Sub-seasonal tropical cyclone genesis prediction and MJO in the S2S dataset

Chia-Ying Lee*

International Research Institute for Climate and Society, Columbia University, Palisades, NY

Suzana J. Camargo

Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY

Fréderic Vitart

European Centre for Medium-Range Weather Forecasts, Reading, United Kingdom

Adam H. Sobel

Department of Applied Physics and Applied Mathematics, Columbia University, New York, NY

Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY

Michael K. Tippett

Department of Applied Physics and Applied Mathematics, Columbia University, New York, NY

Department of Meteorology, King Abdulaziz University, Jeddah, Saudi Arabia

Corresponding author address: Chia-Ying Lee, International Research Institute for Climate and Society, Columbia University, 61 Route, 9W, Palisades, NY 10964.

E-mail: clee@iri.columbia.edu

ABSTRACT

Subseasonal probabilistic prediction of tropical cyclone (TC) genesis is investigated here using models from the seasonal to subseasonal (S2S) dataset. Forecasts are produced for basin-wide TC occurrence at weekly temporal resolution. Forecast skill is measured using the Brier skill score relative to two no-skill climatological forecasts: an annual mean climatology that is constant through the year and a seasonal climatology that varies monthly through the TC season. Skill depends on model characteristics, lead time, and ensemble prediction design. Most forecasts show skill for week one (days one to seven), when initialization is important. Among the six S2S models examined here, the European Centre for Medium-Range Weather Forecasts (ECMWF) model has the best performance, with skill in the Atlantic, western North Pacific, eastern North Pacific, and South Pacific at week two. Similarly, the Australian Bureau of Meteorology (BoM) model is skillful in the western North Pacific, South Pacific, and northern Australia at week two. Compared to statistical models, S2S model skill scores are higher at week one, comparable at week two, and lower from weeks three to five. The Madden-Julian Oscillation (MJO) modulates observed TC genesis, and there is a clear relationship, across models and lead times, between models' skill scores and their ability to represent the MJO–TC relation accurately. Skill scores increase with the ensemble size, as found in previous weather and seasonal prediction studies.

1. Introduction

The Madden Julian Oscillation (MJO, Madden and Julian 1972) modulates tropical cyclone (TC) activity globally. The probability of TC genesis is typically greater during or after a strong convective MJO phase than at other times (Camargo et al. 2009; Klotzbach 2014; Klotzbach and Oliver 2015a). In the Atlantic (ATL, Mo 2000; Maloney and Hartmann 2000b; Klotzbach 2010; Klotzbach and Oliver 2015b), the enhanced storm genesis occurs when a strong active MJO is in the Indian Ocean. In contrast, in the eastern North Pacific (ENP), enhanced TC genesis occurs when a strong MJO is active in the central and eastern North Pacific (Molinari et al. 1997; Maloney and Hartmann 2000a, 2001; Aiyyer and Molinari 2008). Similarly, in the Western North Pacific (WNP, Nakazawa 1988; Liebmann et al. 1994; Sobel and Maloney 2000; Kim et al. 2008; Li and Zhou 2013), North (NI, Nakazawa 1988; Liebmann et al. 1994; Kikuchi and Wang 2010; Krishnamohan et al. 2012) and South (SIN, Bessafi and Wheeler 2006; Ho et al. 2006) Indian Ocean, and South Pacific (SCP, Hall et al. 2001), the number of storms increases when the MJO is active in these basins. Additionally, typhoon tracks shift eastward (westward) when the convective MJO is active in the Indian Ocean (western Pacific) (Kim et al. 2008). Rapidly intensifying storms are more frequent in the ATL when the MJO is active in the Indian Ocean (Klotzbach 2012). An empirical TC genesis index (Camargo et al. 2009) shows the occurrence of a systematic enhanced low-level absolute vorticity and increased midlevel relative humidity associated with a strong active MJO, leading to larger local values of this index.

In these observational studies, it is often stated that accurate predictions of the MJO and knowledge of the MJO–TC relationship offer the potential for forecasts of the probability of TC genesis with a few weeks lead time. Regional statistical models for subseasonal TC prediction have in fact been developed (Leroy and Wheeler 2008; Slade and Maloney 2013) using MJO indices, as well

as other environmental parameters. When an MJO index is included as one of the predictors, there is a significant, albeit small, improvement of these forecasts relative to those produced by a a sub-seasonal statistical model with no MJO index, up to three weeks lead time. For longer leads, the forecast skill is thought to be primary from the climatological seasonal cycle and interannual variability. Reforecasts from the European Centre for Medium-Range Weather Forecasts (ECMWF) also suggest that the accuracy of the MJO prediction has a significant impact on the predicted TC frequency (Vitart 2009). Compared to a southern hemisphere TC statistical forecasts (Leroy and Wheeler 2008), the ECMWF model has greater skill in predicting TC occurrence at week one, while the statistical model performs better for longer leads (Vitart et al. 2010). Furthermore, the ECMWF skill in predicting TC activity is sensitive to the MJO phase and amplitude at the time of the model initialization (Belanger et al. 2010).

With the increasing demand for forecasts on the time scale between weather and seasonal–interannual — the so-called "subseasonal" time scale — an international effort was initiated to improve and develop various aspects of dynamical subseasonal predictions, including TC subseasonal forecasts. A key goal of these efforts is improved understanding of the factors that affect forecast prediction skill. The multi-model seasonal to subseasonal (S2S, Vitart et al. 2017) dataset, containing extensive reforecasts with lead times up to 60 days, is ideal for this task. In this study, we focus on the subseasonal prediction of TC genesis in the S2S reforecasts. While the ability of global models to simulate the MJO-TC modulation (Vitart 2009; Satoh et al. 2012; Kim et al. 2014; Murakami et al. 2015; Xiang et al. 2015) and the prediction skill of TC genesis prediction in subseasonal time-scales have been analyzed (Belanger et al. 2010; Elsberry et al. 2011; Tsai et al. 2013; Elsberry et al. 2014; Nakano et al. 2015; Barnston et al. 2015; Yamaguchi et al. 2015; Li et al. 2016) for various models, this is the first comprehensive multi-model, multi-year analysis of reforecasts of TC genesis prediction on subseasonal time-scales.

Here we use the S2S data to construct reforecasts of the probability of TC occurrence basin-wide with weekly temporal resolution. The prediction skill is evaluated using the Brier skill score. Reforecasts, observations, Brier skill, and other analysis methods used here are described in Section 2. We then discuss TC climatology in the reforecasts in Section 3 to define the tropical storm thresholds and seasonality for prediction skill evaluation. The simulated and observed MJO modulation of TC genesis is examined in Section 4. Then, we analyze the the prediction skill and its connections to forecast lead time, model characteristics, and ensemble prediction system design in Section 5. Results are then summarized in Section 6.

2. Data and Methods

a. S2S reforecasts

Table 1 shows some basic characteristics of the S2S reforecasts used here. They are obtained from coupled, global general circulation models run by six operational centers: the Australian Bureau of Meteorology (BoM), the China Meteorological Administration (CMA), the European Centre for Medium-Range Weather Forecasts (ECMWF), the Japan Meteorological Agency (JMA), the Météo-France/Centre National de Recherche Météorologiques (MetFr), and the National Centers for Environmental Prediction (NCEP). Note that because the design of the ensemble prediction system (specifically, the frequency of forecasts and ensemble size) differs among these agencies, the reforecasts are heterogeneous. We treat such differences in system design as additional factors contributing to prediction skill. Another heterogeneous feature in this dataset is over which years the reforecasts were done. While this might affect the comparison, we do not think it is likely to qualitatively change the relative skill of the forecast systems. Further details of the S2S dataset is described in Vitart et al. (2017).

b. TCs in the S2S models and observations

To track TCs in the S2S models, we use a tracker (Vitart and Stockdale 2001) that detects TCs based on thresholds for the vorticity maximum at 850 hPa, sea surface pressure minimum, a local maximum in the vertically averaged temperature in between 250–500 hPa (warm core), and the thickness maximum between 1000–200 hPa. Additionally, detected storms have to last at least two days to be include in our analysis. The S2S TC tracks contain instantaneous maximum sustained winds and storm locations daily.

Observations of tropical cyclone tracks are from the HURDAT2, produced by the National Hurricane Center (NHC, Landsea and Franklin 2013), and from the Joint Typhoon Warning Center (JTWC, Chu et al. 2002). Both best-track datasets include 1-min maximum sustained wind, minimum sea level pressure, and storm location every 6 hour.

Following the conventional definitions, the TC basins are: Atlantic (ATL), northern Indian Ocean (NI), western north Pacific (WNP), eastern north Pacific (ENP), southern Indian Ocean (SIN, 0-90^{o}E), northern Australia (AUS, 90-160^{o}E), and southern Pacific (SPC, east of 160^{o}E), .

c. MJO definition

The MJO indices, RMM1 and RMM2, are calculated using zonal wind at 200 hPa and 850 hPa, and outgoing long wave radiation (Gottschalck et al. 2010; Wheeler and Hendon 2004; Vitart 2017). Observational RMM indices are calculated using reanalysis data from ERA-Interim reanalysis data.

d. Brier Skill Score

The Brier Skill Score (BSS) is used to assess the relative skill of a probabilistic forecast of TC occurrence in a basin over climatology. The Brier Skill (BS) is defined as:

$$BS = \frac{1}{N}\sum_{i=1}^{N}(p_i - o_i)^2 \tag{1}$$

$$BSS = 1 - \frac{BS}{BS_{ref}}, \tag{2}$$

where N is the total number of forecasts, p_i is the predicting probability for i$^{\text{th}}$ forecast and o_i is the i$^{\text{th}}$ observation. For each ensemble member in p_i and for each o_i, it is 0 for no genesis in the basin and is 1 for more than 1 storm genesis during the forecast period. Thus, the BS is the mean forecast error. The BS_{ref} is similar to the BS, but with forecasts from observed climatology. In this study, two climatologies are used. One is the seasonally varying climatology at monthly time resolution, while the other one is a constant, annual mean climatology. When a model is skillful compared to the climatological predictions, the BSS is larger than 0.

e. Candy–plot analysis

To analyze dependence of TC genesis on MJO phases, the probability density function (PDF) of storm genesis is calculated in each TC basin and binned by MJO phase. To identify the favorable and the unfavorable MJO phases, values of TC number in each week are randomly shuffled in time throughout the entire period to obtain PDFs independent of the MJO. The favorable MJO phases are then defined when the true PDF is larger than the 90$^{\text{th}}$ percentile of the 4000 PDFs obtained from randomly swapping the data, and the unfavorable MJO phases are defined when true PDF is less than 10$^{\text{th}}$ percentile of the randomized PDFs. The PDFs are then organized by the longitude in Y–axis from ATL to ENP and by the MJO phases in X–axis, like a sheet of candies.

The Candy–plot analysis is conducted for observations and the six S2S models. The comparison of global pattern (PDFs from all TC basins) between each S2S model and the observed pattern is then quantified using r^2, i.e., the fraction of the variance of the observed PDFs that is predicted by the model PDFs. In this analysis, we include only storms when the magnitude of the MJO index is larger than one standard deviation. The fractions of storms we used are 60% in observations, and roughly 60%, 50%, 40%, 52%, 65%, and 52% in BoM, CMA, ECMWF, JMA, MetFr, and NCEP reforecasts, respectively.

3. Tropical cyclone climatology

a. Intensity and tropical storm threshold

Since their horizontal resolutions are inadequate to represent TC inner core structure, the global models used here are not able to simulate the highest observed TC intensities. Another factor that impacts the simulated intensities as represented in the S2S archive is that the model outputs are instantaneously archived on a 1.5^{o} by 1.5^{o} grid in the S2S database every 24 hours. As a result, the cumulative density distribution (CDF) of TCs' lifetime maximum intensity (LMI) shows that median LMI for of the observed storms is 50 kt while it is the range of 25–35 kt for the S2S models, except for BoM which has a median LMI of 40 kt (Fig. 1). The BoM model is able to simulate stronger storms than other S2S models that have higher horizontal resolutions, probably due to its physical parameterizations and dynamical cores. Multiple studies with other global climate models have noted that these other factors are important, so that the maximum TC intensities simulated by a model are not a simple function of that model's horizontal resolution (Vitart et al. 2001; Murakami et al. 2012; Zhao et al. 2012; Reed et al. 2015b; Duvel et al. 2017). Although there is a significant low bias in the S2S TC intensities, we can categorize storms using quantile

analysis (Camargo and Barnston 2009). For example, the observed tropical storm wind speed threshold is 34 kt, which in the observed LMI distribution corresponds to the 18th percentile (gray line in Fig. 1). Thus, we define the tropical storm threshold as the 18th percentile of the LMI CDF in each model. These are 34, 23, 24, 24, 27, and 26 kt, respectively, for the BoM, CMA, ECMWF, JMA, MetFr, and NCEP models. In this study, we only consider TCs that reach the tropical storm threshold thus defined in observations and the model reforecasts.

b. Genesis and TC season definition

Genesis time is defined here as the time of the first point recorded on each forecast track. TCs that exist previous to the model initialization time have already undergone genesis. Nevertheless, for purposes of model evaluation, we refer to the first recording time (usually day one, or $t = 24$ hours) of the pre-existing storms as their genesis time. The reason for including pre-existing storms in our analysis will be discussed in Section 5. With the exception of week one — when there is a higher TC occurrence because of the pre-existing storms — the forecast genesis climatology does not change much with lead time. Therefore, while we only show here the genesis climatology for week two forecasts (Fig. 2), our results are also valid at longer lead times.

Globally, the ECMWF model generates 20% more TCs than observed, while CMA, MetFr, and NCEP have genesis rates 140%, 65%, and 80% higher than observed. In contrast, the BoM and JMA models generate 35% and 45% fewer TCs than the observed climatology. Low-resolution models often have unrealistically high TC genesis rates in the subtropics, when storms are detected and tracked using algorithms with model-dependent thresholds as we use here (Camargo 2013). However the difference maps between simulated and observed genesis counts, Fig. 2b–g, suggest that this is not the case for the S2S models, since the errors in the subtropics are much smaller than those in the tropical belt (30^{o}S – 30^{o}N).

Regionally, the strongest observed local maxima of TC genesis rate occur in the ENP and WNP (Fig. 2a). In the three southern hemisphere basins (SIN, AUS, SPC), the observed storms form in an elongated area around 15^{o}S. In general, the S2S models are able to capture the regions of TC genesis (not shown), although they all underestimate the genesis rate in the ENP and tend to under-predict genesis in the ATL, with exception of NCEP in the eastern tropical ATL. Among these models, the ECMWF model has the smallest regional biases, followed by BoM. The JMA model underestimates the rate of TC genesis everywhere, while CMA, MetFr, and NCEP generate too many TCs in the Pacific and Indian Oceans. The CMA model has the largest positive bias in the Pacific Ocean (Fig. 2c), with more than 1 storm per year per grid ($4^{o}\times4^{o}$) between 4 to 12^{o}N, compared to the observed climatological mean of less than 0.2 storms per year per grid (Fig. 2a).

Despite these biases in the total TC counts and genesis spatial distribution, the S2S models represent the annual cycle of TC genesis reasonably well (Fig. 3). We define regionally varying TC seasons that consist of the months with genesis rate higher than 5% of the annual genesis rate in each region. Using this definition, the TC seasons in some models are slightly different than in observations. For example, the observed hurricane season in the ENP is defined as May to October, but it is from July to December in the BoM model. The simulated TC seasons in the ECMWF model have the best match with those from observations.

Although there are differences between simulated and observed TC seasons, our target is to have skillful TC predictions during the observed TC seasons. Therefore, the observed TC seasons are used in our prediction skill evaluation below. Using this definition the TC seasons are: June to November in the ATL, May to October in the ENP, May to December in the WNP, April to June and September to December in the NI, October to April in the SIN, and November to April in the AUS and SPC.

4. MJO–TC modulation

Next, we examine whether the S2S models are capable of simulating the observed MJO–TC modulation. Specifically, we refer to the spatial distribution of genesis as a function of the MJO phase, such as the anomalous fields in Figs. 5 and 4, and the basin-wide probability density functions (PDF) in Fig. 6. Our focus is the week two reforecasts, when all S2S models have skill in predicting the MJO (Vitart 2017).

Global climate models are able to simulate (Zhang 2013; Camargo and Wing 2016) the observed dependence of TC genesis on the MJO (Nakazawa 1988; Liebmann et al. 1994; Mo 2000; Maloney and Hartmann 2000b,a; Kim et al. 2008; Li and Zhou 2013; Krishnamohan et al. 2012; Bessafi and Wheeler 2006). In these studies high horizontal resolution is cited as a necessary condition to capture the MJO–TC modulation. However, the necessary resolution is not precisely defined. The horizontal resolution in these studies varies from 50 km for the EMCWF (Vitart 2009) and the Geophysical Fluid Dynamics Laboratory (GFDL) High Resolution Atmospheric Model (HiRAM) (Jiang et al. 2012), to 14 km for the Japanese Nonhydrostatic ICosahedral Atmospheric Model (NICAM) (Oouchi et al. 2009). Furthermore, good representations of convection and microphysics are key elements in order to simulate well the MJO in global models (Kim et al. 2012; Holloway et al. 2013; Kim et al. 2014; Kang et al. 2016). Changing the horizontal resolution (with the same physical packages for the same model) need not improve the MJO simulation (Jia et al. 2008; Holloway et al. 2013; Hung et al. 2013). Horizontal resolution might be important primarily through its influence on the models' ability to simulate TCs and their interaction with the ambient environment (Kim et al. 2017).

The S2S models have horizontal grid spacings from 0.25 to 2 degrees. All of them are able to capture, at least qualitatively, the observed eastward propagation of TC genesis anomalies in the

southern hemisphere with increasing MJO phase (Fig. 4). While the observed eastward propagating signal is weaker in the northern hemisphere (Fig. 5), we can still see positive TC genesis anomalies propagating from the NI to the ATL, i.e. from MJO phases 2 & 3 to 8 & 1. The northern hemisphere eastward propagation is stronger in most of the S2S models than in observations. The observed positive anomalies in the ENP for MJO phases 6 & 7 are too strong and expand towards the WNP for the ECMWF and MetFr models. Similarly, the WNP anomalies for MJO phases 8 & 1 are over-predicted and expand towards ENP for the BoM, CMA, and ECMWF models. The JMA model is not able to capture the MJO eastward propagation in the northern hemisphere.

To further identify the MJO phases which are favorable for TC genesis in individual basins, we perform a candy–plot analysis (Section 2), which shows the storm genesis rate binned by MJO phase in each TC basin. In observations (Fig. 6a), 31% of observed ATL hurricanes form when the MJO convection center is in the Indian Ocean (MJO phases 2 & 3). Similarly, in the SIN, a higher rate of genesis (50% of all storms) occurs in the MJO phases 2 & 3. The favorable MJO phases for TC occurrence are 3 & 5 for the NI, 3 & 4 for the AUS, 5 & 6, for the WNP, 7 & 8 for SPC, and 7, 8 & 1 for the ENP. The favorable MJO phases are listed by basin with increasing longitudes so that they line up from the lower-left corner to upper-right corner in Fig. 6a.

The favorable MJO phases in S2S models also show a lower-left to upper right trend, with the exception of the JMA model. The ECMWF model with a horizontal resolutions of 0.25–0.5^{o} best simulates the observed MJO–TC pattern (Fig. 6b), and explains 53% (r^2) of the observed variance. The MetFr, NCEP, CMA, and BoM models (in order of model resolution from 0.7^{o} to 1^{o}) explain 44%, 46%, 36%, and 46% of the variance, respectively. The JMA model does not capture the upward trend of the MJO–TC pattern, because the MJO phases 5 and 6 occur much more frequently in the model than in observations, especially during the northern hemisphere TC seasons (Fig. 6e). As a result, despite having 0.5^{o} horizontal grid spacing, the JMA model

explains only 22% of the observed MJO–TC relationship. Our results suggest that while model resolution plays an important role in simulating TCs, it is probably not the most important factor for simulating the observed modulation of the MJO on TC genesis.

5. Genesis forecast skill

The ability of global models to represent TC genesis depends on the model characteristics, including the dynamical core (Reed et al. 2015a), the physical parameterizations (Reed and Jablonowski 2011), and the model resolution (Kajikawa et al. 2016). These characteristics are also responsible for how well the interaction between the MJO and TCs (or interactions between two weather systems in general) is represented in the model. A good TC forecast (at least at the weather scale) strongly relies on the model initialization, i.e., the data assimilation scheme. Additionally, the skill of the ensemble prediction system is sensitive to the design of the forecast, such as the ensemble size and the range of the model spread, topics that have been broadly studied in the context of weather, (Wilson et al. 1999; Richardson 2001) seasonal (Brankovic and Palmer 1997; Deque 1997; Kumar et al. 2001; Kumar and Chen 2015) and decadal (Sienz et al. 2016) predictions. In this section, we discuss the impact of the model characteristics, initialization, and the ensemble size on the subseasonal (week one to five) genesis prediction skill.

a. Probabilistic TC occurrence prediction – the BSS

We start by quantifying the performance of the S2S models in predicting the probability of weekly TC occurrence using the Brier Skill Score (Section 2). Two different Brier Skill Scores are calculated here. The first (called BSS_c, where 'c' stand for 'constant', dashed lines in Fig. 7) compares the Brier Score of the forecast to that of a constant, annual mean climatological forecast. Positive values of BSS_c means that the model is more skillful than a constant climatological

prediction. We consider storms with genesis in all months, and with this score, forecasts receive credit for correctly matching the annual cycle in TC genesis frequency.

The second Brier Skill Score (called BSS without a subscript here, solid lines in Fig. 7) compares Brier Skill of the forecasts to that of a monthly varying climatological forecast. In this case, only storms that formed in the observed TC seasons are considered (as defined in the Section 3). This is a more conventional and strict measure for evaluating TC genesis prediction skill and the S2S models achieve positive BSS values when they capture deviations from the observed seasonality. A positive BSS in this case means that the model is more skillful than the monthly varying climatological prediction.

The values of BSS_c are often positive, indicating that S2S models are more skillful than an annually constant forecast in most of TC basins. They are also often noticeably greater than those of BSS because the reference forecast used for BSS_c is less skillful than that of BSS. Forecasts from the CMA and those for the NI are exceptions. In the CMA forecasts, the values of BSS_c are much close to those of BSS in SIN, AUS, and WNP. This is because the CMA model's climatology is poor with too many storms forming during the observed off season (not shown). The differences between BSS_c and BSS in the NI are much smaller than in other basins, and there are no positive BSS_c values after week one. In other words, none of the S2S models has more skill in predicting the TC season than an annually constant forecast after week one in the NI. In contrast, JMA, ECMWF, NCEP, and MetFr all have positive BSS_c in the ATL up to week five, although ECMWF is the only model that has positive BSS after week one.

The BSS values for all models drop significantly from week one to week two (Fig. 7). This large drop is connected to our genesis definition, which leads to a high correlation between forecasts and observations (therefore high BSS) by accounting for pre-existing storms at the time of model initialization. Without including pre-existing TCs, the week one BSS values are close to but still

slightly higher than those at week two. By keeping the pre-existing storms we acknowledge that initialization is one of the contributing factors to a dynamical model's prediction skill. Most S2S models are skillful at week one in most basins, with exception of the NI. Low or even negative week one BSS values in a basin are related to poor model initialization in those regions. From weeks two to five, most models' BSS values level off, with the forecast errors saturating at week two. This is consistent with the fact that the genesis climatology does not vary significantly after week two. In some cases, such as for the ECMWF in the SPC and the NCEP in the NI, the model error continues to grow and therefore the BSS values decrease with increasing lead time. The MetFr BSS values in the SIN and NI fluctuate, with relatively higher values in weeks three and five than in weeks two and four.

Based on the BSS evaluation, the ECMWF model has skill up to week five in predicting TC occurrence in the ATL and the WNP, the SPC and ENP up to week two, but has no skill in the SIN, NI, and AUS after week one. The BoM model has positive skill in the WNP up to week five and the AUS and SPC up to week two. The MetFr model is skillful up to week two in the WNP. The CMA, JMA, and NCEP models have no skill after week one. Compared to the existing basin-wide statistical models (Leroy and Wheeler 2008; Slade and Maloney 2013)[1], the ECMWF, BoM and MetFr models have comparable prediction skill. At week one, the BSS from these multiple logistic regression models are 0.13 and 0.17 in the ATL and the ENP (Slade and Maloney 2013), and 0.09 in the SIN, 0.06-0.08 near the AUS, 0.045 in the SPC (Leroy and Wheeler 2008). The highest S2S BSS values (from ECMWF) at week one are: 0.7 in the ATL, the WNP, and the ENP,

[1]While the mathematical formula in the statistical models are similar from one basin to another, the predictors and how sensitive they are to varies. Leroy and Wheeler (2008) focused on the southern oceans and used two MJO indices, the ENSO SST index, the Indo-Pacific SST, and the regional TC seasonal climatology. Slade and Maloney (2013), with the focuses on the Atlantic and east Pacific basins, used MJO and ENSO indices, and regional genesis climatology. For Atlantic basin, an additional predictor representing the variability of SST in the Main Development Region are used. We do not distinguish the different statistical models in our discussions but refer the interested readers to these studies.

0.35 in the AUS, 0.25 in the SIN, and 0.06 in the NI. At week two, the statistical models have BSS values of 0.11 and 0.16 in the ATL and the ENP (Slade and Maloney 2013), 0.07 in the SIN, 0.05-0.07 near AUS and 0.001 in the SPC (Leroy and Wheeler 2008). The highest S2S BSS values at same leadtime are 0.15 in the ATL, 0.19 in the WNP, 0.105 in the ENP (from ECMWF) and 0.056 in the SPC and 0.08 in the AUS (from BoM). None of the S2S models has positive skill in the SIN and NI at week two. From week three, the BSS values from the statistical models are overall better than those from the S2S models.

b. BSS and MJO–TC relationship

To examine whether a model's representation of the observed MJO–TC genesis relationship influences its TC genesis prediction skill, we compute the correlation, across lead times and models, between the BSS values (Fig. 7) and the fraction of variance (r^2) of observed MJO–TC relationship explained by the models from the candy-plot analysis (Fig. 6). These two measures have correlation coefficients of 0.4–0.5 in the ATL, WNP, and ENP, 0.2 in the three southern basins, and close to 0 in the NI (Fig. 8). The positive correlation can be partially attributed to the dependence of both quantities on the lead time; the week one BSS and r^2 values are higher than those at week three in the ATL. The accuracy of the simulated MJO–TC relationship in each model also contributes to its prediction skill, as expected – There are positive correlations across models at a fixed lead time up to week three in all basins, and week five in the ATL, WNP, and ENP. The r^2 values given in Fig. 8 are global, varying by lead time and model, but not by basin. When regional values are used, most basins still have positive small correlations, these are 0.1–0.3 in the AUS, WNP, SPC, and ENP, 0.6 in SIN, but are close to zero in the ATL and negative in the NI (not shown).

c. BSS and ensemble size

Next, we consider the impact of the ensemble size on the forecast skill of subseasonal TC predictions. In particular, we are interested in exploring whether the NCEP model low skill scores are due to its small ensemble size (four members) in the S2S reforecasts. In the climatology and candy–plot analyses, the NCEP system performance is as good as the BoM, and MetFr models, but its skill scores are negative after week one. To examine this question, we first reduce the number of ensemble members for all S2S models to four in the BSS calculation. We focus our analysis in the BoM, ECMWF, and MetFr models at weeks two and five, as the CMA and JMA models only have four and five members, respectively. As expected, the ECMWF, BoM and MetFr BSS values drop to below zero when only four ensemble members are used in the calculation (Fig. 9). The ECMWF system is still more skillful than NCEP with four ensemble members, as is the BoM model in the basins where it is skillful with all 33 ensemble members.

We further calculate the BSS values with increasing ensemble size and find that the BoM model reaches a saturation point with roughly 15 ensemble members, i.e. further increases in the ensemble size do not benefit the genesis prediction skill. The ECMWF and MetFr models, with 11 and 15 ensemble members respectively, seem to be close to their saturation points as well, although their BSS values are not flat yet. In other words, the forecast strategy of the ECMWF and MetFr is probably more efficient than that chose by BoM. Thus, we can expect that the NCEP, JMA, and CMA models will have better skill scores using larger ensemble sizes. This is particularly true in the case of the NCEP model, as the JMA and CMA models have larger bias in their simulated TC genesis climatology.

6. Summary and discussion

Dynamical subseasonal TC genesis prediction is very promising. In this study, we show that the prediction skill of probabilistic TC occurrence forecast from six S2S reforecasts. Most forecasts show skill for week one (days one to seven), when initialization is important. For weeks two to five, prediction skill is associated with the models' ability in simulating TC genesis climatology and the interaction between the MJO and TCs. The forecast skill in individual basin is positively related to how well the global MJO–TC relationship is captured by the model, up to week three in SIN, NI, AUS, and SPC, and up to week five in the ATL, WNP, and ENP. The accuracy of the simulated regional MJO–TC relationship also positively contributes to the genesis skill, particularly, in the SIN.

Among the six models, the ECMWF model has the best performance in reproducing the observed genesis TC climatology and is skillful in forecasting TC genesis up to week five in the ATL and WNP, week two in the SPC and ENP, and week one for the SIN, NI, and AUS. The BoM system has positive skill up to week four in the WNP, week two in the SPC and AUS, and week one in the ATL and SIN. The MetFr model has skill in the WNP up to week two, and week one in the other basins, except the NI. The CMA, JMA, and NCEP models show no skill in predicting TC genesis from weeks two to five.

While current skill scores are still low, the forecast is produced directly without any bias correction. As noted by Vitart et al. (2010) the skill of a model can be extended by a few weeks by a bias correction through post-processing techniques based on the past hindcast performance. Even with no bias-correction, we show that the most skillful models (ECMWF and BoM) have higher skill than the existing regional statistical models at week one, and comparable skill at week two. It is

noted that our results may not reflect the maximum potential skill of some models, such as CMA, JMA and NCEP, since their S2S reforecast ensembles are small (four or five members).

Among the TC basins, subseasonal TC predictions in the Indian Ocean and southern oceans are the hardest, and most of the S2S models have less skill than the monthly climatological probabilities after week one. In contrast, more S2S models have positive skill in North Atlantic and North Pacific basins after week one.

Acknowledgments. The research was supported by NOAA S2S project NA16OAR4310079.

References

Aiyyer, A., and J. Molinari, 2008: MJO and tropical cyclogenesis in the Gulf of Mexico and Eastern Pacific: Case study and idealized numerical modeling. *J. Atmos. Sci.*, **65**, 2691–2704.

Barnston, A. G., N. Vigaud, L. N. Long, M. K. Tippett, and J.-K. E. Schemm, 2015: Atlantic tropical cyclone activity in response to the MJO in NOAA's CFS model. *Mon. Wea. Rev.*, **143**, 4905–4927.

Belanger, J. I., J. A. Curry, and P. J. Webster, 2010: Predictability of North Atlantic tropical cyclone activity on intraseasonal time scales. *Mon. Wea. Rev.*, **138**, 4362–4374.

Bessafi, M., and M. C. Wheeler, 2006: Modulation of South Indian Ocean tropical cyclones by the Madden–Julian Oscillation and convectively coupled equatorial waves. *Mon. Wea. Rev.*, **134**, 638–656.

Brankovic, C., and T. N. Palmer, 1997: Atmospheric seasonal predictability and estimates of ensemble size. *Mon. Wea. Rev.*, **125**, 859–874.

Camargo, S. J., 2013: Global and regional aspects of tropical cyclone activity in the CMIP5 models. *J. Climate*, **26**, 9880–9902.

Camargo, S. J., and A. G. Barnston, 2009: Experimental dynamical seasonal forecasts of tropical cyclone activity at IRI. *Wea. Forecasting*, **24**, 472–491.

Camargo, S. J., M. C. Wheeler, and A. H. Sobel, 2009: Diagnosis of the MJO modulation of tropical cyclogenesis using an empirical index. *J. Atmos. Sci.*, **66**, 3061–3074.

Camargo, S. J., and A. A. Wing, 2016: Tropical cyclones in climate models. *WIREs Climate Change*, **(2)**, 211–237.

Chu, J.-H., C. R. Sampson, A. Lavine, and E. Fukada, 2002: *The Joint Typhoon Warning Center tropical cyclone best-tracks, 1945-2000*. 22pp, Naval Research Laboratory Tech. rep. NRL/MR/7540-02-16.

Deque, M., 1997: Ensemble size for numerical seasonal forecasts. *Tellus A*, **49**, 74–86.

Duvel, J.-P., S. J. Camargo, and A. H. Sobel, 2017: Role of convection scheme in modeling initiation and intensification of tropical depressions over the North Atlantic. *Mon. Wea. Rev.*, **145**, 1495–1509.

Elsberry, R. L., M. S. Jordan, and F. Vitart, 2011: Evaluation of the ECMWF 32-day ensemble predictions during the 2009 season of the western North Pacific tropical cyclone events on intraseasonal timescales. *Asia-Pacific J. Atmos. Sci.*, **47**, 305–318.

Elsberry, R. L., H.-C. Tsai, and M. S. Jordan, 2014: Extended-range forecasts of Atlantic tropical cyclone events during 2012 using the ECMWF 32-day ensemble predictions. *Wea. Forecasting*, **29**, 271–288.

Gottschalck, J., and Coauthors, 2010: A framework for assessing operational Madden–Julian Oscillation forecasts: A CLIVAR MJO working group project. *Bull. Amer. Meteor. Soc.*, **91**, 1247–1258.

Hall, J. D., A. J. Matthews, and D. J. Karoly, 2001: The modulation of tropical cyclone activity in the Australian region by the Madden–Julian Oscillation. *Mon. Wea. Rev.*, **129**, 2970–2982.

Ho, C.-H., J.-H. Kim, J.-H. Jeong, and H.-S. Kim, 2006: Variation of tropical cyclone activity in the South Indian Ocean: El Niño-Southern Oscillation and Madden-Julian Oscillation effects. *J. Geophys. Res.*, **111**, D22 191.

Holloway, C. E., S. J. Woolnough, and G. M. S. Lister, 2013: The effects of explicit versus parameterized convection on the MJO in a large-domain high-resolution tropical case study. Part I: Characterization of large-scale organization and propagation. *J. Atmos. Sci.*, **70**, 1342–1369.

Hung, M.-P., J.-L. Lin, W. Wang, D. Kim, T. Shinoda, and S. J. Weaver, 2013: MJO and convectively coupled equatorial waves simulated by CMIP5 climate models. *J. Climate*, **26**, 6185–6214.

Jia, X., C. Li, J. Ling, and C. Zhang, 2008: Impacts of a GCM's resolution on MJO simulation. *Adv. Atmos. Sci.*, **25**, 139–156.

Jiang, X., M. Zhao, and D. E. Waliser, 2012: Modulation of tropical cyclones over the Eastern Pacific by the intraseasonal variability simulated in an AGCM. *J. Climate*, **25**, 6524–6538.

Kajikawa, Y., Y. Miyamoto, R. Yoshida, T. Yamaura, H. Yashiro, and H. Tomita, 2016: Resolution dependence of deep convections in a global simulation from over 10-kilometer to sub-kilometer grid spacing. *Prog. in Earth and Planet. Sci.*, **3**, 16.

Kang, I.-S., M.-S. Ahn, and Y.-M. Yang, 2016: A GCM with cloud microphysics and its MJO simulation. *Geosci. Lett.*, **3**, 16.

Kikuchi, K., and B. Wang, 2010: Formation of tropical cyclones in the Northern Indian Ocean associated with two types of tropical intraseasonal oscillation modes. *J. Meteorol. Soc. Japan*, **88**, 475–496.

Kim, D., A. H. Sobel, A. D. Del Genio, Y. Chen, S. J. Camargo, M.-S. Yao, M. Kelley, and L. Nazarenko, 2012: The tropical subseasonal variability simulated in the NASA GISS general circulation model. *J. Climate*, **25**, 4641–4659.

Kim, D., and Coauthors, 2014: Process-oriented MJO simulation diagnostic: Moisture sensitivity of simulated convection. *J. Climate*, **27**, 5379–5395.

Kim, D., and Coauthors, 2017: Process-oriented diagnosis of tropical cyclones in high-resolution GCMs. *J. Climate*, submitted.

Kim, J.-H., C.-H. Ho, H.-S. Kim, C.-H. Sui, and S. K. Park, 2008: Systematic variation of summertime tropical cyclone activity in the Western North Pacific in relation to the Madden–Julian Oscillation. *J. Climate*, **21**, 1171–1191.

Klotzbach, P. J., 2010: On the Madden Julian Oscillation – Atlantic hurricane relationship. *J. Climate*, **23**, 282–293.

Klotzbach, P. J., 2012: El Niño–Southern Oscillation, the Madden-Julian Oscillation and Atlantic basin tropical cyclone rapid intensification. *J. Geophys. Res.: Atmospheres*, **117**.

Klotzbach, P. J., 2014: The Madden–Julian Oscillation's impacts on worldwide tropical cyclone activity. *J. Climate*, **27**, 2317–2330.

Klotzbach, P. J., and E. C. Oliver, 2015a: Modulation of Atlantic basin tropical cyclone activity by the Madden–Julian Oscillation (MJO) from 1905 to 2011. *J. Climate*, **28**, 204–217.

Klotzbach, P. J., and E. C. Oliver, 2015b: Variations in global tropical cyclone activity and the Madden–Julian Oscillation (MJO) since the midtwentieth century. *Geophys. Res. Lett.*, **42**, 4199–4207.

Krishnamohan, K. S., K. Mohanakumar, and P. V. Joseph, 2012: The influence of Madden–Julian Oscillation in the genesis of North Indian Ocean tropical cyclones. *Theor. Appl. Climatol.*, **109**, 271–282.

Kumar, A., A. G. Barnston, and M. P. Hoerling, 2001: Seasonal predictions, probabilistic verifications, and ensemble size. *J. Climate*, **14**, 1671–1676.

Kumar, A., and M. Y. Chen, 2015: Inherent predictability, requirements on the ensemble size, and complementarity. *Mon. Wea. Rev.*, **143**, 3192–3203.

Landsea, C. W., and J. L. Franklin, 2013: Atlantic hurricane database uncertainty and presentation of a new database format. *Mon. Wea. Rev.*, **141**, 3576–3592.

Leroy, A., and M. C. Wheeler, 2008: Statistical prediction of weekly tropical cyclone activity in the Southern Hemisphere. *Mon. Wea. Rev.*, **136**, 3637–3654.

Li, R. C. Y., and W. Zhou, 2013: Modulation of Western North Pacific tropical cyclone activity by the PISO. Part I: Genesis and intensity. *J. Climate*, **26**, 2904–2918.

Li, W. W., Z. Wang, and M. S. Peng, 2016: Evaluating tropical cyclone forecasts from the NCEP Global Ensemble Forecasting System (GEFS) Reforecast Version 2. *Wea. Forecasting*, **31**, 895–916.

Liebmann, B., H. H. Hendon, and J. D. Glick, 1994: The relationship between tropical cyclones of the western Pacific and Indian oceans and the Madden-Julian oscillation. *J. Meteor. Soc. Japan*, **72**, 401–411.

497 Madden, R. A., and P. R. Julian, 1972: Description of global circulation cells in the tropics with a
498 40–50-day period. *J. Atmos. Sci.*, **29**, 1109–1123.

499 Maloney, E. D., and D. L. Hartmann, 2000a: Modulation of Eastern North Pacific hurricanes by
500 the Madden–Julian Oscillation. *J. Climate*, **13**, 1451–1460.

501 Maloney, E. D., and D. L. Hartmann, 2000b: Modulation of hurricane activity in the Gulf of
502 Mexico by the Madden-Julian Oscillation. *Science*, **287**, 2002–2004.

503 Maloney, E. D., and D. L. Hartmann, 2001: The Madden-Julian Oscillation, barotropic dynamics,
504 and North Pacific tropical cyclone formation. Part I: Observations. *J. Atmos. Sci.*, **58**, 2545–
505 2558.

506 Mo, K. C., 2000: The association between intraseasonal oscillations and tropical storms in the
507 Atlantic basin. *Mon. Wea. Rev.*, **128**, 4097–4107.

508 Molinari, J., D. Knight, M. Dickinson, D. Vollaro, and S. Skubis, 1997: Potential vorticity, easterly
509 waves, and eastern Pacific tropical cyclogenesis. *Mon. Wea. Rev.*, **125**, 2699–2708.

510 Murakami, H., and Coauthors, 2012: Future changes in tropical cyclone activity projected by the
511 new high-resolution MRI-AGCM. *J. Climate*, **25**, 3237–3260.

512 Murakami, H., and Coauthors, 2015: Simulation and prediction of category 4 and 5 hurricanes in
513 the high-resolution GFDL HiFLOR coupled climate model. *J. Climate*, **28**, 9058–9079.

514 Nakano, M., M. Sawada, T. Nasuno, and M. Satoh, 2015: Intraseasonal variability and tropical
515 cyclogenesis in the western North Pacific simulated by a global nonhydrostatic atmospheric
516 model. *Geophys. Res. Lett.*, **42**, 565–571.

517 Nakazawa, T., 1988: Tropical super clusters within intraseasonal variations over the western Pa-
518 cific. *J. Meteor. Soc. Japan*, **64**, 17–34.

519 Oouchi, K., A. T. Noda, M. Satoh, H. M. H. Tomita, T. Nasuno, and S. ichi Iga, 2009: A simulated
520 preconditioning of typhoon genesis controlled by a boreal summer Madden-Julian Oscillation
521 event in a global cloud–system–resolving model. *SOLA*, **5**.

522 Reed, K. A., J. T. Bacmeister, N. A. Rosenbloom, M. F. Wehner, S. C. Bates, P. H. Lauritzen,
523 J. E. Truesdale, and C. Hannay, 2015a: Impact of the dynamical core on the direct simulation
524 of tropical cyclones in a high–resolution global model. *Geophys. Res. Lett.*, **42**, 3603–3608.

525 Reed, K. A., J. T. Bacmeister, N. A. Rosenblum, M. F. Wehner, S. C. Bates, P. H. Lauritzen, J. E.
526 Truesdale, and C. Hannay, 2015b: Impact of the dynamical core on the direct simulation of
527 tropical cyclones in a high-resolution global model. *Geophys. Res. Lett.*, **42**, 3603–3608.

528 Reed, K. A., and C. Jablonowski, 2011: Impact of physical parameterizations on idealized tropical
529 cyclones in the Community Atmosphere Model. *Geophys. Res. Lett.*, **38**.

530 Richardson, D. S., 2001: Measures of skill and value of ensemble prediction system, their interre-
531 lationship and the effect of ensemble size. *Q. J. R. Meteorol. Soc.*, **127**, 2473–2489.

532 Satoh, M., and Coauthors, 2012: The Intra-Seasonal Oscillation and its control of tropical cyclones
533 simulated by high-resolution global atmospheric models. *Clim. Dyn.*, **39**, 2185–2206.

534 Sienz, F., W. A. Mueller, and H. Pohlmann, 2016: Ensemble size impact on the decadal predictive
535 skill assessment. *Meteorol. Z.*, **25**, 645–655.

536 Slade, S. A., and E. D. Maloney, 2013: An intraseasonal prediction model of Atlantic and East
537 Pacific tropical cyclone genesis. *Mon. Wea. Rev.*, **141**, 1925–1942.

538 Sobel, A. H., and E. D. Maloney, 2000: Effect of ENSO and MJO on the western North Pacific
539 tropical cyclones. *Geophys. Res. Lett.*, **27**, 1739–1742.

Tsai, H.-C., R. L. Elsberry, M. S. Jordan, and F. Vitart, 2013: Objective verifications and false alarm analyses of western North Pacific tropical cyclone event forecasts by the ECMWF 32-day ensemble. *Asia-Pacific J. Atmos. Sci.*, **49**, 409–420.

Vitart, F., 2009: Impact of the Madden Julian Oscillation on tropical storms and risk of landfall in the ECMWF forecast system. *Geophys. Res. Lett.*, **36**.

Vitart, F., 2017: Madden-Julian Oscillation prediction and teleconnections in the S2S database. *Quart. J. Roy. Meteor. Soc.*, **143**, 2210–2220.

Vitart, F., J. L. Anderson, J. Sirutis, and R. E. Tuleya, 2001: Sensitivity of tropical storms simulated by a general circulation model to cumulus parameterization. *Q. J. Roy. Meteor. Soc.*, **127**, 25–51.

Vitart, F., A. Leroy, and M. C. Wheeler, 2010: A comparison of dynamical and statistical predictions of weekly tropical cyclone activity in the southern hemisphere. *Mon. Wea. Rev.*, **138**, 3671–3682.

Vitart, F., and T. N. Stockdale, 2001: Seasonal forecasting of tropical storms using coupled GCM integrations. *Mon. Wea. Rev.*, **129**, 2521–2537.

Vitart, F., and Coauthors, 2017: The subseasonal to seasonal (S2S) prediction project database. *Bull. Amer. Meteor. Soc.*, **98**, 163–173.

Wheeler, M. C., and H. H. Hendon, 2004: An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction. *Mon. Wea. Rev.*, **132**, 1917–1932.

Wilson, L. J., W. R. Burrows, and A. Lanzinger, 1999: A strategy for verification of weather element forecasts from an ensemble prediction system. *Mon. Wea. Rev.*, **127**, 956–970.

Xiang, B., and Coauthors, 2015: Beyond weather time-scale prediction for Hurricane Sandy and Super Typhoon Haiyan in a global climate model. *Mon. Wea. Rev.*, **143**, 524–535, doi:10.1175/MWR-D-14-00227.1.

Yamaguchi, M., F. Vitart, S. T. K. Lang, L. Magnusson, R. L. Elsberry, G. Elliot, M. Kyouda, and T. Nakazawa, 2015: Global distribution of the skill of tropical cyclone activity forecasts on short- to medium-range time scales. *Wea. Forecasting*, **30**, 1695–1709, doi:10.1175/WAF-D-14-00136.1.

Zhang, C., 2013: Madden–Julian Oscillation: Bridging weather and climate. *Bull. Amer. Meteor. Soc.*, **94**, 1849–1870.

Zhao, M., I. M. Held, and S.-J. Lin, 2012: Some counter-intuitive dependencies of tropical cyclone frequency on parameters in a GCM. *J. Atmos. Sci.*, **69**, 2272–2283.

LIST OF TABLES

TABLE 1. Characteristics of the six S2S reforecasts used here

Model	forecast time	Resolution	Period	Ens. size	Frequency & Sample size
BoM	0–64 days	2°, L17	1981–2013	33	∼5 days & 2160
CMA	0–61 days	1°, L40	1994–2014	4	daily & 7665
ECMWF	0–46 days	0.25° for first 10 days 0.5° after day 10, L91	1994–2014	11	∼4 days & 2058
JMA	0–33 days	0.5°, L60	1981–2010	5	∼10 days & 1079
MetFr	0–61 days	∼0.7°, L91	1993–2014	15	∼15 days & 528
NCEP	0–44 days	∼1°, L64	1999–2010	4	daily & 4380

LIST OF FIGURES

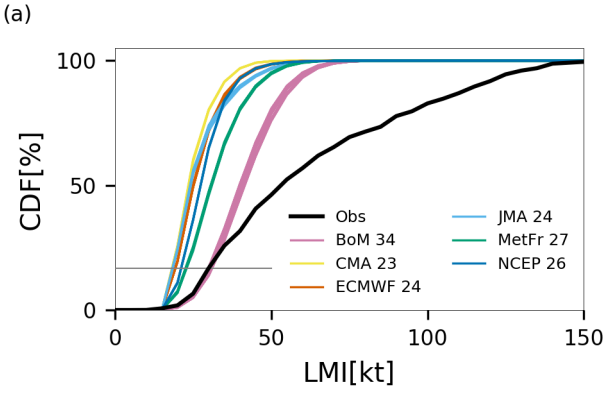


FIG. 1. Cumulative density function (CDF, in percentage) of the observed (black) and the simulated (colored) LMI from each of the ensemble members from the six S2S models.

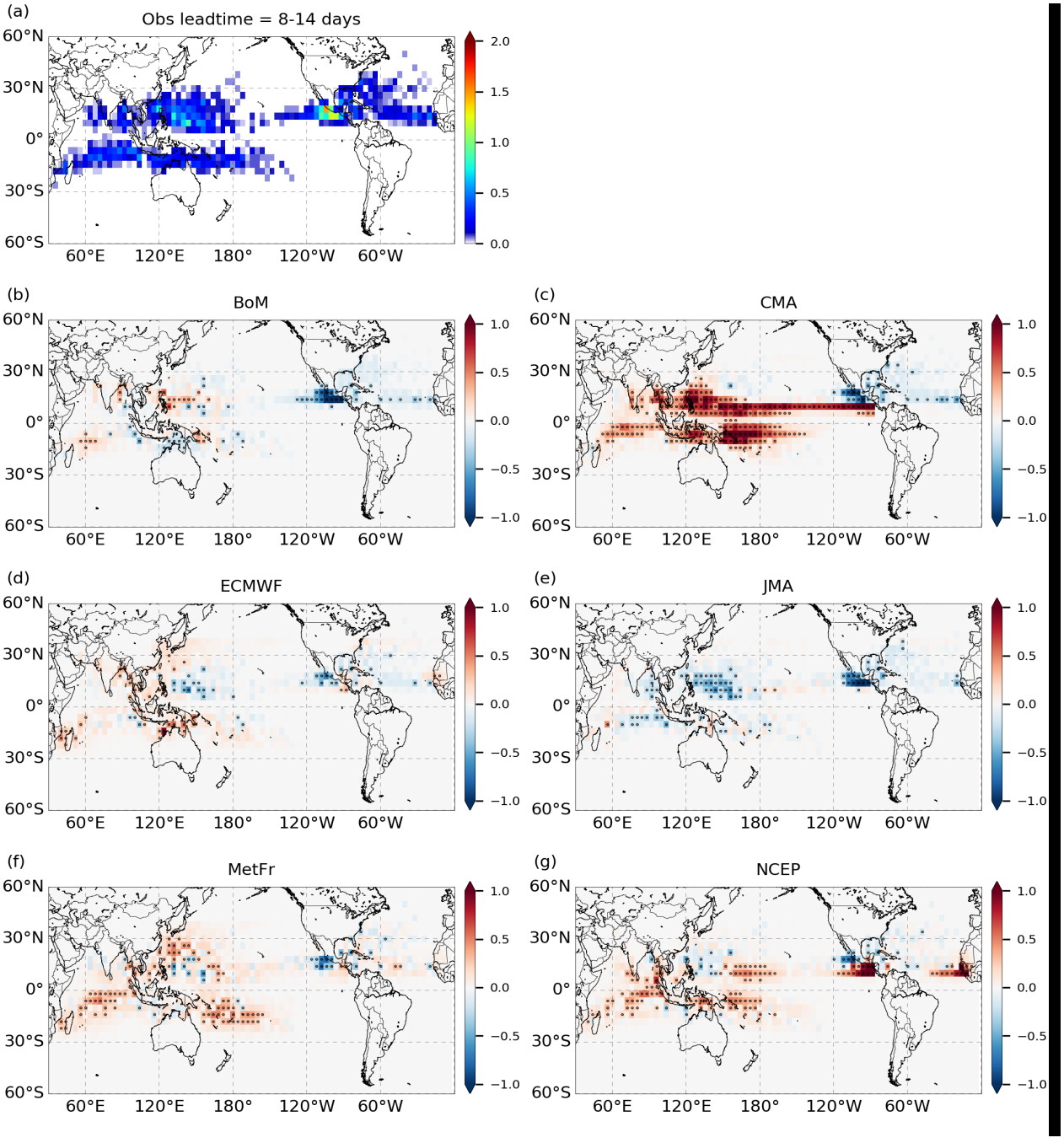
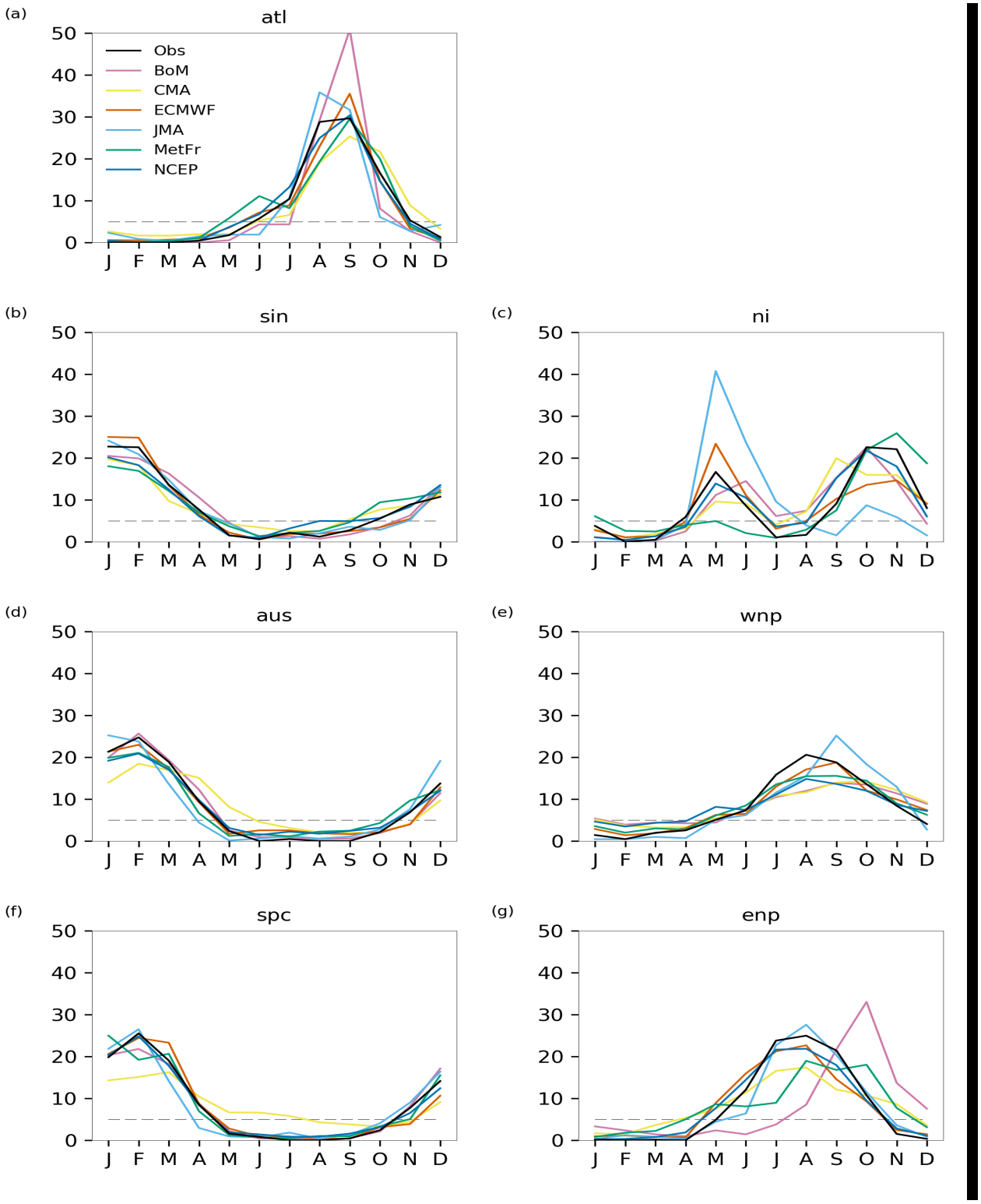


FIG. 2. (a) The observed genesis density in storm numbers per 4 by 4 degree per year. (b–g) Week 2 ensemble mean genesis density biases in BoM, CMA, ECMWF, JMA, MetFr, and NCEP. The black dots indicating where the error is larger than 1 standard deviation of the natural variability.



610 FIG. 3. Observed (black) and simulated (colored) ensemble mean seasonality (in percentage). At each basin,
611 the months when the genesis rate is larger than 5% (gray dash lines) of the annual genesis rate are defined as TC
612 season.

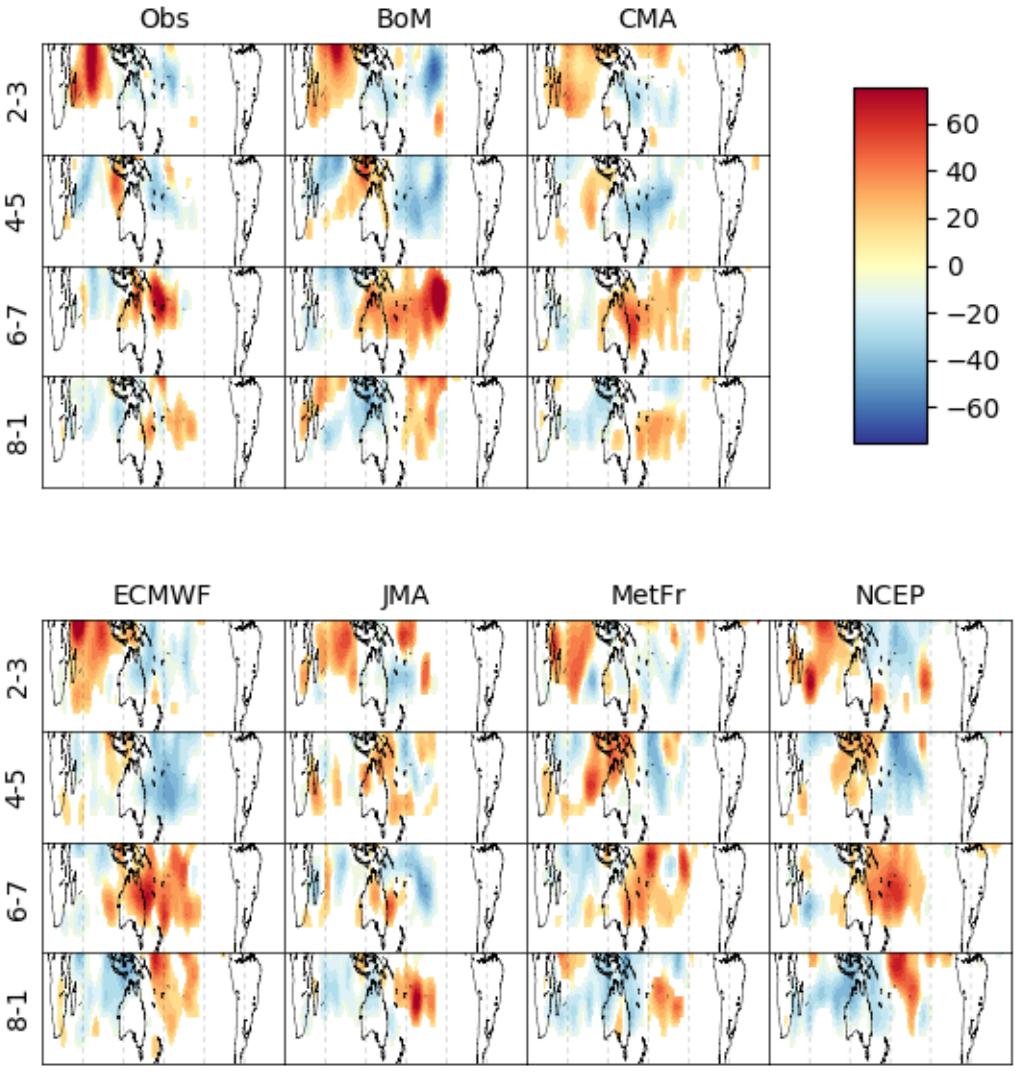
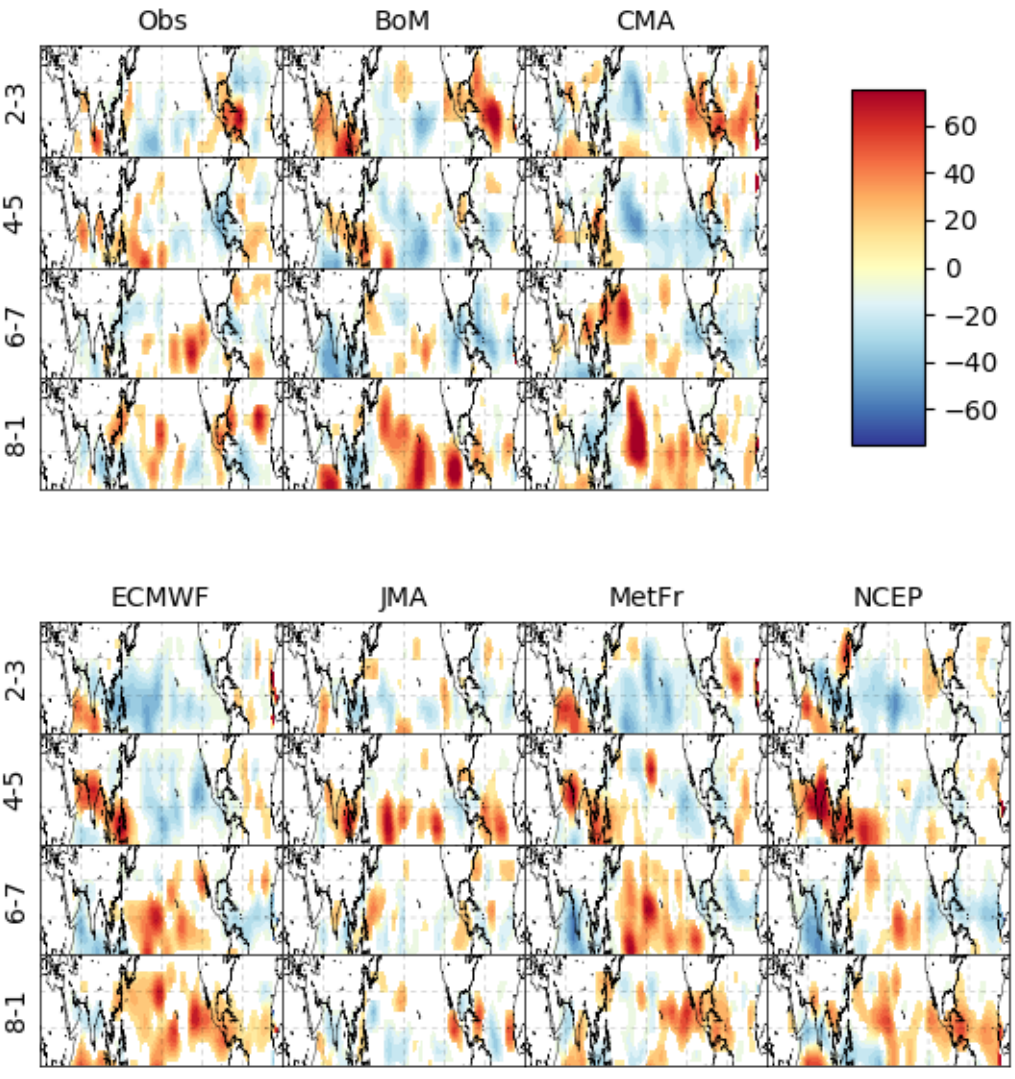


FIG. 4. Ensemble mean genesis anomalies (in percentage) at every 2 MJO phases from 2–3, 4–5, 6–7 and 8–1 (top down) in southern hemisphere from observations and the week 2 forecasts from six S2S models.



615 FIG. 5. Ensemble mean genesis anomalies (in percentage) at every 2 MJO phases from 2–3, 4–5, 6–7 and 8–1
616 (top down) in northern hemisphere from observations and the the week 2 forecasts from six S2S models.

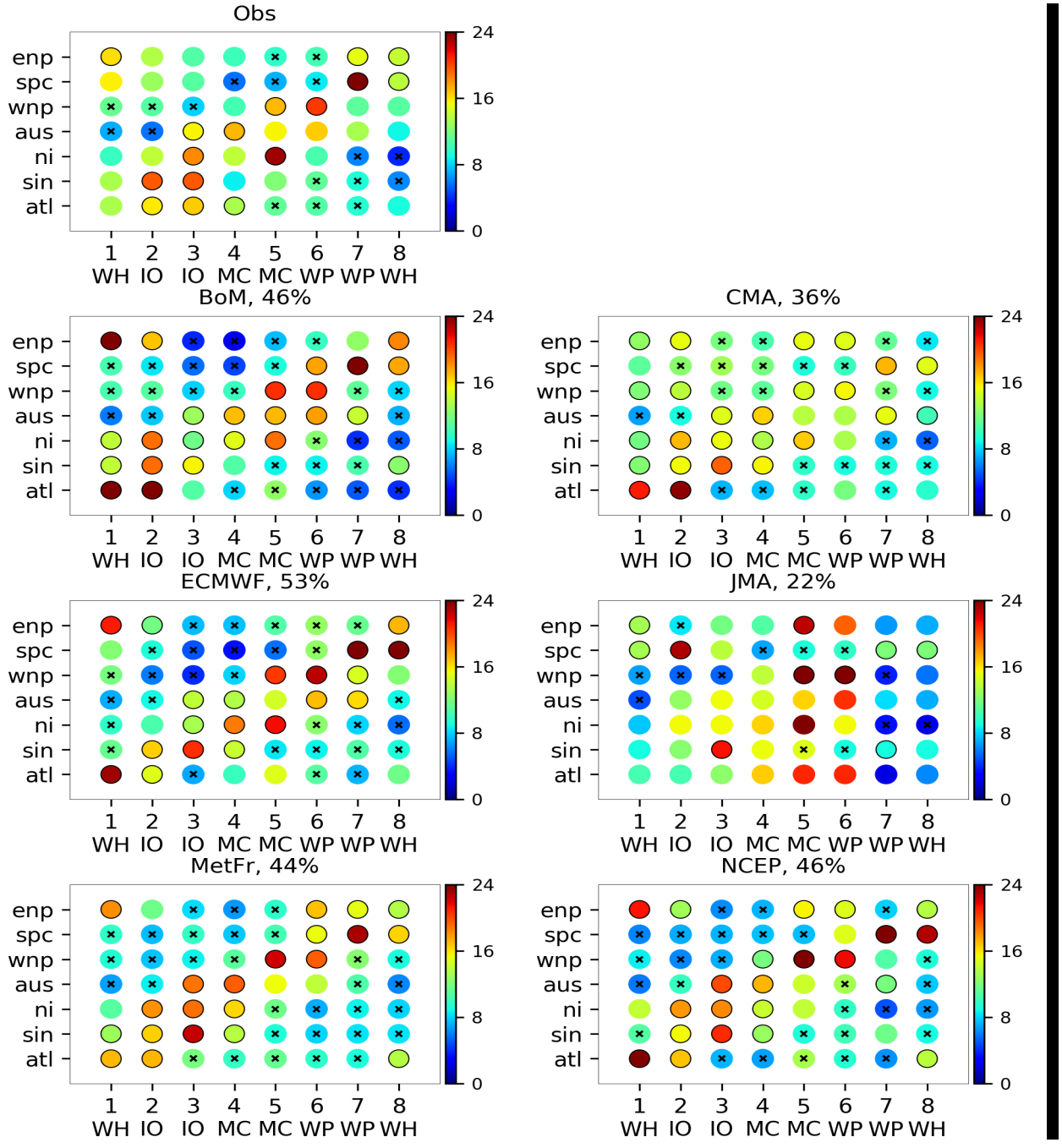
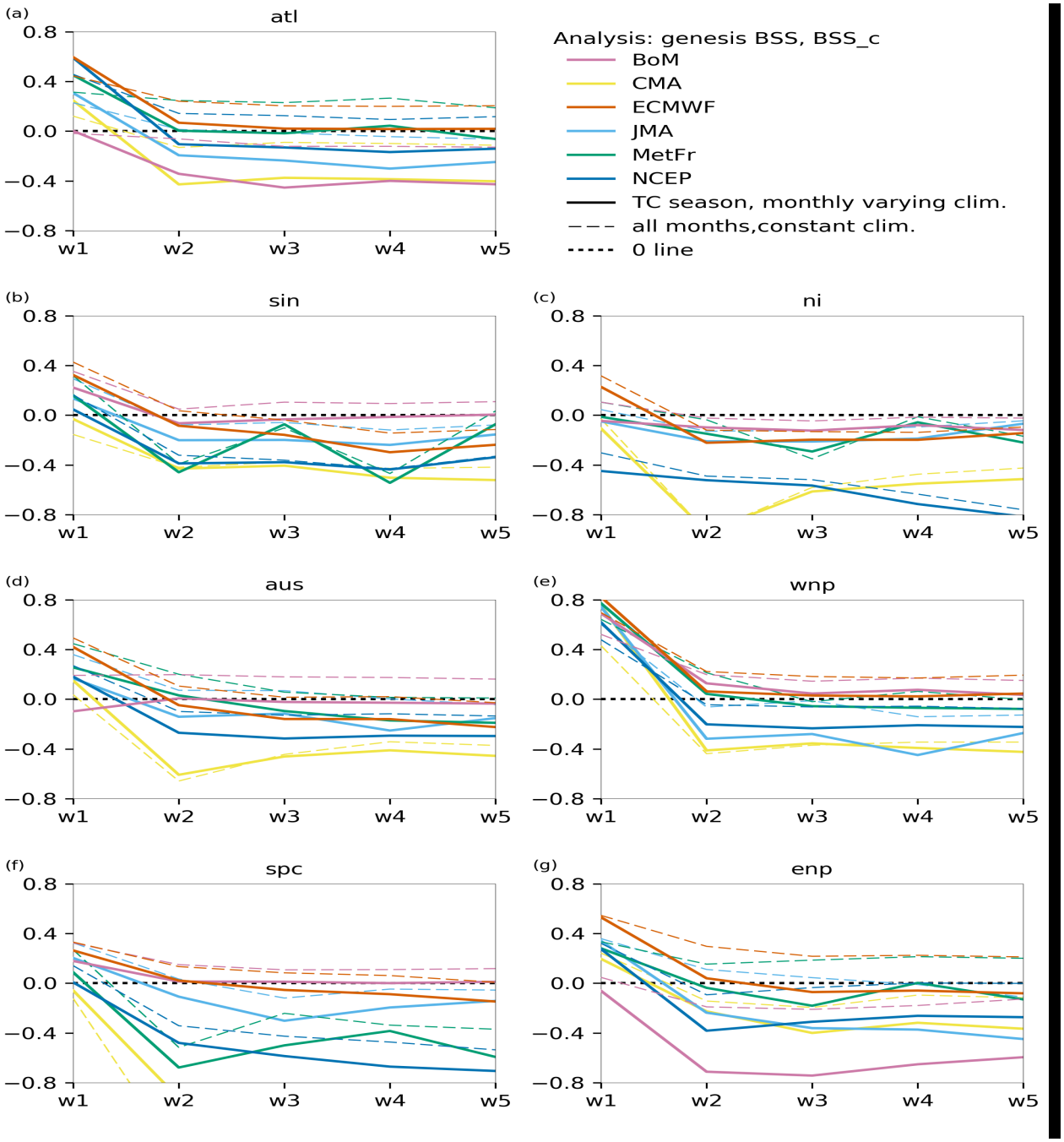
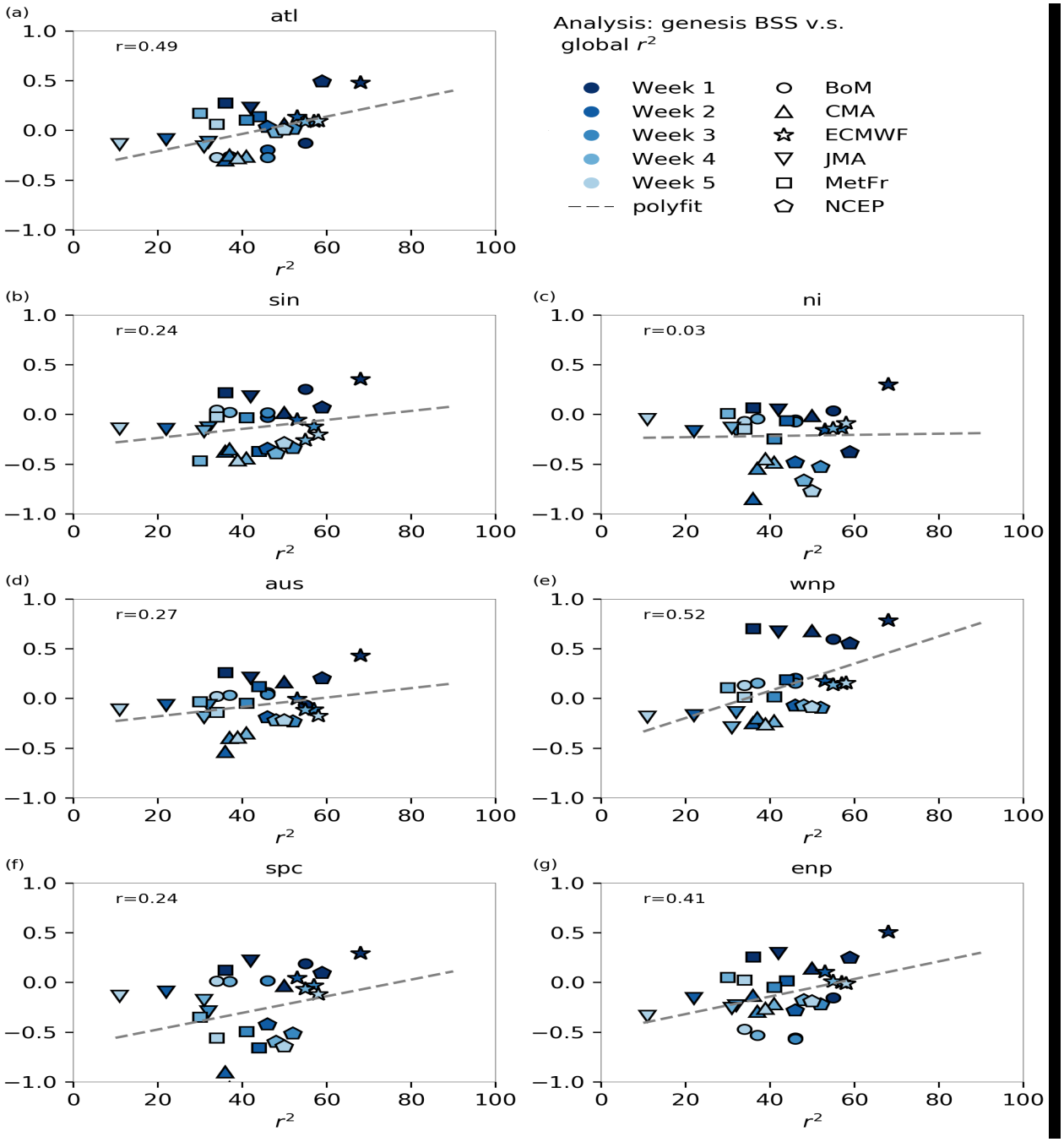


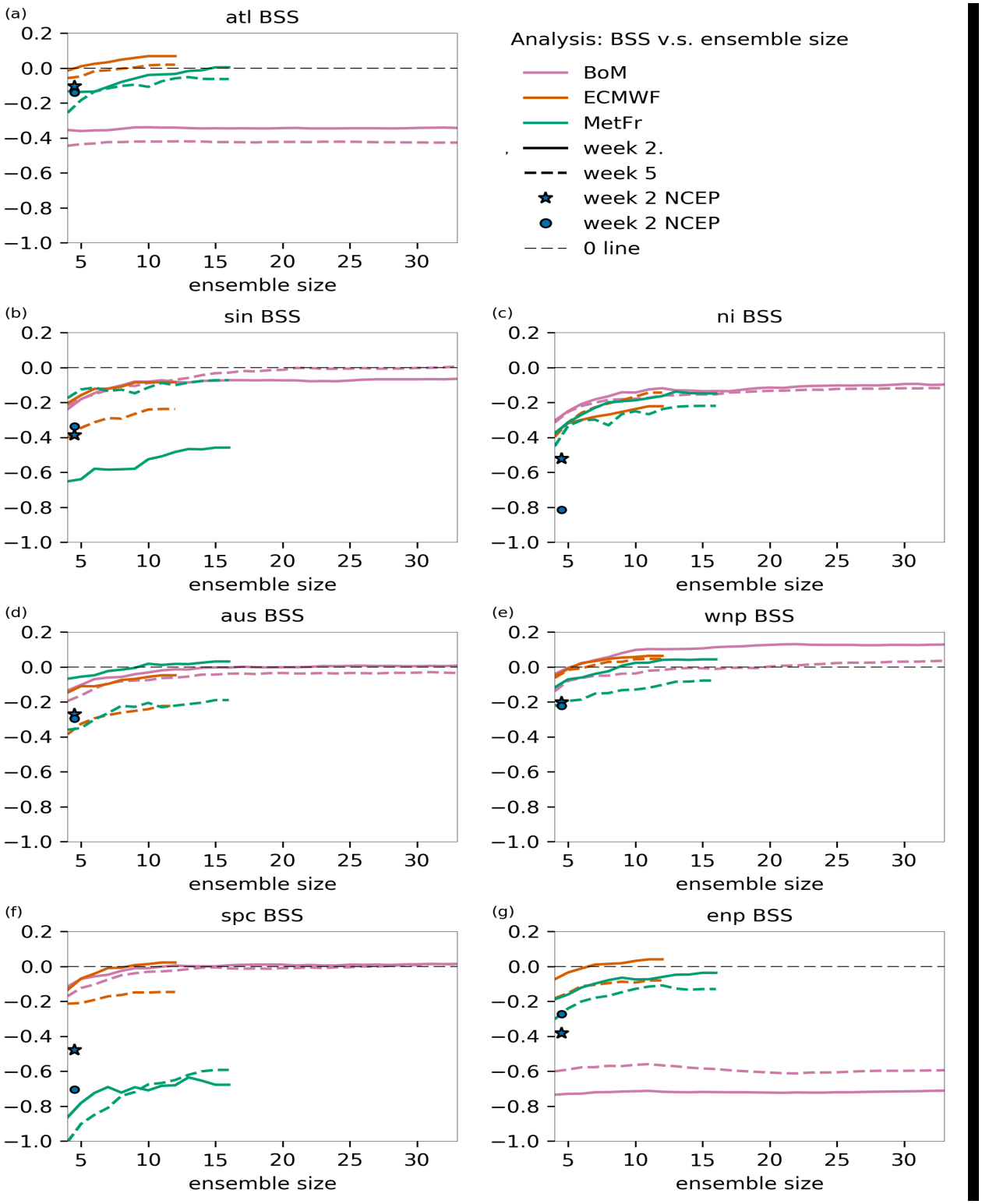
FIG. 6. Candy plot for MJO–TC relationship in observations and six S2S models from week 2 forecasts. Color of each candy indicates the probability density function (PDF in percentage) in the corresponding MJO phase in the basin. Sum of the circles across the MJO phases in each basin is 100%. The black circle at the edge indicates the value is above 90% significant level while the cross symbol, $\times$, at the center means the value is below 10% significant level. In the title of each subplot from simulations we label the r^2, which represents the fraction of the observed pattern explained by the model simulation.



623 FIG. 7. Weekly, basin–wide BSS in the S2S model relative to the constant (dashed lines) and to the monthly
624 varying climatological (solid) predictions. See Sections 2 and 5 for details.



625 FIG. 8. Basin-wide BSS (from Fig. 7) as a function of the r^2 from candy–plot analyses from week 1 to week
626 5 (colors) in all six S2S models (symbols). The dashed line shows the best-fit line and the correlation coefficient
627 between BSS and the r^2 is labeled at the upper left corner in each subplot.



628 FIG. 9. Basin-wide BSS at week 2 and week 5 as a function of ensemble size in BoM, ECMWF, and MetFr.

629 The blue star and circle indicate the BSS from NCEP (4 ensemble members) at week 2 and 5, respectively.