Deep learning for bias correction of MJO prediction

H. Kim, Y. G. Ham, Y. S. Joo & S. W. Son

Producing accurate weather prediction beyond two weeks is an urgent challenge due to its ever-increasing socioeconomic value. The Madden-Julian Oscillation (MJO), a planetary-scale tropical convective system, serves as a primary source of global subseasonal (i.e., targeting three to four weeks) predictability. During the past decades, operational forecasting systems have improved substantially, while the MJO prediction skill has not yet reached its potential predictability, partly due to the systematic errors caused by imperfect numerical models. Here, to improve the MJO prediction skill, we blend the state-of-the-art dynamical forecasts and observations with a Deep Learning bias correction method. With Deep Learning bias correction, multi-model forecast errors in MJO amplitude and phase averaged over four weeks are significantly reduced by about 90% and 77%, respectively. Most models show the greatest improvement for MJO events starting from the Indian Ocean and crossing the Maritime Continent.
Accurate prediction beyond the two-week limit of atmospheric predictability is extremely valuable to society and the economy. In particular, reliable forecasts in the subseasonal range (i.e., timescale of 3–4 weeks) provide vital information about hazardous weather threats, such as floods, heat waves, and cold spells, which are extremely important for risk managers, stakeholders, and policymakers. The value of subseasonal forecasts has been recognized by society and the scientific community, and tremendous international efforts towards making reliable subseasonal forecasts are underway. The Madden-Julian Oscillation (MJO), the dominant mode of tropical subseasonal variability, is a planetary-scale organized convection-circulation coupled system with a typical period of 30–60 days, characterized by an eastward propagation, especially during boreal winter. As the MJO develops and propagates, anomalous diabatic heating leads to the formation of an anomalous Rossby wave source. This excites the Rossby waves to propagate into the extratropics, modulating the weather events therein. For example, studies have shown significant influences of the MJO on tropical cyclones, extreme temperature and precipitation, storm tracks, atmospheric blocking events, atmospheric rivers, tornadoes, and weather in the Arctic and Antarctic, among many others. Due to its far-reaching global impacts and quasiperiodic nature, the MJO is recognized as one of the leading sources of global climate predictability for the subseasonal timescale that bridges the gap between the traditional weather (i.e., from one day to 2 weeks) and seasonal (i.e., from 2 months to 1 year) forecast ranges.

Recent advances in theoretical understanding, improved numerical models, and international collaborative activities on field campaigns and forecast experiments have promoted advances in MJO forecasting. Now, the state-of-the-art dynamical forecast systems are able to predict the MJO up to 3 weeks in advance, a remarkable improvement since the early 2000s. However, due to errors originating from imperfect numerical models, the MJO prediction skill has not reached its theoretical predictability, which is known to be ~7 weeks. This indicates that there is considerable room for improvement.

One of the greatest challenges in current dynamical forecast systems is the fast damping of the MJO signal as the forecast lead time increases, which results in a rapid decrease of forecast skill. This systematic damping of the MJO convection signal is particularly apparent when the MJO starts over the Indian Ocean and is expected to propagate through the Maritime Continent and move further into the western Pacific. The frequency of MJO events not crossing the Maritime Continent in forecast models is more than twice as large as it is in observations, known as the Maritime Continent prediction barrier. Given that the MJO prediction alone presents considerable systematic biases, the global weather forecast beyond 2 weeks is an even more daunting task.

Model deficiencies in simulating realistic MJO events are partially due to our poor understanding of the underlying physics. Therefore, more efforts on process-level diagnostics are suggested to further improve MJO simulation and prediction. Concurrently, post-processing of numerical forecasts has been proven to be a powerful tool to improve forecasts when models display systematic biases. A recent study has shown an increase of MJO prediction skill by correcting model biases with a linear statistical model. Deep learning (DL) techniques have been proven to be a powerful tool for identifying weather and climate patterns, sub-grid scale parameterizations, weather and climate predictions, and post-processing of numerical weather forecasts (shorter than 7 days). However, post-processing with DL methods has not yet been applied to MJO forecasts.

In this study, we utilize DL as a bias correction method to improve MJO forecasts. We demonstrate that the DL post-processing substantially reduces the MJO forecast errors from the state-of-the-art dynamical forecasting systems, thus making strides towards improving global extended range forecasts.

**Results**

**Improved MJO prediction with deep learning bias correction.**

Figure 1 highlights the advantage of Deep learning bias correction (DL-correction) for MJO forecasts. It shows the multi-model mean of predicted Real-time Multivariate MJO indices (RMMs) composite on a phase-space diagram as a function of initial MJO phases and forecast lead days from day 1 to day 28 (4 weeks). Predictions from the original Subseasonal-to-seasonal (S2S) reforecasts and DL-correction for each forecast target years are composited and compared with observations (see “Methods” section). The composite results of individual models are displayed in Supplementary Fig. 2. Several key results strongly demonstrate the benefit of the DL-correction on MJO forecasts throughout all MJO phases. A large discrepancy between S2S reforecasts and observations is clearly shown on day 1. Most S2S models forecast either weaker (phases 2 and 3) or stronger amplitude (phases 6 and 7), or phase (θ) shifts relative to the observations on day 1. The DL-correction reduces those systematic errors, making the day 1 and the following forecasts closer to the observations in all models and throughout all MJO phases (Fig. 1 and Supplementary Fig. 2).

To evaluate the forecast errors quantitatively, the bivariate root-mean-squared error (BMSE, see “Methods” section) is calculated as a function of initial MJO phases from the
composites shown in Fig. 1. The amplitude error (BMSEa) appears in the S2S models from the beginning of the forecasts for most phases, with predominant errors in phases 2 and 3 (Fig. 2a). This amplitude error reduces substantially when DL-correction is applied (Fig. 2c). Whether amplitude errors are large or small in individual S2S models, they all become similar after the DL-correction (Supplementary Fig. 3). For example, during the first 2 weeks, the ECMWF-Cy43r3 has the largest MJO amplitude error in phases 2 and 3, while the NASA-GEOS5 possesses a large amplitude error in phases 6 and 7 (Supplementary Figs. 2 and 3). Nevertheless, these amplitude errors become negligible after DL-correction (Supplementary Fig. 3). The BMSEa averaged over 4 weeks and eight phases is summarized in Fig. 3. After DL-correction, BMSEa in S2S reforecasts is about 90% reduced in multi-model mean and about 70–94% reduced in individual models (Fig. 3).

The majority of S2S models show large errors emanating from MJO phase (BMSEP) (Fig. 2b, Supplementary Fig. 4), indicating the inability of current forecasting systems to predict the main location of the MJO realistically. This phase error reduces substantially when DL-correction is applied (Fig. 2d). The BMSEP averaged over 4 weeks and eight phases is summarized in Fig. 3. After DL-correction, BMSEP in S2S reforecasts is about 90% reduced in multi-model mean and about 70–94% reduced in individual models (Fig. 3).

To assess prediction skill and predictability of the MJO, two additional verification metrics are applied (see “Methods” section). The multi-model mean BCOR of the DL-correction is consistently higher than the S2S models up to 4 weeks (Supplementary Fig. 7). An increased signal and reduced noise after the DL-correction, which results in a higher MJO predictability than the original ECMWF-Cy43r3 reforecasts, is also shown (Supplementary Fig. 8).

Eastward propagation of the predicted MJO. As mentioned earlier, dynamical forecasts have struggled to accurately forecast the MJO propagation over the Maritime Continent, especially when the forecast is initialized in the Indian Ocean (phases 2 and 3). Such a prediction barrier can be partly explained by the basic state moisture biases that degrade the physical processes associated with the eastward propagation of the MJO. Even the ECMWF-Cy43r3, the best MJO prediction model, simulates an exaggerated Maritime Continent barrier (Supplementary Fig. 2a). Figure 4 shows the reconstructed Outgoing Longwave Radiation (OLR) and zonal wind at 850 hPa (U850) anomalies obtained by projecting the RMMs starting from phase 2 onto the

![Fig. 2](https://example.com) **Fig. 2 Forecast errors in multi-model mean S2S reforecasts and DL-corrections.** Multi-model mean bivariate root-mean-squared error (BMSE) as a function of initial MJO phases and forecast lead days for (top) Subseasonal-to-seasonal (S2S) reforecasts and (bottom) Deep learning (DL)-corrections. BMSE is divided into (a, c) amplitude error (BMSEa) and (b, d) phase error (BMSEP) for the Real-time Multivariate MJO indices (RMMs) composite shown in Fig. 1. Values are multiplied by 100. Note that a three-day moving average is applied.
© 2021 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

Figure 3: Averaged forecast errors in individual models. Bivariate root-mean-squared amplitude error (BMSEA) (darker colours) and phase error (BMSEP) (lighter colours) for Subseasonal-to-seasonal (S2S) reforecasts (blue) and Deep learning (DL)-corrections (red) averaged over 4 weeks and eight phases shown in Fig. 2, Supplementary Figs. 3 and 4. For ECMWF-Cy43r3 and NCAR-CESM1, the orange error bar denotes the 95% confidence interval based on the bootstrap method. NCEP-GEFS = National Centres for Environmental Prediction Environmental Modelling Centre Global Ensemble Forecast System; NASA-GEOSS = National Aeronautics and Space Administration Global Modelling and Assimilation Office Goddard Earth Observing System; Navy-ESPC = Naval Research Laboratory Navy Earth System Prediction Capability; RSMAS-CCSM4 = Community Climate System Model version 4 run at the University of Miami Rosenstiel School for Marine and Atmospheric Science; ESRL-FIM = Earth System Research Laboratory Flow-Following Icosahedral Model; NCAR-CESM1 = National Centre for Atmospheric Research Community Earth System Model Version 1; KMA/UKMO-GloSea5 = Korea Meteorological Administration-UK Met Office coupled Global Seasonal forecast; ECMWF-Cy43r3: European Centre for Medium-Range Weather Forecast models version Cy43r3.

Discussion

This study demonstrates the power of Deep learning to be used as a post-processing tool to correct the systematic biases that evolve during MJO forecasts. The errors emanating from MJO amplitude and phase in the dynamical model forecasts are both reduced substantially by DL-correction. The results show that the performance of poor models becomes comparable to the best model after DL-correction. This implies that the differences in the model’s performance mainly originate from the systematic errors, rather than the random errors, and the DL method effectively works to minimize them regardless of their amplitude as long as the forecast errors are systematic. Given that huge efforts have been made in operational centres to upgrade their models to reach the level of the world’s best model for MJO forecasting (i.e., the ECMWF model), our results show promise for conserving both human and computational resources. Moreover, while the characteristics of systematic biases can change as operational forecast models undergo periodic upgrades, the DL-correction method developed here can be easily adapted to appropriately reflect the changes. Most importantly, because the model is simple, the approach developed in this study can easily be implemented into real-time MJO forecasts, which in turn can help end-user preparedness and eventually protect lives and properties vulnerable to various hazardous extreme events related to the MJO.

However, although the Deep learning approach can assist in correcting model biases, continuous effort towards developing the dynamical forecast system to minimize the inherent errors is the key for making MJO forecasts reach their potential predictability. Note that the improved MJO prediction with DL-correction does not guarantee an improved prediction of MJO-related phenomena such as tropical cyclones, monsoons, and midlatitude teleconnections, because they rely on both the MJO and the background state within the model. Therefore, further improvements in dynamical models and initialization are fundamental to ultimately improve the S2S predictions.
Methods

Hindcast and validation datasets. In this study, we use long-term reforecasts from the international Subseasonal-to-Seasonal prediction (S2S) project, and from the NCAR Community Earth System Model v1.43, which follows the SubX protocol. Hereafter, we refer to these reforecasts as S2S reforecasts for convenience. Supplementary Table 1 provides information on the eight S2S reforecast models, including initialization interval, ensemble size, reforecast period, and sample size. Note that the reforecast period of ECMWF-Cy43r3 and KMA/UKMO-GloSea5 are different compared to the rest of the models (Supplementary Table 1). These eight models have shown good performance in MJO prediction48,49.

To identify the MJO events, the Real-time Multivariate MJO (RMM)50 index is calculated with the daily mean zonal wind at 850 hPa (U850) and 200 hPa (U200) from the ECMWF Interim Reanalysis51 and Outgoing Longwave Radiation (OLR) from the NOAA Advanced Very High-Resolution Radiometer52; these are referred to as observation for brevity. All S2S reforecasts and observations are interpolated onto a 1° longitude and 1° latitude grid. The method for calculating anomalies and RMM indices follows previous studies48,49.

Deep learning bias-correction model. The Deep learning bias correction (hereafter, DL-correction) model utilizes the Long Short-Term Memory (LSTM), which has been proven to be powerful for time sequence modelling53,54. The LSTM model is often used for making predictions on data not used in the training procedure. In this study, we use long-term reforecasts (1997–2016) are used to train the LSTM model, and so on. The LSTM model is built at every target year, forecast lead time, MJO phase, and each model individually due to the limited training sample size. In this real forecast procedure, biases are still evaluated 10,000 times and the 2.5th and 97.5th percentile values are used to define the confidence level if it lies outside the 2.5th or 97.5th percentile. The same process is performed for the individual forecasts around the ensemble mean (i.e., the forecast spread), and both quantities depend on the forecast lead time (τ), the reference quantity (e.g., OLR and zonal winds), and the model ensemble size.

DL-correction procedure. The one-year-out cross-validation (LOOCV) procedure is often used for making predictions on data not used in the training period and for a relatively small dataset. For example, to process DL-correction on the target year 1997 in ECMWF-Cy43r3, the modelled/observed RMMs of MJO events from the rest of 19 years (from 1998 to 2016) are used to train the LSTM model. Then, the weighting coefficients and biases obtained during the training period are directly applied to the modelled RMMs of MJO events in 1997 (target year). This results in DL-corrected MJO predictions in 1997. For the target year 1998, MJO events from the rest of the 19 years (1997 and from 1999 to 2016) are used to train the LSTM model, and so on. The LSTM model is built at every target year, forecast lead time, MJO phase, and each model individually due to their unique systematic biases. Note that, for given input data sets that were randomly selected, the LOOCV produces very similar results for every target year, indicating that the LSTM model is stable.

We also perform the DL-correction in a real forecast manner. The ECMWF-Cy43r3 reforecasts during the first 10 years (1997–2006) are used as the training period to build the LSTM model, and the remaining independent 10 years (2007–2016) are evaluated. In this real forecast procedure, biases are still significantly reduced compared to the raw S2S reforecasts (not shown), but larger than those by the LOOCV approach due to the limited training sample size.

The MJO amplitude (A) and phase (θ) for the observation (Ao) and reforecast (Ar) are defined as:55:

\[ A_o(t) = \sqrt{O_1(t)^2 + O_2(t)^2} \]

\[ A_r(t, τ) = \sqrt{M_1(t, τ)^2 + M_2(t, τ)^2} \]

\[ θ_o(t) = \tan^{-1} \frac{O_1(t)}{O_2(t)} \]

\[ θ_r(t, τ) = \tan^{-1} \frac{M_1(t, τ)}{M_2(t, τ)} \]

where \( O_1(t) \) and \( O_2(t) \) are the observed RMM1 and RMM2 at time t, and \( M_1(t, τ) \) and \( M_2(t, τ) \) are the modelled RMM1 and RMM2 at time t with a lead time τ days. On the two-dimensional phase-space diagram, the MJO phase is defined as the azimuth of the RMM1 and RMM2 combination and is usually divided into eight phases depending on the location of the MJO convection48. The MJO amplitude is determined based on the distance of the azimuth point from the origin, and an MJO event is defined when the observed MJO amplitude (\( A_o \)) exceeds 1.0 on initial day 0. Although the MJO is most active during boreal winter and thus has the highest forecast skill in this season56, we use MJO events from all seasons due to the limited sample size. For the same reason, we group two MJO phases (2 & 3, 4 & 5, 6 & 7, and 8 & 1) when training the LSTM model.

The Deep learning bias correction (hereafter, DL-correction) model utilizes the Long Short-Term Memory (LSTM), which has been proven to be powerful for time sequence modelling53,54. The LSTM model is often used for making predictions on data not used in the training procedure. In this study, we use long-term reforecasts (1997–2016) are used to train the LSTM model, and so on. The LSTM model is built at every target year, forecast lead time, MJO phase, and each model individually due to the limited training sample size. In this real forecast procedure, biases are still evaluated 10,000 times and the 2.5th and 97.5th percentile values are used to define the confidence level if it lies outside the 2.5th or 97.5th percentile. The same process is performed for the individual forecasts around the ensemble mean (i.e., the forecast spread), and both quantities depend on the forecast lead time (τ), the reference quantity (e.g., OLR and zonal winds), and the model ensemble size.

\[ \text{BMSE}(t) = \frac{1}{N} \sum_{i=1}^{N} (O_i(t) - M_i(t, τ))^2 + (O_i(t) - M_i(t, τ))^2 \]

where N is the number of MJO events. The BMSE can be separated into the error emanating from amplitude error (BMSEA) and phase error (BMSEP) as:

\[ \text{BMSEA}(t) = \frac{1}{N} \sum_{i=1}^{N} A_o(i, t) - A_r(i, t)^2 \]

\[ \text{BMSEP}(t) = \frac{1}{N} \sum_{i=1}^{N} 2A_o(i, t)A_r(i, t)\cos(θ_o(i, t) - θ_r(i, t)) \]

where \( A_o(i, t) \) and \( A_r(i, t) \) denote ensemble mean and prime presents individual ensembles' deviations from the ensemble mean. In this formulation, the signal refers to the variability of the ensemble mean while the noise refers to the variability of individual forecasts around the ensemble mean (i.e., the forecast spread), and both quantities depend on the forecast lead time (τ). The ECMWF-Cy43r3 is used to estimate the signal and noise due to its relatively large ensemble size and high MJO skill.

Assessment of MJO predictions. To evaluate MJO forecast quality, the bi-variate correlation coefficient (BCOR)57 and bivariate root-mean-square-error (BMSE)57 are calculated between the predicted and observed RMM indices as a function of forecast lead days as follows:

\[ \text{BCOR}(τ) = \frac{\sum_{i=1}^{N} O_i(τ)M_i(t, τ) - \sum_{i=1}^{N} O_i(τ)\sum_{i=1}^{N} M_i(t, τ)}{\sqrt{\sum_{i=1}^{N} O_i(τ)^2\sum_{i=1}^{N} M_i(t, τ)^2}} \]

\[ \text{BMSE}(t) = \frac{1}{N} \sum_{i=1}^{N} (O_i(t) - M_i(t, τ))^2 + (O_i(t) - M_i(t, τ))^2 \]

where N is the number of MJO events. The BMSE can be separated into the error emanating from amplitude error (BMSEA) and phase error (BMSEP) as:

\[ \text{BMSEA}(τ) = \frac{1}{N} \sum_{i=1}^{N} A_o(i, τ) - A_r(i, τ)^2 \]

\[ \text{BMSEP}(τ) = \frac{1}{N} \sum_{i=1}^{N} 2A_o(i, τ)A_r(i, τ)\cos(θ_o(i, τ) - θ_r(i, τ)) \]

where \( A_o(i, τ) \) and \( A_r(i, τ) \) denote ensemble mean and prime presents individual ensembles' deviations from the ensemble mean. In this formulation, the signal refers to the variability of the ensemble mean while the noise refers to the variability of individual forecasts around the ensemble mean (i.e., the forecast spread), and both quantities depend on the forecast lead time (τ). The ECMWF-Cy43r3 is used to estimate the signal and noise due to its relatively large ensemble size and high MJO skill.

Confidence interval. The statistical significance test is performed with ECMWF-Cy43r3 and NCAR-CESSM1 only, due to their relatively large ensemble sizes. The confidence interval of DL-correction results is calculated using the bootstrap method. We randomly select 11 ensemble members from the S2S reforecasts with allowing overlap to calculate the ensemble-averaged BMSE. This process is repeated 10,000 times and the 2.5th and 97.5th percentile values are used to define the 95% confidence interval. The ensemble-averaged BMSE of DL-correction value (Fig. 3, Supplementary Figs. 3, 4) is significant at the 95% confidence level if it lies outside the 2.5th or 97.5th percentile. The same process is performed for the
reconstructed OLR and U850 to check whether the composited anomalies from DL-correction is significantly different from the S2S forecast results (Fig. 4).

Data availability
Data related to this paper can be downloaded from: ERA-Interim, http://apps.ecmwf.int/datasets/data/interim_full_daily; NOAA OLR, https://www.esrl.noaa.gov/psd/data/grided/data.interp.OLRhtml; the SubX and NCAR-CESM1 reforecasts, http://irdl. ldeo.columbia.edu/SOURCES/.Models/.SubX; The S2S reforecasts, https://apps.ecmwf. int/datasets/data/s2s/.

Code availability
TensorFlow (https://www.tensorflow.org) libraries were implemented to formulate the forecast model using the LSTM. The codes used in the current analysis are available at https://zenodo.org/record/4601794 and from the corresponding author upon reasonable request.

Received: 29 September 2020; Accepted: 22 April 2021; Published online: 25 May 2021

References
1. E. National Academies of Sciences, Medicine, Next Generation Earth System Prediction: Strategies for Subseasonal to Seasonal Forecasts. (The National Academies Press, Washington, DC, 2016), pp. 350.
2. Mariotti, A. et al. Windows of opportunity for skilful forecasts seasonal to seasonal and beyond. B Am. Meteorol. Soc. 101, E608–E625 (2020).
3. Madden, R. & Julian, P. Further evidence of global-scale, 5-day pressure waves. J. Atmos. Sci. 29, 1464 (1972).
4. Sardeshmukh, P. D. & Hoskins, B. J. The generation of global rotational by steady idealized tropical divergence. J. Atmos. Sci. 45, 1228–1251 (1988).
5. Hoskins, B. J. & Karoly, D. J. The steady linear response of a spherical to thermal and orographic forcing. J. Atmos. Sci. 38, 1179–1196 (1981).
6. Sardeshmukh, P. D. & Hoskins, B. J. The generation of global rotational by steady idealized tropical divergence. J. Atmos. Sci. 45, 1228–1251 (1988).
7. Klotzbach, P. J. On the Madden–Julian oscillation in the intraseasonal variability hindcast experiment (ISVHE)∗. J. Clim. 27, 4531–4543 (2014).
8. I. S. Kang, J. Shukla, in The Asian Monsoon. (Springer Praxis Books. Springer, Berlin, Heidelberg, 2006), pp. 585–612.
9. Wu, J. & Jin, F.-F. Improving the MJO forecast of S2S operation models by correcting their biases in linear dynamics. Geophys. Res. Lett. 48, e2020GL091930 (2021).
10. Barnes, E. A., Hurrell, J. W., Ebert-Uphoff, I., Anderson, C. & Anderson, D. Viewing forced climate patterns through an AI lens. Geophys. Res. Lett. 46, 13389–13398 (2019).
11. Neena, J. M., Lee, J. Y., Waliser, D., Wang, B. & Jiang, X. Deep learning for scientific inference from geophysical data: the Madden–Julian oscillation as a test case. arXiv preprint arXiv:1902.04621 (2019).
12. Barnes, E. et al. Indicator patterns of forced change learned by an artificial neural network. J. Adv. Modeling Earth Syst. 12, e2020MS002195 (2020).
13. Gentile, P., Pritchard, M., Rasp, S., Reinaudi, G. & Yacalis, G. Can machine learning break the convection parameterization deadlock? Geophys. Res. Lett. 45, 5742–5751 (2018).
14. Rasp, S., Pritchard, M. S. & Gentile, P. Deep learning to represent subgrid processes in climate models. Proc. Natl. Acad. Sci. 115, 9684–9689 (2018).
15. Arcomano, T. et al. A machine learning-based global atmospheric forecast model. Geophys. Res. Lett. 47, e2020GL087776 (2020).
16. Weny, J. A., Durran, D. R. & Caruana, R. Can machines learn to predict weather? using deep learning to predict gridred 500-hPa geopotential height from historical weather data. J. Adv. Modeling Earth Syst. 11, 2680–2693 (2019).
17. Harn, Y.-C., Kim, J.-H. & Luo, J.-J. Deep learning for multi-year ENSO forecasts. Nature 573, 568–572 (2019).
18. Rasp, S. & Lerch, S. Neural networks for postprocessing ensemble weather forecasts. Mon. Weather Rev. 146, 3885–3900 (2018).
19. Chapman, W. E., Subramanian, A. C., Dele Monache, L., Xie, S. P. & Ralph, F. M. Improving atmospheric river forecasts with machine learning. Geophys. Res. Lett. 46, 10627–10635 (2019).
20. Wheeler, M. C. &endon, 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).
21. Jiang, X. Key processes for the eastward propagation of the Madden–Julian Oscillation based on multimodel simulations. J. Geophys. Res.: Atmospheres 122, 755–770 (2017).
22. Lim, Y., Son, S.-W. & Kim, D. MJO prediction skill of the seasonal-to-seasonal prediction models. J. Clim. 31, 4075–4094 (2018).
23. Kim, H.-M., Webster, P. J., Toma, V. E. & Kim, D. Predictability and prediction skill of the MJO in two operational forecasting systems. J. Clim. 27, 3564–3578 (2014).
24. Vitart, F. et al. The subseasonal to seasonal (S2S) prediction project database. Bull. Am. Meteorol. Soc. 98, 163–173 (2017).
25. Pegion, K. et al. The subseasonal experiment (SubX): a multimodel subseasonal prediction experiment. Bull. Am. Meteorol. Soc. 100, 2043–2060 (2019).
26. Richter, J. et al. Subseasonal prediction with and without a well-represented stratosphere in CESM1. Weather Forecast. 35, 2589–2602 (2020).
Acknowledgements

H.K. was supported by NSF grant AGS-1652289 and the Brain Pool program funded by the Ministry of Science and ICT through the National Research Foundation of Korea (NRF) 2019H1D3A2A01102234. Y.G.H. and Y.S.J. are supported by the NRF under Grant No. NRF-2020R1A2C1010205. S.W.S. was supported by the Korea Meteorological Administration Research and Development Program under Grant KMI2020-01010.

Author contributions

H.K. designed the research, performed model run, analyzed the model output, and generated figures. H.K., Y.G.H. and Y.S.J. built the LSTM model. H.K., Y.G.H., Y.S.J. and S.W.S. discussed the results and contributed to the writing of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-021-23406-3.

Correspondence and requests for materials should be addressed to H.K.

Peer review information Nature Communications thanks Samson Hagos, Aneesh Subramanian and Baoqiang Xiang for their contribution to the peer review of this work. Peer reviewer reports are available.

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