29**, 2146–2155 (2009).
- Wu, J., Ren, H.-L., Zuo, J., Zhao, C., Chen, L. & Li, Q. Mjo prediction skill, predictability, and teleconnection impacts in the Beijing climate center atmospheric general circulation model. *Dyn. Atmos. Oceans* **75**, 78–90 (2016).
- Liu, Z., Jin, Y. & Rong, X. A theory for the seasonal predictability barrier: threshold, timing, and intensity. *J. Clim.* **32**, 423–443 (2019).
- RMM data. *LDEO/IRI Data Library* <https://iridl.ldeo.columbia.edu/SOURCES/BoM/MJO/RMM/index.html?Set-Language=en> (2021).
- MJO indices data. *NOAA/ESRL Physical Sciences Laboratory* <https://psl.noaa.gov/mjo/mjoindex/> (2021).
- Lin, H., Brunet, G. & Derome, J. Forecast skill of the Madden–Julian oscillation in two Canadian atmospheric models. *Mon. Weather Rev.* **136**, 4130–4149 (2008).
- Cho, K. et al. Learning phrase representations using rnn encoder–decoder for statistical machine translation. In *Proc. 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* 1724–1734 (Association for Computational Linguistics, 2014).
- Abadi, M. et al. Tensorflow: large-scale machine learning on heterogeneous systems <https://www.tensorflow.org/> (2015).
- Silini, R. MJO–prediction–networks. *GitHub* <https://github.com/riccardosilini/MJO-prediction-networks> (2021).

## ACKNOWLEDGEMENTS

This work received funding from the European Union’s Horizon 2020 research and innovation program under the Marie Skłodowska Curie Grant Agreement No 8138444. C.M. also acknowledges funding by the Spanish Ministerio de Ciencia, Innovación y Universidades (PGC2018-099443-B-I00), and the ICREA ACADEMIA program of Generalitat de Catalunya. We would like to thank Laura Ferranti, Linus Magnusson, and Nikolaos Mastrantonas for their expertise and useful discussions. Portions of this work are modifications based on work created and shared by Google and used according to terms described in the Creative Commons 4.0 Attribution License.

## AUTHOR CONTRIBUTIONS

R.S. and M.B. designed the study. R.S. performed the study and wrote the manuscript. M.B. and C.M. supervised the study. All authors reviewed the manuscript.

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Correspondence** and requests for materials should be addressed to Riccardo Silini.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021