



**RESEARCH ARTICLE**

# Configuration and hindcast quality assessment of a Brazilian global sub-seasonal prediction system

Bruno S. Guimarães<sup>1,2</sup>  | Caio A. S. Coelho<sup>1</sup> | Steven J. Woolnough<sup>2</sup> | Paulo Y. Kubota<sup>1</sup>  | Carlos F. Bastarz<sup>1</sup> | Silvio N. Figueroa<sup>1</sup> | José P. Bonatti<sup>1</sup> | Dayana C. de Souza<sup>1</sup>

<sup>1</sup>Center for Weather Forecast and Climate Studies, National Institute for Space Research, Cachoeira Paulista, Brazil

<sup>2</sup>National Centre for Atmospheric Science, Department of Meteorology, University of Reading, Reading, UK

**Correspondence**

B.S. Guimarães, CPTEC, Rodovia Presidente Dutra, Km 39, Cachoeira Paulista, SP, 12630-000, Brazil.  
Email: guimara.bruno@gmail.com

**Funding information**

CASC thanks CNPq, 304586/2016-1; Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and University of Reading, GS18-179; Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), 2015/50687-8

**Abstract**

This article presents the Centre for Weather Forecast and Climate Studies (CPTEC) developments for configuring a global sub-seasonal prediction system and assessing its ability in producing retrospective predictions (hindcasts) for meteorological conditions of the following 4 weeks. Six Brazilian Global Atmospheric Model version 1.2 (BAM-1.2) configurations were tested in terms of vertical resolution, deep convection and boundary-layer parametrizations, as well as soil moisture initialization. The aim was to identify the configuration with best performance when predicting weekly accumulated precipitation, weekly mean 2 m temperature (T2M) and the Madden–Julian Oscillation (MJO) daily evolution. Hindcasts assessment was performed for 12 extended austral summers (November–March, 1999/2000–2010/2011) with two start dates for each month for the six configurations and two ensemble approaches. The first approach, referred to as Multiple Configurations Ensemble (MCEN), was formed of one ensemble member from each of the six configurations. The second, referred to as Initial Condition Ensemble (ICEN), was composed of six ensemble members produced with the chosen configuration as the best using an empirical orthogonal function (EOF) perturbation methodology. The chosen configuration presented high correlation and low root-mean-squared error (RMSE) for precipitation and T2M anomaly predictions at the first week and these indices degraded as lead time increased, maintaining moderate performance up to week-4 over the tropical Pacific and northern South America. For MJO predictions, this configuration crossed the 0.5 bivariate correlation threshold in 18 days. The ensemble approaches improved the correlation and RMSE of precipitation and T2M anomalies. ICEN improved precipitation and T2M predictions performance over eastern South America at week-3 and over northern South America at week-4. Improvements were also noticed for MJO predictions. The time to cross the above-mentioned threshold increased to 21 days for MCEN and to 20 days for ICEN.

**KEYWORDS**

forecast verification, intraseasonal variability, MJO

## 1 | INTRODUCTION

Forecasting for the time-scale between 2 weeks and 2 months is known as sub-seasonal prediction (Vitart *et al.*, 2017). This type of forecast is a major challenge because the predictability contribution from the atmospheric initial conditions is reduced compared to shorter (weather) time-scales, and the predictability from slowly varying boundary conditions is small for 1–2 week averages, typically the focus of sub-seasonal prediction, compared to seasonal time-scales (Kumar *et al.*, 2011; Lin *et al.*, 2016). The main source of predictability for sub-seasonal forecasting is the Madden–Julian Oscillation (MJO) (Zhang, 2013). However, general-circulation models (GCMs) still show limitations in simulating this oscillation (Green *et al.*, 2017; Wang *et al.*, 2018), even with important improvements achieved in recent years (Saha *et al.*, 2014; Vitart, 2014). As a consequence of these limitations, the predictive ability of GCMs in the sub-seasonal scale is lower than in the weather and seasonal scales (Zhu *et al.*, 2014). For example, de Andrade *et al.* (2019) showed the limited predictive ability of GCMs for sub-seasonal precipitation predictions for lead times beyond 15 days. In general, the GCMs show modest performance in specific areas such as the equatorial regions of the Atlantic and Pacific Oceans and over a few regions in South America at this lead time.

In spite of these results, a tendency toward improvements in GCMs for sub-seasonal predictions is seen and several meteorological centres currently operationally produce this type of forecast (Vitart, 2004; Hudson *et al.*, 2011; Mastrangelo *et al.*, 2012; Liu *et al.*, 2017; Weber and Mass, 2017; Liang and Lin, 2018). The Centre for Weather Forecast and Climate Studies [Centro de Previsão de Tempo e Estudos Climáticos (CPTEC)], which plays a leading role in South America with respect to weather and seasonal forecasts, is now following the trend of these meteorological centres and started to develop a sub-seasonal prediction system. This is motivated by the fact that, as in seasonal forecasting, South America is located in a privileged region for sub-seasonal prediction, with GCMs showing better predictive ability in this region when compared to other continental regions (Li and Robertson, 2015; de Andrade *et al.*, 2019).

The identified evolution in sub-seasonal predictions is mainly due to improvements in the representation of the MJO in GCMs. For example, The European Centre for Medium-range Weather Forecasts (ECMWF) showed a mean gain of one day in MJO prediction performance per year (Vitart, 2014). This indicates that in addition to improvements in predictive ability for

a phenomenon that manifests in the tropical region, there is also associated improvement in the extratropics due to teleconnections generated by the MJO (Vitart, 2017).

These findings are documented, in large part, thanks to the effort generated by the Sub-seasonal to Seasonal (S2S) Prediction Project. This project was launched jointly by the World Weather Research Programme (WWRP) and the World Climate Research Programme (WCRP) of the World Meteorological Organization (WMO) and aims to improve forecast skill and understanding on the sub-seasonal to seasonal time-scales and also to promote its uptake by operational centres and by the applications community. Currently, the S2S Prediction Project stores and disseminates near-real-time forecasts and hindcasts of eleven operational and research centres for research purposes (Vitart *et al.*, 2017).

The Brazilian Global Atmospheric Model (BAM: Figueroa *et al.*, 2016) is the current CPTEC global atmospheric model for weather forecasting. The performance of this model for sub-seasonal predictions has not been documented yet. Therefore, this study presents the first outcomes of this model for sub-seasonal predictions and aims to determine which model configuration presents the best performance for this time-scale. Special attention is given to characteristics such as vertical resolution, deep convection and boundary-layer parametrizations, as well as initialization of the soil moisture, because they have an important influence on the MJO and sub-seasonal predictions. A similar approach was taken by Green *et al.* (2017) in order to identify a model configuration with best performance when producing MJO predictions. Green *et al.* (2017) evaluated the MJO predictive ability in multi-physics and multimodel global ensembles, by performing two sets of hindcasts in order to test the impact of using the Grell and Freitas (2014) versus the revised simplified Arakawa–Schubert (Han and Pan, 2011) deep convection parametrization. They revealed that the Grell and Freitas (2014) convection parametrization showed better MJO prediction performance than the revised simplified Arakawa–Schubert scheme.

The article is organized as follows. Section 2 presents the model description, datasets used for model initialization and hindcast quality assessment, the definition of the experiments, ensemble approaches and the metrics used for evaluation. The retrospective performance of the produced precipitation, 2 m temperature (T2M) and MJO predictions with different BAM configuration experiments, including two ensemble approaches, is shown in Section 3. The final section is intended for the conclusion.

## 2 | MODEL DESCRIPTION, DATASETS AND EXPERIMENTAL CONFIGURATIONS, EVALUATION METRICS AND ENSEMBLE APPROACHES

### 2.1 | Model description

The model version used in this study is the current operational CPTEC global spectral atmospheric model developed for numerical weather forecasting, which is known as BAM version 1.2 (BAM-1.2). This model version has different options for dynamical and physics parametrizations. The Eulerian advection scheme option with a two-time-level semi-Lagrangian scheme for moisture transport and microphysics prognostic variables is used in this study. The physical processes of this recent operational version are similar to the previous version (BAM-1.0) and are described in Figueroa *et al.* (2016), which are: microphysics from Morrison *et al.* (2009), the IBIS-CPTEC surface model (Kubota, 2012), the long-wave radiation scheme developed by Chou *et al.* (2001) (CLIRAD-LW), the short-wave radiation scheme developed by Chou and Suarez (1999) (CLIRAD-SW), the latter modified by Tarasova and Fomin (2000), the modified Mellor–Yamada diffusion scheme for the planetary boundary layer (PBL), which is based on Mellor and Yamada (1982) and is referred to as dry-PBL, and the modified Grell–Dévényi deep convection scheme, which is based on Grell and Dévényi (2002). The two new BAM-1.2 components are the Bretherton–Park moist diffusion scheme (Bretherton and Park, 2009) for the PBL, which is referred to as moist-PBL, and the revised version of the simplified Arakawa–Schubert deep convection scheme (Han and Pan, 2011), which were recently implemented. Following Yu *et al.* (2006), aerosol optical depth in the first 2 km of the atmosphere is specified as 0.22 over the continents and as 0.14 over the oceans. The horizontal resolution used in this study is triangular truncation at 126 waves (TQ126, corresponding to a grid of approximately 1.0° in latitude and longitude) and two vertical resolutions are examined: 42 (L42) and 64 (L64) sigma vertical levels.

One of the objectives of this study is to investigate the performance of the two PBL and deep convection schemes mentioned above for sub-seasonal predictions. The main difference of the newly implemented moist-PBL Bretherton–Park scheme compared to the dry-PBL modified Mellor–Yamada diffusion scheme is the use of moist conserved variables and an explicit entrainment closure for convective layers. Regarding the convection schemes, the revised simplified Arakawa–Schubert and the previously implemented modified Grell–Dévényi deep convection parametrization schemes were both derived from

**TABLE 1** Hindcasts initialization dates for the six BAM-1.2 configurations

Days	Months	Years
03 and 14	November	1999–2010
01 and 15	December	1999–2010
04 and 14	January	2000–2011
01 and 15	February	2000–2011
03 and 14	March	2000–2011

Grell (1993), in which the cloud spectrum of the original Arakawa and Schubert (1974) scheme is reduced to a single cloud using a single mass flux closure. The main differences between these convection schemes implemented in BAM-1.2 are the fractional entrainment rate and convection trigger formulations; see Han and Pan (2011) and Figueroa *et al.* (2016) for additional information.

### 2.2 | Datasets and experimental configurations

Sub-seasonal hindcasts were performed over the period defined as the extended austral summer (from November to March) over the 1999/2000–2010/2011 period. Two hindcasts for two selected start dates were produced for each month of a given year. Start dates vary from one month to the next and are presented in Table 1. Each hindcast was run for the following 35 days after the start date (35 days of lead time). For the production of these hindcasts, BAM-1.2 was not coupled with an ocean model. Instead, the total sea-surface temperature (SST) field (not the anomaly) of each start date was kept constant during the 35 days of integration (persisted SST). It is worth highlighting that coupled ocean–atmosphere processes are recognized as being important on these time-scales (Reichler and Roads, 2005; Chen *et al.*, 2010; Kumar *et al.*, 2011; Shelly *et al.*, 2014), but a number of centres contributing to the S2S database [e.g. Japan Meteorological Agency (JMA) and Environment and Climate Change Canada (ECCC)] produce operational sub-seasonal forecasts using uncoupled systems (Vitart *et al.*, 2017). The CPTEC coupled ocean–atmosphere model version which uses BAM-1.2 as atmospheric component is under development. The sub-seasonal hindcast quality assessment of this coupled model version will be reported in future work.

ERA-Interim reanalyses (Dee *et al.*, 2011) produced by ECMWF were used in two ways. Firstly, the reanalyses were used as atmospheric initial conditions for the hindcasts produced with BAM-1.2. The variables required for initialization are zonal and meridional wind, specific humidity, virtual temperature and ozone in 35 vertical

levels between 1,000 and 50 hPa, surface pressure and SST. The horizontal resolution chosen for initialization was  $1.5^\circ \times 1.5^\circ$  degrees in latitude and longitude, which was interpolated to the model spectral resolution (TQ126L42,  $\sim 100$  km). Secondly, ERA-Interim data were used as reference to assess the quality of the produced hindcasts. The variables selected for this assessment are T2M and zonal and meridional winds at 850 and 200 hPa.

To assess precipitation hindcasts quality, daily data from the Global Precipitation Climatology Project (GPCP: Huffman *et al.*, 2001) were used. GPCP is a product derived from observed rainfall data and precipitation estimates by geostationary and polar-orbiting satellites. The spatial resolution of GPCP is  $1^\circ \times 1^\circ$  degrees in latitude and longitude. Additionally, estimates of outgoing long-wave radiation (OLR) from National Oceanic and Atmospheric Administration (NOAA), with a spatial resolution of  $2.5^\circ \times 2.5^\circ$  degrees in latitude and longitude, were used for assessing the model ability to represent the MJO in conjunction with zonal wind at 850 and 200 hPa from ERA-Interim. This OLR estimation is generated through interpolations in time of polar-orbiting satellite data (for additional information, see Liebmann and Smith, 1996).

Six BAM-1.2 configurations for sub-seasonal prediction have been defined for evaluation. Characteristics such as vertical resolution, convection and boundary-layer parametrizations were evaluated as well as the impact of soil moisture initialization. Single member hindcasts over the 1999/2000–2010/2011 extended austral summer period were produced for each configuration. Five of the configurations were defined by combining two convection schemes, the revised simplified Arakawa–Schubert and the modified Grell–Dévényi, and two vertical diffusion schemes for the PBL, the moist-PBL Bretherton–Park scheme and the dry-PBL modified Mellor–Yamada, and two vertical resolutions, 42 and 64 sigma levels. These physical processes and vertical resolutions of the model were selected because they have an important influence on the predictive ability of the MJO (Vitart, 2014; Boyle *et al.*, 2015; Wang and Chen, 2017) and consequently in the sub-seasonal precipitation and T2M predictions. It is important to highlight that other aspects such as horizontal resolution, radiation and microphysics parametrizations are also important for the good representation of the MJO (Zhang, 2005; Vitart, 2014; Wang *et al.*, 2018). However, such characteristics were not evaluated in the present work.

The sixth configuration evaluates the impact of soil moisture. This characteristic is a source of predictability for the sub-seasonal time-scale and has a positive impact on GCM predictive ability, especially in longer lead times such as when predicting weeks 3 and 4 (Koster *et al.*, 2010). In this part of the study, the mean soil moisture from the

previous month of the start date of each hindcast from the Global Land Data Assimilation System (GLDAS) version 2 product (Rui and Beaudoin, 2017) was used to initialize the soil moisture rather than using the monthly climatological soil moisture estimate in order to assess whether a more realistic soil moisture condition has an impact on the predictive ability of BAM-1.2. The monthly climatological soil moisture data estimates used in this study were obtained from the balance analyses of Willmott *et al.* (1985). Both GLDAS and climatological soil moisture data estimates were interpolated to the model Gaussian grid and converted to soil moisture fraction for hindcasts initialization.

The six examined configurations are defined in Table 2 and are summarized below:

- 42ABC: BAM-1.2 with 42 vertical levels, revised simplified Arakawa–Schubert deep convection parametrization, moist Bretherton–Park boundary-layer parametrization, and climatological soil moisture initialization.
- 64ABC: BAM-1.2 with 64 vertical levels, revised simplified Arakawa–Schubert deep convection parametrization, moist Bretherton–Park boundary-layer parametrization, and climatological soil moisture initialization.
- 42ABG: BAM-1.2 with 42 vertical levels, revised simplified Arakawa–Schubert deep convection parametrization, moist Bretherton–Park boundary-layer parametrization, and soil moisture initialized through the GLDAS version 2 product.
- 42GBC: BAM-1.2 with 42 vertical levels, modified Grell–Dévényi deep convection parametrization, moist Bretherton–Park boundary-layer parametrization, and climatological soil moisture initialization.
- 64GBC: BAM-1.2 with 64 vertical levels, modified Grell–Dévényi deep convection parametrization, moist Bretherton–Park boundary-layer parametrization, and climatological soil moisture initialization.
- 42AMC: BAM-1.2 with 42 vertical levels, revised simplified Arakawa–Schubert deep convection parametrization, dry modified Mellor–Yamada boundary-layer parametrization, and climatological soil moisture initialization.

### 2.3 | Evaluation metrics and ensemble approaches

We assessed the ability of the six BAM-1.2 configurations to predict precipitation, T2M and the MJO. For

**TABLE 2** Summary of the six BAM-1.2 configurations investigated for sub-seasonal predictions

Configurations	Vertical levels		Convection parametrization		Boundary-layer parametrization	Modified Mellor-Yamada	Soil moisture	
	42	64	Revised simplified arakawa-schubert	Modified Grell-Dévényi	Bretherton-Park (moist scheme)	(dry scheme)	Climatology	GLDAS
42ABC	■		■		■		■	
64ABC		■	■		■		■	
42ABG	■		■		■			■
42GBC	■			■	■		■	
64GBC		■		■	■		■	
42AMC	■		■			■	■	

precipitation and T2M, the deterministic assessment consists of computing the Pearson correlation and root-mean-square error (RMSE) between the prediction and observed anomalies. Each metric was calculated for each grid point and for four lead times: days 1–7 (week-1), 8–14 (week-2), 15–21 (week-3) and 22–28 (week-4). The results were evaluated in the form of weekly averages because the model is expected to have a greater ability to predict weekly anomalies than daily values when producing sub-seasonal predictions (Vitar, 2014).

The performance of MJO prediction was evaluated using the Real-time Multivariate MJO indices (RMMs: Wheeler and Hendon, 2004). Reference RMMs were calculated using the meridional wind at 850 and 200 hPa from the ERA-Interim reanalyses and satellite-observed OLR. RMMs for hindcasts were calculated as proposed by Rashid *et al.* (2011). The metrics used for the MJO prediction quality assessment were bivariate correlation and RMSE (Lin *et al.*, 2008) with lead times in days.

In addition to the single-member deterministic prediction assessment, we evaluated the ability of the Multiple Configurations Ensemble (MCEN) mean prediction formed by the six here-investigated BAM-1.2 configurations with each configuration representing one ensemble member. This was compared to an Initial Condition Ensemble (ICEN) produced with an empirical orthogonal function (EOF) perturbation methodology (Mendonça and Bonatti, 2009), using the configuration that showed the best performance among the six evaluated configurations for producing six ensemble members. The EOF-based perturbation methodology is in operation at CPTEC for extended range forecasts up to 15 days. The methodology produces optimally perturbed analyses by applying the EOFs to  $n$  time series formed by the differences between a model run initialized with a control initial condition and  $n$  model runs initialized with randomly perturbed initial

conditions. The initial random perturbations are drawn from a Gaussian distribution with zero mean and standard deviation comparable to the model short length forecast error, e.g.  $3 \text{ m}\cdot\text{s}^{-1}$  for the horizontal wind field components, 0.6 K for the air temperature field, 1 hPa for the surface pressure field and a vertical standard deviation profile for the specific humidity derived from the ECMWF background error covariance matrix (Derber and Bouttier, 1999). The EOF analysis is performed over the Northern and Southern Hemispheres, over the tropical domain and also regionally over southern and northern South America. The EOF perturbations are the ones associated with the fast growth coefficients. To be used as optimal perturbations, these fast growth modes are rescaled in order to have a standard deviation of the same order of magnitude as the initial perturbations. Finally, the optimal perturbation is added and subtracted to/from the control analysis and an ensemble of  $2n$  initial perturbed states is produced. A more detailed revision of the EOF-based perturbation methodology used at CPTEC can be found in Cunningham *et al.* (2015).

The above-mentioned deterministic metrics for precipitation and T2M anomaly hindcasts, as well as for the hindcast MJO indices, were calculated for the ensemble mean of the two equal-size (six members) ensembles (MCEN and ICEN) to assess and compare the value of utilizing multiple sub-seasonal predictions using two approaches.

In order to have an assessment of the differences in the obtained scores for the investigated model configurations and the two ensemble mean approaches, 95% confidence intervals were computed for the mean correlation and RMSE (for precipitation and T2M anomalies), globally averaged between  $60^\circ\text{N}$  and  $60^\circ\text{S}$ , and for the bivariate correlation and RMSE (for the MJO) using a bootstrap resampling procedure with replacement with 1,000 samples.

### 3 | HINDCAST QUALITY ASSESSMENT

#### 3.1 | Deterministic evaluation of the six investigated BAM-1.2 configurations

Figure 1 shows the correlation between predicted and observed (GPCP) precipitation anomalies for the six BAM-1.2 configurations (first six rows) and four lead times (four columns representing week-1, week-2, week-3 and week-4). The 10 hindcasts per extended austral summer (5 months times two start dates) over 12 austral summers produce a sample with a total of 120 hindcasts. Applying a two-sided Student's *t*-test, the correlation value of 0.2 is statistically significant, different from zero at the 5% level.

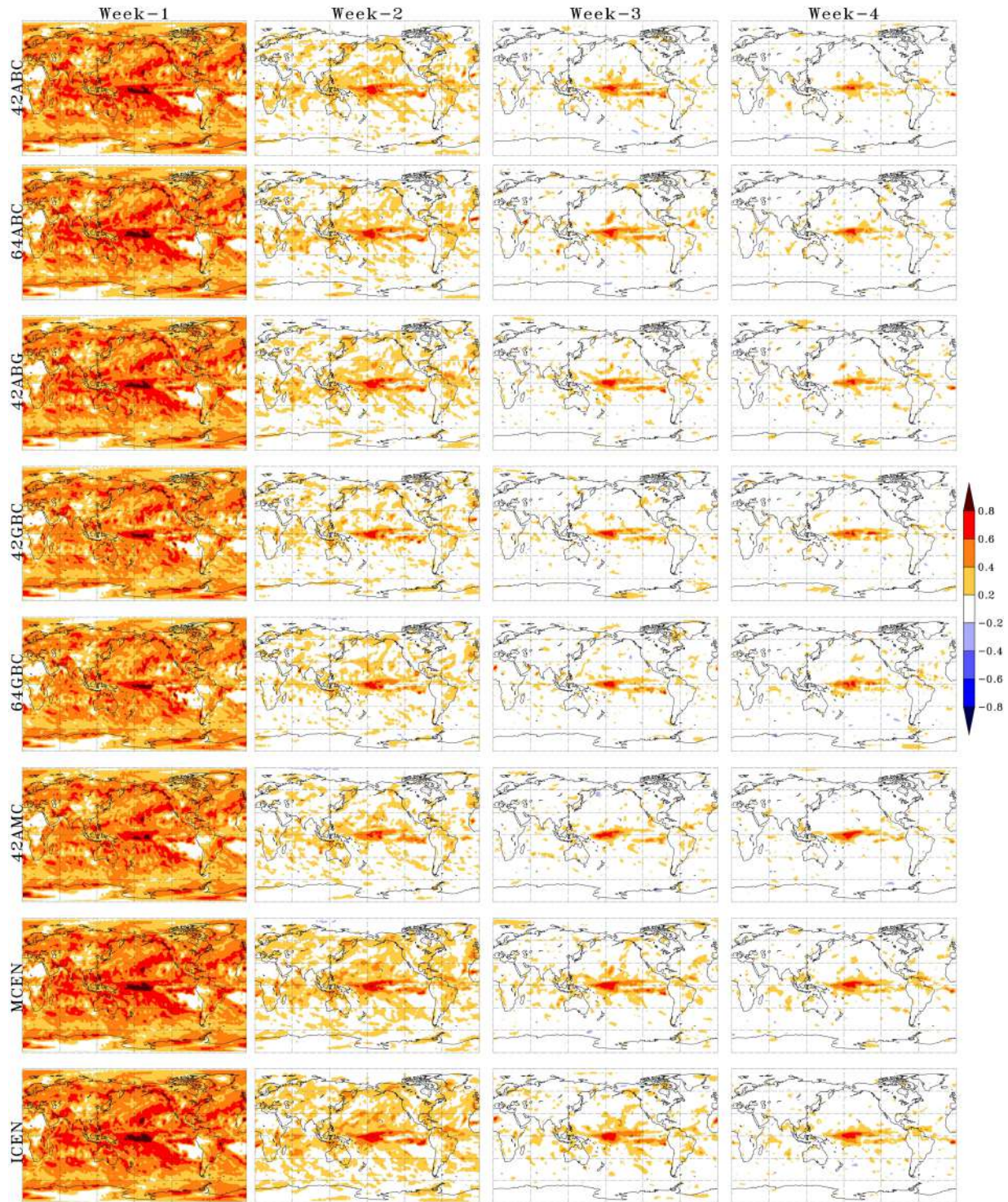
For the six configurations, in general, the correlation is high during the first week in most regions and drops rapidly as lead time increases. This fall is more pronounced between the first and second week for all six configurations as the forecasts extend beyond the deterministic limit of predictability for many scales and we are considering only a single member for this initial analysis. It is noted that all configurations show greater correlation over the Northern Hemisphere than the Southern Hemisphere during week-1 and week-2. This is because GCMs are more likely to predict winter baroclinic weather systems and associated fronts (Zhu *et al.*, 2014). As of the third week, BAM-1.2 correlation values are smaller than 0.2 in practically all extratropical regions. This illustrates that the predictive ability of BAM-1.2 over midlatitudes beyond 15 days is limited for single-member hindcasts. For weeks 3 and 4, significant correlation values are only seen in the tropical Pacific Ocean, over a few areas in northern South America and in the equatorial Atlantic Ocean. The high correlation values in the first two lead times, especially at week-1, are associated to the predictability provided by the initial conditions, and the high correlation values observed in the last two lead times over the equatorial Pacific Ocean are mainly associated to the predictability provided by the El Niño/Southern Oscillation (ENSO) and the MJO (Li and Robertson, 2015; de Andrade *et al.*, 2019). All six configurations show negligible correlation values near the subtropical highs and desert regions from week-1. These characteristics are also noticed in other GCMs configured for sub-seasonal predictions (Zhu *et al.*, 2014; Li and Robertson, 2015; Wheeler *et al.*, 2017; de Andrade *et al.*, 2019) and are associated with the low capacity of GCMs to simulate small precipitation rates in these regions.

The spatial correlation pattern is similar for the four weeks of each of the six BAM-1.2 configurations. However, this pattern for the configurations with revised simplified Arakawa–Schubert (deep convection) and moist

Bretherton–Park (boundary layer) parametrizations shows slightly larger values in the first two weeks than for the other configurations (Figure 1, first two columns of configurations 42ABC, 64ABC and 42ABG vs. first two columns of configurations 42GBC, 64GBC and 42AMC). In week-3 and week-4 (last two columns of Figure 1), correlation levels are similar in terms of both spatial pattern and intensity for the six configurations. Increasing the vertical resolution shows very little change in the precipitation correlation levels at any lead time. For example, the 42 vertical level configuration, 42ABC (first row of Figure 1), and the 64 vertical level configuration, 64ABC (second row of Figure 1), have nearly identical correlation values for most regions. The same is noticed for configurations 42GBC (fourth row of Figure 1) and 64GBC (fifth row of Figure 1). Initialization of the soil moisture also shows no increase in the correlation values for all four investigated weeks. Hindcasts initialised with climatological soil moisture (42ABC, first row of Figure 1) have the same correlation levels as 42ABG (third row of Figure 1) hindcasts, which were initialized with GLDAS soil moisture.

Figure 2 shows the precipitation anomaly RMSE spatial features for the six BAM-1.2 configurations. Highest RMSE values are found over the Intertropical Convergence Zone (ITCZ), Indian Ocean, Maritime Continent, South Pacific Convergence Zone (SPCZ) and South American Convergence Zone (SACZ), which are regions of strong sub-seasonal variability (Liu *et al.*, 2014). The errors grow as lead time increases. As for the correlation assessment, the errors grow more between week-1 and week-2 than from weeks 2 to 3. Again, configurations with revised simplified Arakawa–Schubert and moist Bretherton–Park parametrizations have the fewest errors and do not differ greatly from each other (rows 42ABC, 64ABC and 42ABG in Figure 2). Configurations with modified Grell–Dévényi parametrizations (rows 42GBC and 64GBC in Figure 2) also do not differ much from each other and have larger errors when compared to configurations with revised simplified Arakawa–Schubert parametrizations. That is, the revised simplified Arakawa–Schubert parametrization seems to be better than the modified Grell–Dévényi parametrization and increase of the vertical resolution and soil moisture initialization do not reduce the errors of the hindcasts in any lead time.

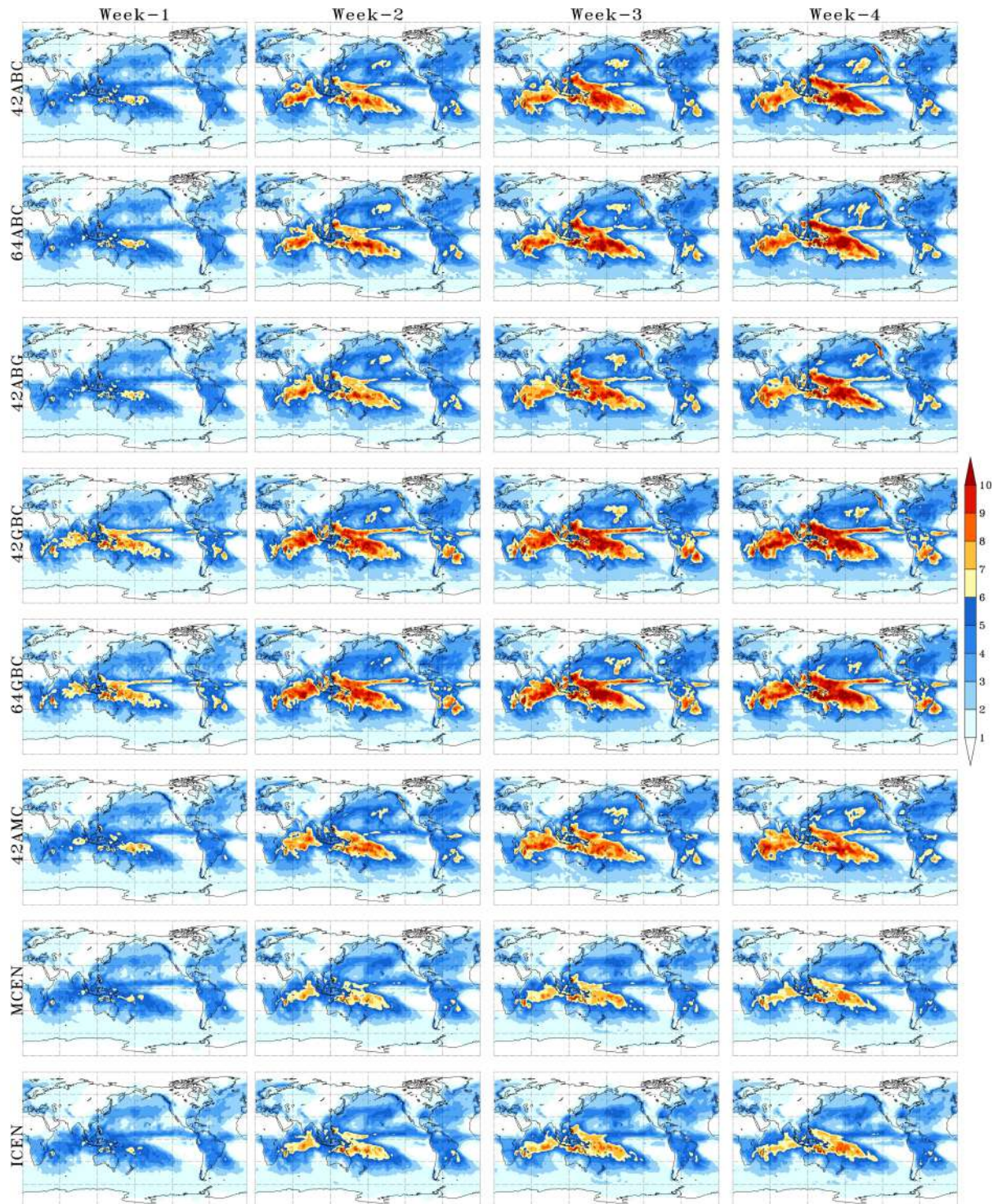
To better note the differences between the six configurations, the mean global correlation between 60°N and 60°S was calculated as a function of lead time (Figure 3a). Vertical bars represent bootstrap 95% confidence intervals. The six configurations show a near-exponential drop in correlation as a function of lead time. The configurations with the revised simplified Arakawa–Schubert and moist Bretherton–Park parametrizations (black, orange and blue lines) have the largest correlation values when compared



**FIGURE 1** Correlation between the predicted and observed (GPCP) precipitation anomalies for the six BAM-1.2 configurations (42ABC, 64ABC, 42ABG, 42GBC, 64GBC and 42AMC) and two ensemble means (MCEN and ICEN) (rows) for week-1, week-2, week-3 and week-4 (columns). The hindcasts were initialized within the extended austral summer period (November–March, 1999/2000–2010/2011) on the dates shown in Table 1 [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

to the other configurations in week-1 and week-2. Important improvements are noticed when comparing the configurations with revised simplified Arakawa–Schubert and modified Grell–Dévényi deep convection parametrizations

at the first two lead times. For example, 42ABC (black line) has a global mean correlation equals to 0.45 at week-1 and drops to 0.18 at week-2, whereas 42GBC (yellow line) has a global mean correlation values equals to 0.40 at week-1

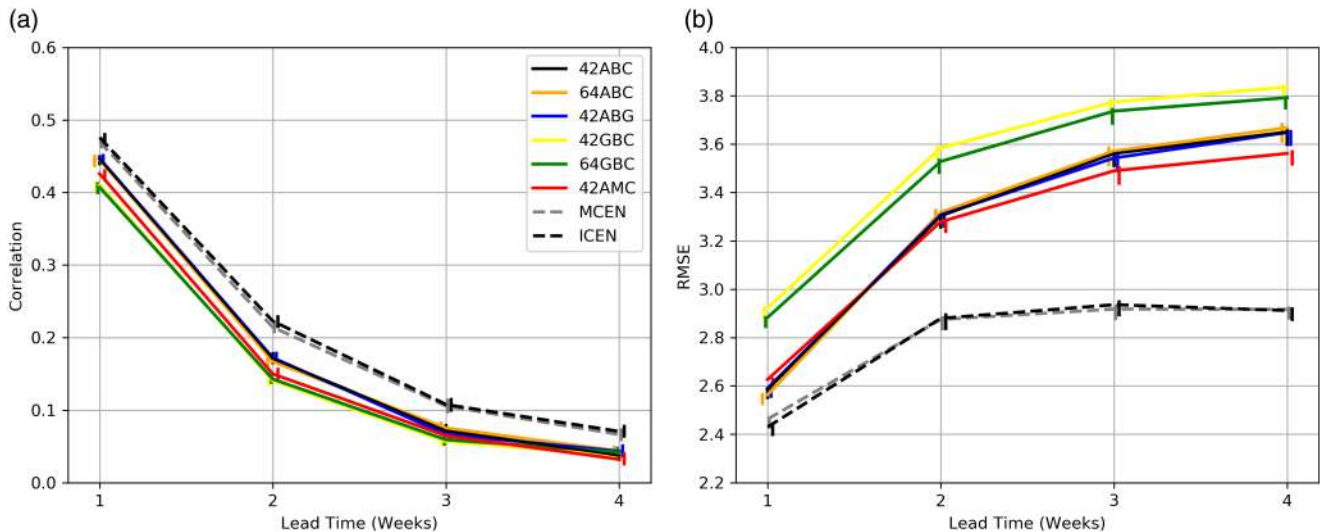


**FIGURE 2** Six configurations (42ABC, 64ABC, 42ABG, 42GBC, 64GBC and 42AMC) and two ensemble means (MCEN and ICEN) (rows) RMSE precipitation anomaly (units are  $\text{mm}\cdot\text{day}^{-1}$ ) for week-1, week-2, week-3 and week-4 (columns). The hindcasts were initialized within the extended austral summer period (November– March, 1999/2000–2010/2011) on the dates shown in Table 1 [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

and drops to 0.14 at week-2. The 95% confidence intervals for the 42ABC (black vertical bars on top of solid black line) do not overlap the 95% confidence intervals for the 42GBC (yellow vertical bars on top of solid yellow line),

illustrating the superiority of 42ABC over 42GBC. However, the six configurations show similar correlation levels at week-4. As noted in the previous figures, the increase of vertical resolution does not result in an increase in





**FIGURE 3** Global mean (a) correlation between predicted and observed precipitation anomalies and (b) RMSE, for six BAM-1.2 configurations (42ABC, 64ABC, 42ABG, 42GBC, 64GBC and 42AMC) and two ensemble approaches (MCEN and ICEN) assessed against GPCP averaged over the latitudinal band 60°N–60°S for four lead times (weeks 1 to 4). The hindcasts were initialized within the extended austral summer period (November–March, 1999/2000–2010/2011) on the dates shown in Table 1. The vertical bars plotted around the four lead times represent 95% confidence intervals produced using a bootstrap resampling procedure with replacement with 1,000 samples. These vertical bars are slightly displaced from the exact lead time location in the horizontal axis to facilitate visualization [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

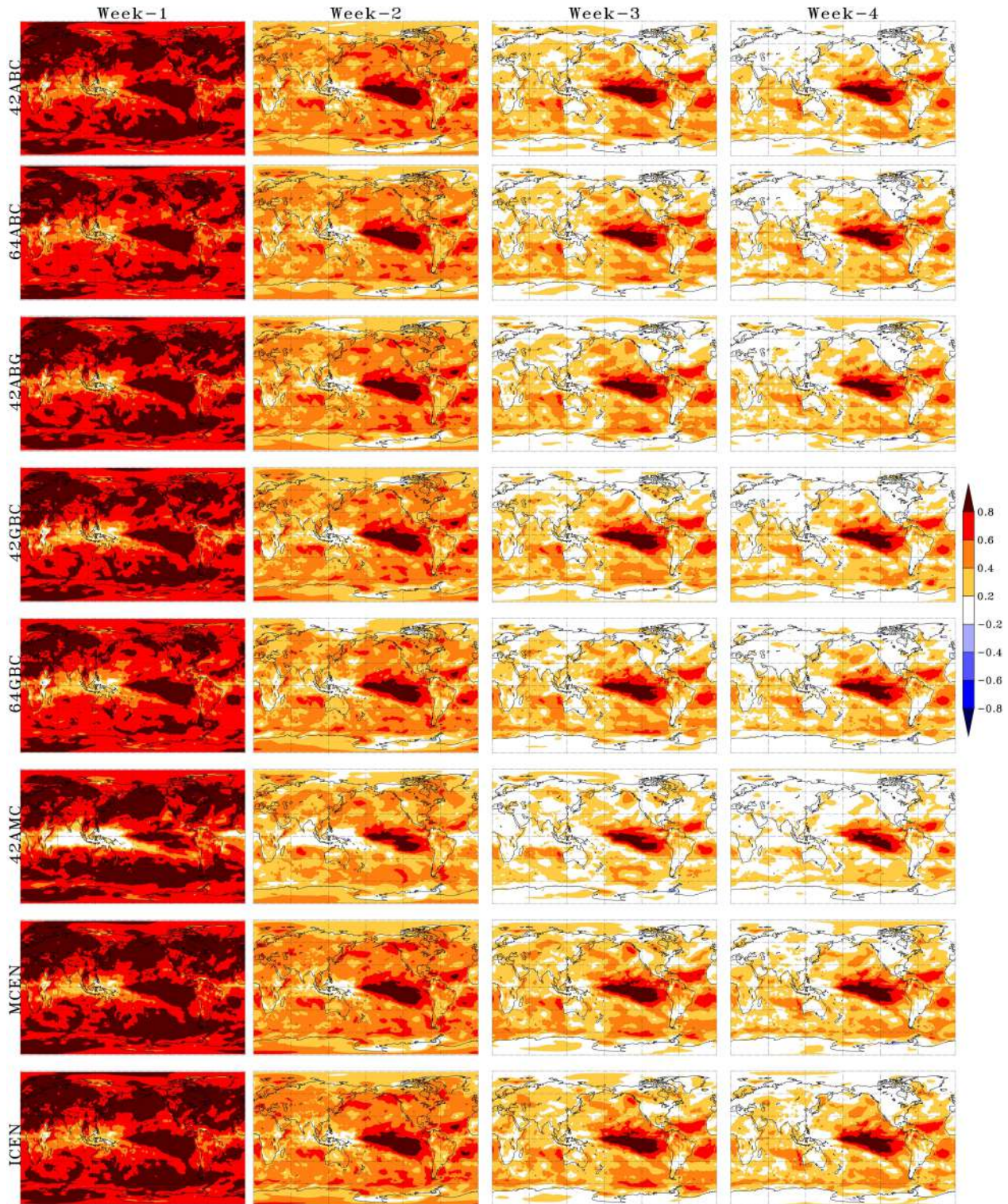
the correlation values. This feature is noticed when we compare the 42ABC (black line) and 64ABC (orange line) configurations or the 42GBC (yellow line) and 64GBC (green line) configurations, which show practically the same behaviour, with the differences between configurations smaller than the 95% confidence intervals (vertical bars). This is also noticed for soil moisture initialization, where the 42ABG configuration (blue line) shows similar correlation levels to the 42ABC configuration (black line).

The global RMSE mean between 60°N and 60°S (Figure 3b) further emphasizes the differences between the revised simplified Arakawa–Schubert and modified Grell–Dévényi deep convection parametrizations. The modified Grell–Dévényi parametrization (green and yellow lines) produces larger errors than the revised simplified Arakawa–Schubert parametrization (other lines) at all lead times. The differences between the errors of these two parametrizations are much larger than the 95% confidence intervals (vertical bars) shown in Figure 3b, illustrating the superiority of the revised simplified Arakawa–Schubert over the modified Grell–Dévényi parametrization. The configurations 42ABC (black line) and 64ABC (orange line) show very similar values at the four lead times, with overlapping 95% confidence intervals. This is also noticed with the 42GBC (yellow line) and 64GBC (green line) configurations. These results suggest that increasing the vertical resolution does not decrease the RMSE. The initialization of soil moisture also does not contribute to the reduction of error (black vs. blue lines). An

interesting aspect is that the configuration with dry modified Mellor–Yamada boundary-layer parametrization (red line) has the smallest error in the last two lead times.

Figure 4 shows the correlation between predicted and reanalyses (ERA-Interim) T2M anomalies for the six configurations and four lead times. The six BAM-1.2 configurations show better prediction performance for T2M anomalies than precipitation anomalies (see Figures 1 and 4). The correlation values decrease with lead time. The highest correlation values are seen over cloud-free oceanic regions (e.g. 42ABC row in Figure 4). However, significant sub-seasonal correlation values exist over a large portion of the global land domain. Over extratropical continental regions, strong correlation values are observed in restricted regions at weeks 3 and 4, for example, over the southeast of the United States and some regions of Asia. Over tropical regions, correlation values are low in regions with high convective activity (e.g. over the Maritime Continent).

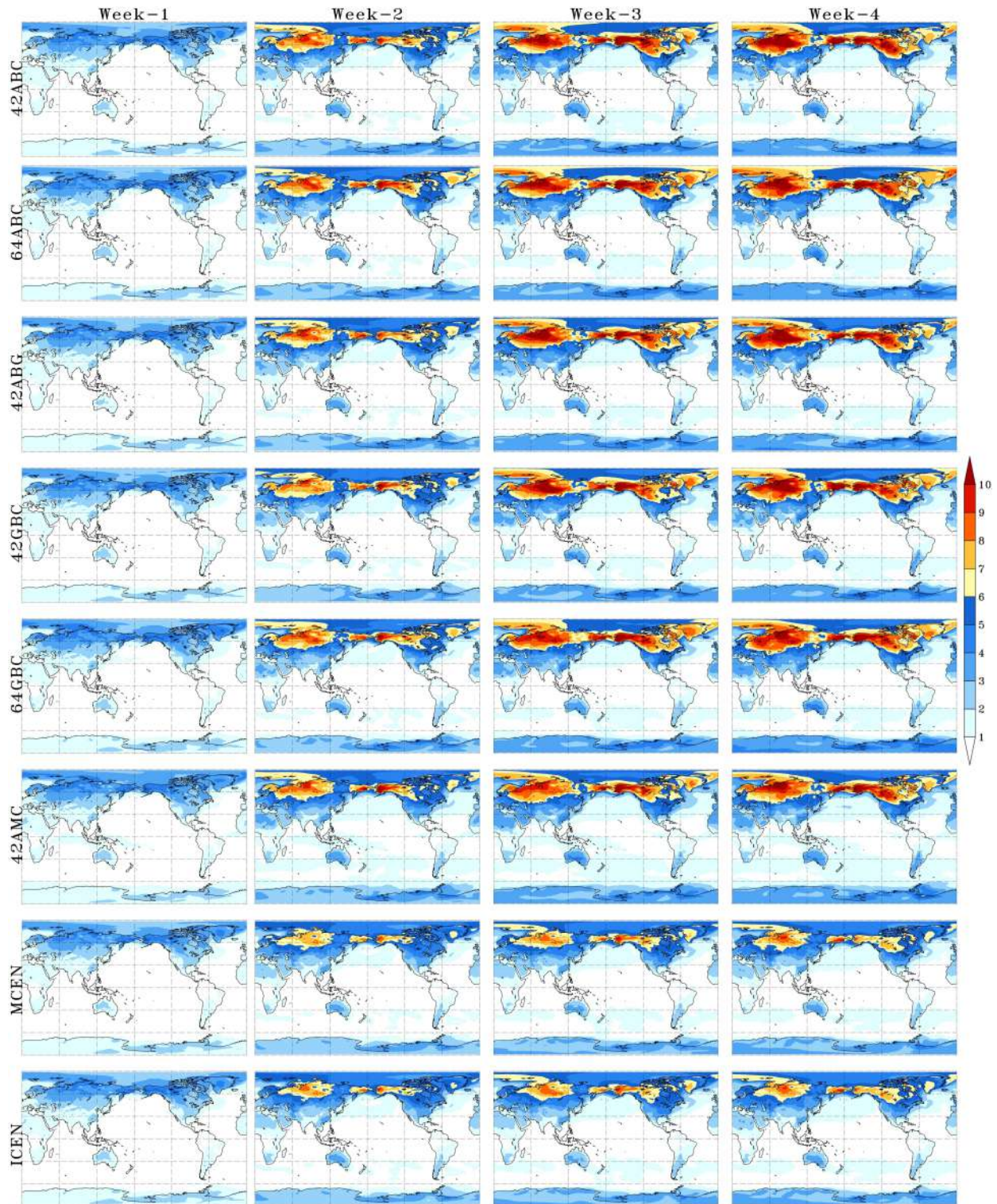
The spatial correlation pattern is similar for the four weeks of the six BAM-1.2 configurations. The difference in performance between configurations with revised simplified Arakawa–Schubert and modified Grell–Dévényi deep convection parametrizations is not observed for T2M (e.g. row 42ABC vs. row 42GBC in Figure 4). Configurations with these two parametrizations have the same performance level for the four weeks of lead time. There are differences when comparing the two boundary layer parametrizations. The configuration with the dry modified Mellor–Yamada parametrization (row 42AMC



**FIGURE 4** Same as Figure 1, except for T2M anomaly [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

in Figure 4) shows reduced performance than the other five configurations at all lead times, which were configured with the moist Bretherton–Park boundary-layer parametrization. Increasing vertical resolution from 42 (rows 42ABC and 42GBC in Figure 4) to 64 (rows 64ABC

and 64GBC in Figure 4) levels seems to slightly reduce prediction performance of extratropical T2M anomalies in the first two weeks of lead times. Predictions with the initialisation of soil moisture (row 42ABG in Figure 4), rather than the climatology soil moisture (row 42ABC in

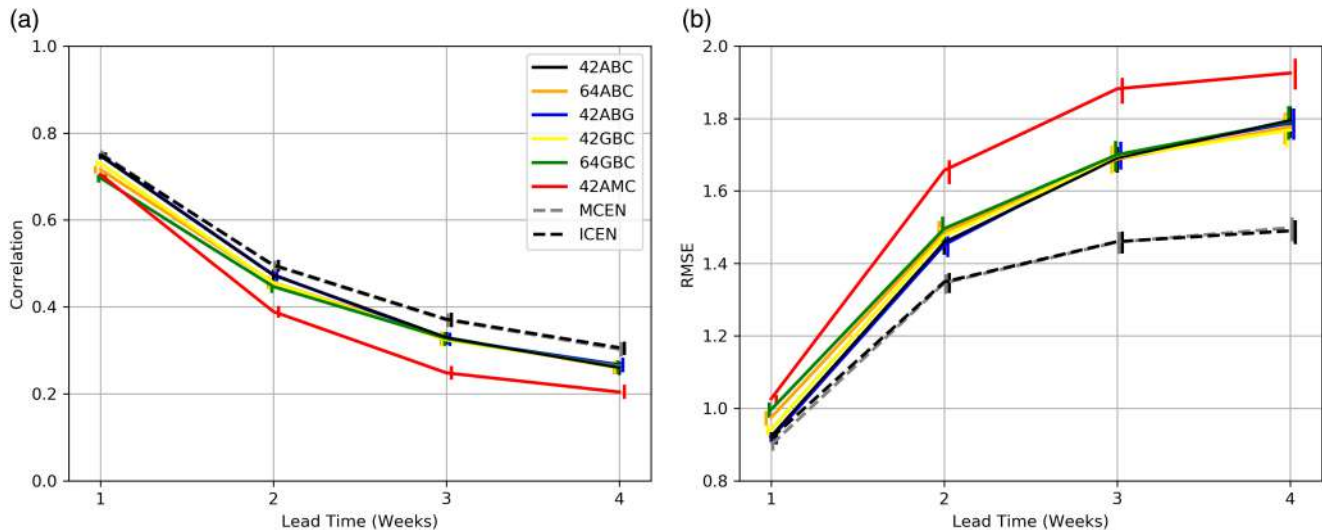


**FIGURE 5** Same as Figure 2, except for T2M anomaly (units are °C) [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

Figure 4), show a slight improvement in correlation in the continental regions (Australia, South and North Americas and Africa) at weeks 2 and 3 lead time.

Figure 5 shows the T2M anomaly RMSE spatial features for the six BAM-1.2 configurations. In all configurations, RMSE values grow with the lead time

and are generally lower over oceanic regions than over continental regions for all four lead times. The highest RMSE values are noticed over Northern Hemisphere regions where there are interactions between midlatitude baroclinic systems, and tropical convective anomalies, which are usually associated with the MJO and circulation



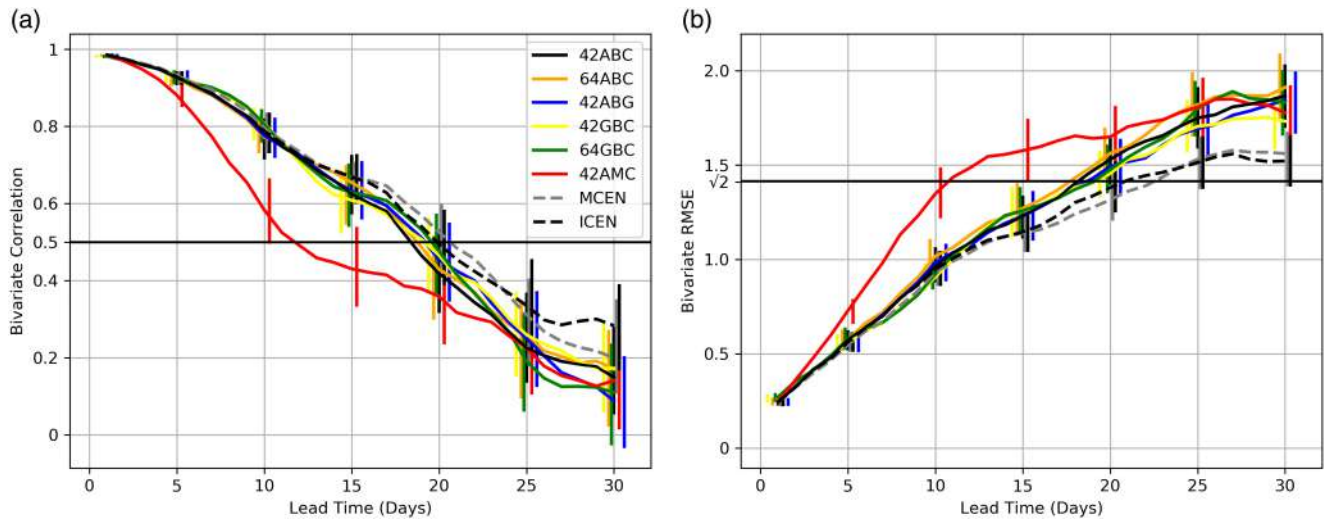
**FIGURE 6** Same as Figure 3, except for T2M anomaly [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

teleconnections through Rossby waves (Stan *et al.*, 2017; Hu *et al.*, 2019). Over northern Asia, high RMSE values are also noticed. The RMSE values are lower over the Southern Hemisphere because there are fewer continental regions than over the Northern Hemisphere, and baroclinic instability is lower at this time of the year in the Southern Hemisphere. The latter makes the interaction between the convective anomalies over tropical regions and circulation over midlatitude regions less pronounced. As a result, the sub-seasonal variability over the Southern Hemisphere extratropical regions is also reduced during the austral summer.

Configurations with the revised simplified Arakawa–Schubert and modified Grell–Dévényi deep convection parametrizations present similar RMSE patterns (Figure 5). Some differences are found in specific regions. For example, the 42GBC configuration shows slightly lower RMSE values over southern South America than the 42ABC configuration at week-3. The opposite is noticed over the Iberian Peninsula. Similar features are noticed for the increase in vertical resolution. Concerning the initialization of soil moisture, subtle differences are noticed between 42ABC and 42ABG, with slight improvements over continental regions such as Australia, South America, southern Africa and North America with initialized soil moisture (42ABG). Large differences are found when comparing configurations with moist Bretherton–Park and dry modified Mellor–Yamada boundary-layer parametrizations. For example, the RMSE values are lower in high-latitude regions over North America and Asia for the 42AMC configuration when compared to the 42ABC configuration. The opposite is over tropical and medium latitude regions.

Figure 6 shows the global mean T2M anomaly correlation (Figure 6a) and RMSE (Figure 6b) averaged between 60°N and 60°S as a function of lead time with 95% confidence intervals (vertical bars). The six configurations show a similar drop (rise) in correlation (RMSE) as a function of lead time. The increase of vertical resolution from 42 to 64 levels, change of deep convection scheme and soil moisture initialization do not influence the levels of correlation and error values for T2M anomaly predictions for the global perspective. This feature is noticed by the proximity or overlap of correlation and RMSE lines of most investigated configurations shown in Figure 6, with overlapping 95% confidence intervals. Differences in performance levels are noticed when comparing moist Bretherton–Park and dry modified Mellor–Yamada boundary-layer parametrizations. The dry modified Mellor–Yamada parametrization (red line) produces smaller correlation values and larger errors than the other five configurations at all lead times, with the differences between these configurations and the others larger than the 95% confidence intervals (vertical bars) illustrating the superiority of the other configurations.

Figure 7a,b show MJO bivariate correlation and bivariate RMSE of all six BAM-1.2 configurations, respectively. Vertical bars represent bootstrap 95% confidence intervals. The MJO predictive ability is determined when the bivariate correlation is lower than 0.5 and the bivariate RMSE grows to  $\sqrt{2}$ . The lead times for these two thresholds to be reached are usually found to be close (Rashid *et al.*, 2011). The bivariate correlation decreases with the increase in lead time and crosses the threshold of 0.5 in 18–19 days for all configurations, except for the 42AMC configuration (red line), which uses the dry modified Mellor–Yamada



**FIGURE 7** (a) Bivariate correlation and (b) bivariate RMSE, for six BAM-1.2 configurations (42ABC, 64ABC, 42ABG, 42GBC, 64GBC and 42AMC) and the two ensemble approaches (MCEN and ICEN) as a function of forecast lead time (in days). The hindcasts were initialized within the extended austral summer period (November– March, 1999/2000–2010/2011) on the dates shown in Table 1. The vertical bars around lead times 1 to 30 days plotted every 5 days represent 95% confidence intervals produced using a bootstrap resampling procedure with replacement with 1,000 samples. These vertical bars are slightly displaced from the exact lead time location in the horizontal axis to facilitate visualization. Note that two black vertical bars are plotted every 5 days, with the first of these bars corresponding to 42ABC and the second corresponding to ICEN [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

boundary-layer parametrization and has a much reduced performance, with the bivariate correlation reaching the 0.5 threshold in 12 days. The bivariate RMSEs (Figure 7b) increase with lead time and each configuration reaches the bivariate RMSE value of  $\sqrt{2}$  at approximately the same lead time as the bivariate correlation. The 42AMC configuration crosses the threshold value of  $\sqrt{2}$  in 11 days, whereas the other five configurations cross the threshold value of  $\sqrt{2}$  in around 18 to 19 days. The overlap of the 95% confidence intervals (vertical bars) for most configurations (except 42AMC) illustrates their similarity in MJO predictive ability.

### 3.2 | Deterministic assessment of two investigated ensemble approaches

With the precipitation anomalies, T2M anomalies and MJO hindcast evaluation of the six configurations shown in the previous section, a preferred BAM-1.2 configuration was determined for defining an ensemble sub-seasonal forecasting system for CPTEC. The increase of the vertical resolution from 42 levels to 64 levels did not result in increase in predictive ability, therefore a vertical resolution of 42 levels was selected. The moist Bretherton–Park boundary-layer parametrization was selected because it contributed to a better performance than the dry modified Mellor–Yamada boundary-layer parametrization, especially for T2M anomalies and MJO predictions.

The revised simplified Arakawa–Schubert and modified Grell–Dévényi deep convection parametrizations showed similar ability for T2M anomalies and MJO prediction with a slight advantage to the modified Grell–Dévényi parametrization for MJO prediction. On the other hand, the revised simplified Arakawa–Schubert parametrization showed a large advantage for sub-seasonal precipitation, with higher correlation and smaller error values than the modified Grell–Dévényi parametrization. Based on this assessment, the revised simplified Arakawa–Schubert deep convection parametrization was chosen. Soil moisture initialization instead of the climatology led to subtle improvements in T2M anomaly predictions in specific regions (e.g. Australia). However, these improvements were lower than expected and given limitations in the availability of accurate real-time soil moisture data, the use of climatological soil moisture was selected for the BAM-1.2 system. Therefore, the chosen BAM-1.2 version for ensemble sub-seasonal forecasting was the 42ABC configuration.

The possible physical reasons why the selected configuration (42ABC) performed better than the other investigated configurations, particularly in terms of the tested boundary-layer and deep convection parametrizations are as follows. The use of the moist Bretherton–Park boundary-layer parametrization resulted in better MJO and T2M predictions performance than the use of the dry modified Mellor–Yamada boundary-layer parametrization. This is because the moist Bretherton–Park has several

advantages compared to the dry modified Mellor–Yamada parametrization. The main contribution of the moist Bretherton–Park parametrization is to improve the representation of the stable night boundary layer, where the predominant physical process in flat areas such as the oceans is the surface radiative cooling. The evolution of the stable nocturnal boundary layer depends on the radiative cooling rate, and therefore the presence of clouds is essential for reducing radiative loss. At sunrise, the state of the stable boundary layer will be important for the evolution of vertical instability and the mixing boundary layer. Therefore, with higher or lower energy released during the evolution of the stable boundary layer, this energy surplus or deficit will contribute to the formation of shallow and deep clouds, and consequently will impact the daytime temperature and precipitation cycle. The energy scales produced by these processes directly or indirectly impact the atmospheric conditions on the sub-seasonal time-scale. As for the comparative performance of sub-seasonal precipitation predictions, important differences were noted when changing the deep convection parametrizations. The revised simplified Arakawa–Schubert parametrization showed better performance than the modified Grell–Dévényi parametrization. This is likely due to the revision made by Han and Pan (2011) in the simplified Arakawa–Schubert parametrization to suppress unrealistic grid-point storms due to remaining instability in the atmospheric column. We next further assess BAM-1.2 sub-seasonal hindcast quality through a deterministic evaluation of ensemble mean predictions. Two ensemble types are evaluated and presented here. The first ensemble consists of one ensemble member from each of the six configurations previously presented, which was denominated Multiple Configurations Ensemble (MCEN). The second ensemble was denominated Initial Condition Ensemble (ICEN) and is composed of six members produced with the chosen 42ABC configuration consisting of a control member and five perturbed members produced with an EOF method (Mendonça and Bonatti, 2009; Cunningham *et al.*, 2015).

The assessment of the ensemble mean of the two ensemble types (MCEN and ICEN) revealed important increase in global mean performance at four lead times for precipitation anomaly predictions when compared to the single-member assessment of the six investigated BAM-1.2 configurations (Figure 3), with the increase in performance larger than the 95% confidence intervals (vertical bars). The two ensemble mean approaches show similar correlation levels (dashed lines in Figure 3a) and overlapping 95% confidence intervals (vertical grey and black bars on top of dashed lines). This shows that BAM-1.2 performance increases when more (six) members are used to form an ensemble with the 42ABC configuration or

when using the six configurations as an ensemble. The predictive ability of GCMs increases as the number of members increases because the ensemble mean acts as a filter for decreasing the uncertainties of the initial conditions used to run the model (Cheung, 2001). This is noticed over several regions in the four investigated lead times (e.g. Figure 1). For example, over extratropical regions at week-2, over eastern South America at week-3 and over northern South America at week-4. Precipitation anomaly hindcasts also show lower RMSE values for both ICEN and MCEN at all four lead times (dashed lines in Figure 3b), with the reduction of error much larger than the 95% confidence intervals (vertical bars). Improvements are noticed primarily over the ITCZ, Indian Ocean, Maritime Continent, SPCZ and SACZ regions (last two rows in Figure 2). This suggests that the ensemble mean helps BAM-1.2 to better represent the sub-seasonal variability in these regions. The same feature is noticed for the T2M hindcasts ensemble means. The two ensemble means show improved T2M anomaly performance when compared to the single-member performance with increased correlation values and decreased error (see last two rows in Figure 4 and 5 and dashed lines in Figures 6). Improvements in MJO forecast performance are also noticed when using the two ensemble approaches. The prediction ability limits are around lead times 18 and 19 days for the single-member configurations, except for the 42AMC which is much reduced (solid lines in Figures 7). For the MCEN this limit increases to 21 days (dashed grey line in Figures 7a,b) and to 20 days (dashed black line in Figures 7a,b) for the ICEN. However, these improvements are less prominent than those identified for precipitation and T2M, because the 95% confidence intervals of the two ensemble approaches largely encompass the 95% confidence intervals of the single members of the individual investigated configurations.

## 4 | CONCLUSIONS

Vertical resolution and physical parametrizations (deep convection and boundary layer) were changed in BAM-1.2 to form five model configurations to determine the model configuration with greater performance for sub-seasonal predictions. These components were selected because these parameters have an important impact on GCMs-simulated MJO (Zhang, 2005; Wang and Chen, 2017; Wang *et al.*, 2018). Given the soil moisture initialization potential to increase the sub-seasonal prediction performance (Koster *et al.*, 2010; 2011; Guo *et al.*, 2012), a further configuration initialized with monthly soil moisture from the previous month, rather than the climatological mean soil moisture, was formed to investigate

the impact of soil moisture initialization on BAM-1.2 predictive ability. The configuration with the best result was selected to form an initial condition ensemble (ICEN) with six members, one control member and five perturbed members produced using an EOF-based methodology. The six configurations, individually evaluated in the first part of this work, were also used to form another ensemble (multiple configurations ensemble, MCEN) to compare the improvements inherent in the ensemble mean between ICEN and MCEN.

All six BAM-1.2 configurations produced high precipitation and T2M anomaly correlation levels for the first week and decreased correlation levels for weeks 2–4. For weeks 3–4, moderate precipitation anomaly correlation levels were restricted to the equatorial Pacific Ocean region. This feature was also noticed in other models (e.g. Li and Robertson, 2015; de Andrade *et al.*, 2019). Precipitation anomaly RMSE increased with lead time and the highest RMSE values were found over regions with strong sub-seasonal variability, for example, over the ITCZ, Indian Ocean, Maritime Continent, SPCZ and SACZ (Liu *et al.*, 2014). For T2M, this characteristic was identified over the Northern Hemisphere where interaction between anomalous convection and midlatitude circulation anomalies are noticed (Hu *et al.*, 2019). The six BAM-1.2 configurations showed better prediction performance for T2M anomalies than for precipitation anomalies.

The increase of the vertical resolution from 42 to 64 levels did not result in an increase in predictive ability. Comparing 42ABC with 64ABC (revised simplified Arakawa–Schubert deep convection configurations with 42 and 64 levels, respectively) and 42GBC with 64GBC (modified Grell–Dévényi deep convection configurations with 42 and 64 levels, respectively), it was noticed that the correlation and RMSE showed nearly identical levels for all lead times for precipitation anomalies, T2M anomalies and MJO predictions. These results may sound contradictory since other studies have shown that the increase in vertical resolution contributes to improvements in predictive ability in the sub-seasonal time-scale (Zhang, 2005; Vitart, 2014). Other factors might be contributing to this finding, for example, the use of initial conditions with only 37 vertical levels, which had to be interpolated to cover 42 and 64 levels.

BAM-1.2 configurations with revised simplified Arakawa–Schubert deep convection parametrization showed better performance than BAM-1.2 configurations with modified Grell–Dévényi deep convection parametrization for sub-seasonal precipitation prediction, with the largest correlation levels found in the first two weeks and the smallest RMSE in the four lead times. However, these two parametrizations showed practically

the same performance for T2M anomalies and MJO, with the commonly used performance thresholds reached at 18/19 days of lead time. The fact that the BAM-1.2 model presents increased ability for sub-seasonal precipitation forecast with the revised simplified Arakawa–Schubert deep convection parametrization compared to modified Grell–Dévényi and very similar ability for T2M and MJO is intriguing. Han and Pan (2011) provided a revision of the simplified Arakawa–Schubert deep convection in the National Centers for Environmental Prediction's (NCEP's) global forecast system. This revision aimed to suppress unrealistic grid-point storms due to remaining instability in the atmospheric column. This might be a possible reason for the better BAM-1.2 performance in sub-seasonal precipitation anomaly forecasting with revised simplified Arakawa–Schubert deep convection parametrization.

The moist Bretherton–Park boundary layer parametrization produced better performance for precipitation anomalies, T2M anomalies and MJO predictions than the dry modified Mellor–Yamada boundary-layer parametrization. The greatest differences were noticed for MJO predictions, where the bivariate correlation decreased more sharply as a function of lead time with the configuration using the dry modified Mellor–Yamada parametrization (42AMC) than for the other configurations. For this, the 42AMC has correlation below 0.5 and RMSE above  $\sqrt{2}$  around the 11th day of lead time. Large discrepancies were also noticed for T2M anomaly predictions when comparing 42AMC with the other configurations.

We did not find important impacts of soil moisture initialization when compared to climatological initialization on the predictive ability of precipitation anomalies in the four investigated lead times. Slight improvements were seen for T2M anomaly predictions in some continental regions. These improvements were smaller than those found in Koster *et al.* (2010; 2011), Guo *et al.* (2012) and van den Hurk *et al.* (2012), and might be related to differences in the investigated seasons, time window, experiments or/and even to low BAM-1.2 sensitivity to soil moisture initialization.

With the evaluation of the six configurations, it was possible to determine a configuration for use as CPTEC sub-seasonal ensemble system. For this, the determined configuration was the 42ABC. This configuration consists of a model version at TQ126 spatial resolution, 42 vertical sigma levels, revised simplified Arakawa–Schubert deep convection parametrization, moist Bretherton–Park boundary-layer scheme, initialization with climatological soil moisture, CLIRAD-LW, CLIRAD-SW, Morrison microphysics and the IBIS-2.6-CPTEC surface model.

The deterministic evaluation of ensembles (MCEN and ICEN), through the computation of the ensemble

means, presented considerable improvements when compared to the single (control) member evaluation (42ABC). For precipitation and T2M anomaly predictions, this improvement was noticed mainly in extratropical continental regions. For MJO predictions, the ensemble means extended by two days the MJO prediction ability limit (e.g. up to 20 days). It is interesting to note that the MCEN mean, formed from the six configurations, showed very similar level of improvements to ICEN when compared to the control member.

This work focused on determining a global CPTEC model configuration for sub-seasonal prediction through a deterministic assessment using a limited number of ensemble members (six). The results presented in this article suggest that BAM-1.2 has competitive performance to other S2S models (Vitart *et al.*, 2017; Lim *et al.*, 2018; de Andrade *et al.*, 2019). In a forthcoming article, we plan to perform a probabilistic assessment of the defined configuration with an increased number of ensemble members and a more detailed comparison of BAM-1.2 with other S2S models. It is worth mentioning that although the extended austral summer is a fundamental season for Brazil (particularly in terms of precipitation), further work is needed in order to evaluate the performance of the Brazilian model during other seasons for identifying regions where best to trust the model for issuing operational sub-seasonal predictions.

## ACKNOWLEDGEMENTS

We thank two anonymous reviewers for providing valuable comments and suggestions that contributed to improving the quality of this manuscript. The first author was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and University of Reading (ref. GS18-179). CASC thanks CNPq, process 304586/2016-1, and Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), process 2015/50687-8 (CLIMAX Project) for the support received. SJW was supported by the National Centre for Atmospheric Science ODA national capability programme ACREW (NE/R000034/1), which is supported by NERC and the GCRF. This research was partially supported by the Climate Science for Services Partnership Brazil project (CSSP-Brazil) funded by the Newton Fund. DCS was supported by CNPq (process 167804/2018-9).

## ORCID

Bruno S. Guimarães  <https://orcid.org/0000-0003-3855-5362>

Paulo Y. Kubota  <https://orcid.org/0000-0003-4858-1337>

## REFERENCES

- Arakawa, A. and Schubert, W.H. (1974) Interaction of a cumulus cloud ensemble with the large-scale environment, Part I. *Journal of the Atmospheric Sciences*, 31, 674–701.
- Boyle, J.S., Klein, S.A., Lucas, D.D., Ma, H.Y., Tannahill, J. and Xie, S. (2015) The parametric sensitivity of CAM5's MJO. *Journal of Geophysical Research: Atmospheres*, 120, 1424–1444.
- Bretherton, C.S. and Park, S. (2009) A new moist turbulence parameterization in the Community Atmosphere Model. *Journal of Climate*, 22, 3422–3448.
- Chen, M., Wang, W. and Kumar, A. (2010) Prediction of monthly-mean temperature: the roles of atmospheric and land initial conditions and sea surface temperature. *Journal of Climate*, 23, 717–725.
- Cheung, K.K.W. (2001) A review of ensemble forecasting techniques with a focus on tropical cyclone forecasting. *Meteorological Applications*, 8, 315–332.
- Chou, M.D. and Suarez, M.J. (1999) *A solar radiation parameterization (CLIRAD-SW) for atmospheric studies*. Suarez, M.J. (Ed.). Series on Global Modeling and Data Assimilation. Report number: NASA/TM-1999-104606/VOL15, pp 40.
- Chou, M.-D., Suarez, M.J., Liang, X.-Z., Yan, M.M.-H. and Cote, C. (2001) *A thermal infrared radiation parameterization for atmospheric studies*. Report number: NASA/TM-2001-104606/VOL19.
- Cunningham, C., Bonatti, J.P. and Ferreira, M. (2015) Assessing improved CPTEC probabilistic forecasts on medium-range timescale. *Meteorological Applications*, 22, 378–384.
- de Andrade, F.M., Coelho, C.A. and Cavalcanti, I.F. (2019) Global precipitation hindcast quality assessment of the Subseasonal to Seasonal (S2S) prediction project models. *Climate Dynamics*, 52, 5451–5475.
- Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, D.P., Bechtold, P., ACM, B., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E.V., Isaksen, I., Kållberg, P., Köhler, M., Matricardi, M., AP, M.N., Monge-Sanz, B.M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N. and Vitart, F. (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137, 553–597.
- Derber, J. and Bouttier, F. (1999) A reformulation of the background error covariance in the ECMWF global data assimilation system. *Tellus A*, 51, 195–221.
- Figueroa, S.N., Bonatti, J.P., Kubota, P.Y., Grell, G.A., Morrison, H., Barros, S.R., Fernandez, J.P., Ramirez, E., Siqueira, L., Luzia, G. and Silva, J. (2016) The Brazilian global atmospheric model (BAM): performance for tropical rainfall forecasting and sensitivity to convective scheme and horizontal resolution. *Weather and Forecasting*, 31, 1547–1572.
- Green, B.W., Sun, S., Bleck, R., Benjamin, S.G. and Grell, G.A. (2017) Evaluation of MJO predictive skill in multiphysics and multimodel global ensembles. *Monthly Weather Review*, 145, 2555–2574.
- Grell, G.A. (1993) Prognostic evaluation of assumptions used by cumulus parameterizations. *Monthly Weather Review*, 121, 764–787.



- Grell, G.A. and Dévényi, D. (2002) A generalized approach to parameterizing convection combining ensemble and data assimilation techniques. *Geophysical Research Letters*, 29(14), 38–1–38–4.
- Grell, G.A. and Freitas, S.R. (2014) A scale and aerosol aware stochastic convective parameterization for weather and air quality modeling. *Atmospheric Chemistry and Physics*, 14, 5233–5250.
- Guo, Z., Dirmeyer, P.A., DelSole, T. and Koster, R.D. (2012) Rebound in atmospheric predictability and the role of the land surface. *Journal of Climate*, 25, 4744–4749.
- Han, J. and Pan, H.L. (2011) Revision of convection and vertical diffusion schemes in the NCEP global forecast system. *Weather and Forecasting*, 26, 520–533.
- Hu, W., Liu, P., Zhang, Q. and He, B. (2019) Dominant patterns of winter-time intraseasonal surface air temperature over the CONUS in response to MJO convections. *Climate Dynamics*, 53, 3917–3936.
- Hudson, D., Alves, O., Hendon, H.H. and Marshall, A.G. (2011) Bridging the gap between weather and seasonal forecasting: intraseasonal forecasting for Australia. *Quarterly Journal of the Royal Meteorological Society*, 137, 673–689.
- Huffman, G.J., Adler, R.F., Morrissey, M.M., Bolvin, D.T., Curtis, S., Joyce, R., McGavock, B. and Susskind, J. (2001) Global precipitation at one-degree daily resolution from multisatellite observations. *Journal of Hydrometeorology*, 2, 36–50.
- Koster, R.D., Mahanama, S.P.P., Yamada, T.J., Balsamo, G., Berg, A.A., Boisserie, M., Dirmeyer, P.A., Doblas-Reyes, F.J., Drewitt, G., Gordon, C.T., Guo, Z., Jeong, J.-H., Lawrence, D.M., Lee, W.-S., Li, Z., Luo, L., Malyshev, S., Merryfield, W.J., Seneviratne, S.I., Stanelle, T., van den Hurk, B.J.J.M., Vitart, F. and Wood, E.F. (2010) Contribution of land surface initialization to subseasonal forecast skill: first results from a multi-model experiment. *Geophysical Research Letters*, 37(2), 1–6. <https://doi.org/10.1029/2009GL041677>.
- Koster, R.D., Mahanama, S.P.P., Yamada, T.J., Balsamo, G., Berg, A.A., Boisserie, M., Dirmeyer, P.A., Doblas-Reyes, F.J., Drewitt, G., Gordon, C.T., Guo, Z., Jeong, J.-H., Lee, W.-S., Li, Z., Luo, L., Malyshev, S., Merryfield, W.J., Seneviratne, S.I., Stanelle, T., van den Hurk, B.J.J.M., Vitart, F. and Wood, E.F. (2011) The second phase of the global land–atmosphere coupling experiment: soil moisture contributions to subseasonal forecast skill. *Journal of Hydrometeorology*, 12, 805–822.
- Kubota, P.Y. (2012) *Variability of storage energy in the soil-canopy system and its impact on the definition of precipitation standard in South America* (in Portuguese with abstract in English). PhD thesis, São José dos Campos, Brazil, Instituto Nacional de Pesquisas Espaciais (INPE).
- Kumar, A., Chen, M. and Wang, W. (2011) An analysis of prediction skill of monthly mean climate variability. *Climate Dynamics*, 37, 1119–1131.
- Li, S. and Robertson, A.W. (2015) Evaluation of submonthly precipitation forecast skill from global ensemble prediction systems. *Monthly Weather Review*, 143, 2871–2889.
- Liang, P. and Lin, H. (2018) Sub-seasonal prediction over East Asia during boreal summer using the ECCO monthly forecasting system. *Climate Dynamics*, 50, 1007–1022.
- Liebmann, B. and Smith, C.A. (1996) Description of a complete (interpolated) outgoing longwave radiation dataset. *Bulletin of the American Meteorological Society*, 77, 1275–1277.
- Lim, Y., Son, S.W. and Kim, D. (2018) MJO prediction skill of the subseasonal-to-seasonal prediction models. *Journal of Climate*, 31(10), 4075–4094.
- Lin, H., Brunet, G. and Derome, J. (2008) Forecast skill of the Madden–Julian Oscillation in two Canadian atmospheric models. *Monthly Weather Review*, 136, 4130–4149.
- Lin, H., Gagnon, N., Bearegard, S., Muncaster, R., Markovic, M., Denis, B. and Charron, M. (2016) GEPS-based monthly prediction at the Canadian Meteorological Centre. *Monthly Weather Review*, 144, 4867–4883.
- Liu, X., Wu, T., Yang, S., Li, T., Jie, W., Zhang, L., Wang, Z., Liang, X., Li, Q., Cheng, Y., Ren, H., Fang, Y. and Nie, S. (2017) MJO prediction using the sub-seasonal to seasonal forecast model of Beijing Climate Center. *Climate Dynamics*, 48, 3283–3307.
- Liu, X., Yang, S., Li, Q., Kumar, A., Weaver, S. and Liu, S. (2014) Sub-seasonal forecast skills and biases of global summer monsoons in the NCEP Climate Forecast System version 2. *Climate Dynamics*, 42, 1487–1508.
- Mastrangelo, D., Malguzzi, P., Rendina, C., Drofa, O. and Buzzi, A. (2012) First outcomes from the CNR-ISAC monthly forecasting system. *Advances in Science and Research*, 8, 77–82.
- Mellor, G.L. and Yamada, T. (1982) Development of a turbulence closure model for geophysical fluid problems. *Reviews of Geophysics*, 20, 851–875.
- Mendonça, A.M. and Bonatti, J. (2009) Experiments with EOF-based perturbation methods and their impact on the CPTEC/INPE ensemble prediction system. *Monthly Weather Review*, 137, 1438–1459.
- Morrison, H., Thompson, G. and Tatarskii, V. (2009) Impact of cloud microphysics on the development of trailing stratiform precipitation in a simulated squall line: comparison of one- and two-moment schemes. *Monthly Weather Review*, 137, 991–1007.
- Rashid, H.A., Hendon, H.H., Wheeler, M.C. and Alves, O. (2011) Prediction of the Madden–Julian Oscillation with the POAMA dynamical prediction system. *Climate Dynamics*, 36, 649–661.
- Reichler, T. and Roads, J.O. (2005) Long-range predictability in the Tropics. Part I: Monthly averages. *Journal of Climate*, 18, 619–633.
- Rui, H. and Beaudoin, H. (2017) *Readme document for GLDAS Version 2 data products*. NASA Goddard Space Flight Center, 610, 1–21.
- Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y.T., Chuang, H.Y., Iredell, M. and Ek, M. (2014) The NCEP climate forecast system version 2. *Journal of Climate*, 27, 2185–2208.
- Shelly, A., Xavier, P., Copsey, D., Johns, T., Rodríguez, J.M., Milton, S. and Klingaman, N. (2014) Coupled versus uncoupled hindcast simulations of the Madden–Julian Oscillation in the Year of Tropical Convection. *Geophysical Research Letters*, 41, 5670–5677.
- Stan, C., Straus, D.M., Frederiksen, J.S., Lin, H., Maloney, E.D. and Schumacher, C. (2017) Review of tropical–extratropical teleconnections on intraseasonal time scales. *Reviews of Geophysics*, 55, 902–937.
- Tarasova, T.A. and Fomin, B.A. (2000) Solar radiation absorption due to water vapor: advanced broadband parameterizations. *Journal of Applied Meteorology*, 39, 1947–1951.
- van den Hurk, B., Doblas-Reyes, F., Balsamo, G., Koster, R.D., Seneviratne, S.I. and Camargo, H., Jr. (2012) Soil moisture effects on seasonal temperature and precipitation forecast scores in Europe. *Climate Dynamics*, 38, 349–362.

- Vitart, F. (2004) Monthly forecasting at ECMWF. *Monthly Weather Review*, 132, 2761–2779.
- Vitart, F. (2014) Evolution of ECMWF sub-seasonal forecast skill scores. *Quarterly Journal of the Royal Meteorological Society*, 140, 1889–1899.
- Vitart, F. (2017) Madden–Julian Oscillation prediction and teleconnections in the S2S database. *Quarterly Journal of the Royal Meteorological Society*, 143(706), 2210–2220.
- Vitart, F., Ardilouze, C., Bonet, A., Brookshaw, A., Chen, M., Codorean, C., Déqué, M., Ferranti, L., Fucile, E., Fuentes, M. and Hendon, H. (2017) The subseasonal to seasonal (S2S) prediction project database. *Bulletin of the American Meteorological Society*, 98, 163–173.
- Wang, B. and Chen, G. (2017) A general theoretical framework for understanding essential dynamics of Madden–Julian Oscillation. *Climate Dynamics*, 49, 2309–2328.
- Wang, B., Lee, S.S., Waliser, D.E., Zhang, C., Sobel, A., Maloney, E., Li, T., Jiang, X. and Ha, K.J. (2018) Dynamics-oriented diagnostics for the Madden–Julian Oscillation. *Journal of Climate*, 31, 3117–3135.
- Weber, N.J. and Mass, C.F. (2017) Evaluating CFSv2 subseasonal forecast skill with an emphasis on tropical convection. *Monthly Weather Review*, 145, 3795–3815.
- Wheeler, M.C. and Hendon, H.H. (2004) An all-season real-time multivariate MJO index: development of an index for monitoring and prediction. *Monthly Weather Review*, 132, 1917–1932.
- Wheeler, M.C., Zhu, H., Sobel, A.H., Hudson, D. and Vitart, F. (2017) Seamless precipitation prediction skill comparison between two global models. *Quarterly Journal of the Royal Meteorological Society*, 143(702), 374–383.
- Willmott, C.J., Rowe, C.M. and Mintz, Y. (1985) Climatology of the terrestrial seasonal water cycle. *Journal of Climatology*, 5, 589–606.
- Yu, H., Kaufman, Y.J., Chin, M., Feingold, G., Remer, L.A., Anderson, T.L., Balkanski, Y., Bellouin, N., Boucher, O., Christopher, S. and DeCola, P. (2006) A review of measurement-based assessments of the aerosol direct radiative effect and forcing. *Atmospheric Chemistry and Physics*, 6, 613–666.
- Zhang, C. (2005) Madden–Julian Oscillation. *Reviews of Geophysics*, 43(2), 1–36. <https://doi.org/10.1029/2004RG000158>.
- Zhang, C. (2013) Madden–Julian Oscillation: bridging weather and climate. *Bulletin of the American Meteorological Society*, 94, 1849–1870.
- Zhu, H., Wheeler, M.C., Sobel, A.H. and Hudson, D. (2014) Seamless precipitation prediction skill in the tropics and extratropics from a global model. *Monthly Weather Review*, 142, 1556–1569.

**How to cite this article:** Guimarães BS, Coelho CAS, Woolnough SJ, *et al.* Configuration and hindcast quality assessment of a Brazilian global sub-seasonal prediction system. *QJR Meteorol Soc.* 2020;1–18. <https://doi.org/10.1002/qj.3725>