

Seamless Precipitation Prediction Skill in the Tropics and Extratropics from a Global Model

HONGYAN ZHU AND MATTHEW C. WHEELER

Centre for Australian Weather and Climate Research, Melbourne, Australia

ADAM H. SOBEL

*Department of Applied Physics and Applied Mathematics, Department of Earth and Environmental Sciences,
Lamont-Doherty Earth Observatory, Columbia University, New York, New York*

DEBRA HUDSON

Centre for Australian Weather and Climate Research, Melbourne, Australia

(Manuscript received 12 July 2013, in final form 3 December 2013)

ABSTRACT

The skill with which a coupled ocean–atmosphere model is able to predict precipitation over a range of time scales (days to months) is analyzed. For a fair comparison across the seamless range of scales, the verification is performed using data averaged over time windows equal in length to the lead time. At a lead time of 1 day, skill is greatest in the extratropics around 40°–60° latitude and lowest around 20°, and has a secondary local maximum close to the equator. The extratropical skill at this short range is highest in the winter hemisphere, presumably due to the higher predictability of winter baroclinic systems. The local equatorial maximum comes mostly from the Pacific Ocean, and thus appears to be mostly from El Niño–Southern Oscillation (ENSO). As both the lead time and averaging window are simultaneously increased, the extratropical skill drops rapidly with lead time, while the equatorial maximum remains approximately constant, causing the equatorial skill to exceed the extratropical at leads of greater than 4 days in austral summer and 1 week in boreal summer. At leads longer than 2 weeks, the extratropical skill flattens out or increases, but remains below the equatorial values. Comparisons with persistence confirm that the model beats persistence for most leads and latitudes, including for the equatorial Pacific where persistence is high. The results are consistent with the view that extratropical predictability is mostly derived from synoptic-scale atmospheric dynamics, while tropical predictability is primarily derived from the response of moist convection to slowly varying forcing such as from ENSO.

1. Introduction

Extratropical and tropical weather have different characteristics. Extratropical weather is dominated by baroclinic disturbances that obtain their energy from the vertical shear in the mean flow and the available potential energy associated with the horizontal temperature gradients that balance that shear (Charney 1947; Lorenz 1955). Precipitation tends to be strongly forced at large scales by isentropic uplift along fronts and

dynamical lifting due to the advection of quasi-balanced upper-level potential vorticity anomalies (Bluestein 1993). Tropical weather exists in an environment of much weaker pressure and temperature gradients and (at least in the zones of greater climatological precipitation) higher humidity (Charney 1963, 1969). Tropical precipitation is typically a result of deep convection and closely associated stratiform rain (Schumacher and Houze 2003). The convection is often organized into wavelike disturbances (Wheeler and Kiladis 1999), but it is unclear to what extent these disturbances are dynamically independent entities that organize the convection (as is the case for many extratropical disturbances) as opposed to resulting from spontaneous “self-aggregation” (Mapes 1993; Bretherton et al. 2005) of the convection itself.

Corresponding author address: Dr. Matthew Wheeler, CAWCR/
Bureau of Meteorology, GPO Box 1289, Melbourne, VIC 3001,
Australia.
E-mail: m.wheeler@bom.gov.au

The lead time at which tropical weather becomes inherently unpredictable is not well known, but is generally thought to be shorter than that for extratropical weather (Shukla 1989; Boer 1995).

At the same time, a substantial literature has established that the tropics are the source of most potential predictability globally on seasonal to interannual time scales (Charney and Shukla 1981; Goddard et al. 2001). There are some extratropical sources of predictability on time scales of weeks to months, such as from stratospheric effects (Baldwin and Dunkerton 1999; Polvani and Kushner 2002), snow cover (Cohen and Entekhabi 1999), sea ice (Holland et al. 2013), and soil moisture (Koster and Suarez 2003). It appears, however, that tropical sea surface temperature variations, particularly those resulting from the El Niño–Southern Oscillation (ENSO) phenomenon, have a greater impact on the global climate (Hoerling and Kumar 2002). The impacts of ENSO are felt strongly not only in the tropics, but also in many extratropical regions (Kiladis and Diaz 1989), because of atmospheric teleconnections (DeWeaver and Nigam 2004). Of course, at seasonal-to-interannual time scales one is not predicting the daily weather, but only the averages over a month or a season.

The intraseasonal time scale lies between daily weather and seasonal climate. On that intermediate time scale we expect the Madden–Julian oscillation (MJO) to be a source of predictability in the tropics (Waliser et al. 2006), while there may be some additional predictability associated with low-frequency extratropical modes driven by eddy–mean flow interactions (Baldwin et al. 2003).

We expect, then, that the extratropics are more predictable than the tropics at lead times of a day to a week, while the tropics are more predictable at climate scales of months to a year (Shukla 1989; Sobel 2012). Our interest here is in testing whether that expectation is correct, and in studying the transition between the two time scales, in both tropical and extratropical latitudes.

We study the relative skill of a particular prediction system in predicting tropical versus extratropical weather and climate across a range of time scales, from daily to monthly. We use a coupled ocean–atmosphere ensemble forecast system that is used operationally for prediction on a range of time scales, from a few days to seasons. We focus on precipitation as it is of interest and has utility in both tropical and extratropical regions (as opposed to pressure and temperature, which vary much less in the tropics than extratropics and are therefore of less interest there). The model used in the forecast system contains some representation of the main sources of predictability described above (Marshall et al. 2011, 2012, 2013; Wang et al. 2011; Cottrill et al. 2013; Hudson et al. 2013), with the exception of the stratospheric

sources (Roff et al. 2011) and sea ice variations. Thus, while this is not a true study of potential predictability limits, the prediction skill from the current model should be somewhat comparable to those limits, at least within the realms of our current knowledge. Further, it is of interest to determine the comparative prediction skill that is currently available from an operational system.

The essence of our approach is as follows. We compute the prediction skill at a range of lead times, from 1 day to 1 month. As the lead time increases, we also increase the length of the time window over which the data are averaged for verification. This is intended to capture the fact that we are transitioning from weather to climate prediction as the lead time increases, and to allow the transition to occur smoothly. The skill is computed for both total precipitation and anomalies, and comparison is made with the skill achievable by a persistence forecast of the precipitation anomalies. For comparison we also evaluate the forecasts at varying lead time but with a fixed verification window of 1 day.

2. Data and method

a. POAMA-2 ensemble forecast system

We use the Bureau of Meteorology’s dynamical Predictive Ocean Atmosphere Model for Australia (POAMA; Alves et al. 2003) version 2 configured for multiweek predictions (“POAMA-2 multi-week” is abbreviated to P2-M; Hudson et al. 2013). Earlier versions of POAMA were designed for seasonal forecasting; however, improvements to the generation of initial conditions to use perturbed atmosphere and ocean initial conditions and a burst ensemble (i.e., an ensemble starting from a single initial time as opposed to a lagged ensemble), as well as the use of three different model configurations to form a multimodel ensemble, have made P2-M applicable for shorter-range forecasts as well, especially at the intraseasonal time scale (Hudson et al. 2013).

The atmospheric component of P2-M is run in spectral space with a triangular truncation at wavenumber 47 (approximately a 250-km grid) and 17 vertical levels. It includes a land component that is a simple bucket for soil moisture and three soil levels for temperature. The ocean model has a zonal resolution of 2°, a meridional resolution of 0.5° within 8° of the equator increasing to 1.5° near the poles, and 25 levels. While the atmospheric model has a relatively coarse resolution compared to modern numerical weather prediction models, it is comparable to what has commonly been used for seasonal prediction over the last decade and is considered adequate to resolve the key sources of predictability discussed in the introduction. Further details of these model

components are provided in Hudson et al. (2013) and references therein.

Also important are the methods employed for producing initial conditions and perturbations to the initial conditions to generate a forecast ensemble. The unperturbed initial conditions are provided by separate data assimilation schemes for the ocean versus the atmosphere and land. The atmosphere and land initial conditions are created by nudging zonal wind, meridional wind, atmospheric temperature, and humidity in the atmosphere–land component of the model (when run prior to hindcasts or forecasts being made, and forced with observed sea surface temperatures) toward an observationally based analysis (Hudson et al. 2011). The analysis used is the 40-yr European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-40; Uppala et al. 2005) for the period from 1980 to August 2002, and the Bureau of Meteorology's operational global numerical weather prediction (NWP) analysis thereafter. Ocean initial conditions are derived using a pseudoensemble Kalman filter data assimilation system (Yin et al. 2011). In situ ocean temperature and salinity observations are assimilated and corrections to currents are generated based on the ensemble cross-covariances with temperature and salinity.

Perturbations to the initial conditions of the central member are generated using a coupled breeding scheme that produces perturbations to all components of the coupled system in a consistent fashion. Ten different perturbed states are produced, which, together with the unperturbed central member, provides for 11 different initial states from which to start a burst ensemble (Hudson et al. 2013).

In addition to the perturbed initial states, which allow for an estimate of forecast uncertainty due to sensitivity to initial condition errors, a multimodel ensemble comprising three different model configurations is used to provide a sample of model uncertainty. The three configurations are differentiated by their use of 1) standard physics with no flux correction, 2) bias correction of fluxes at the air–sea interface, and 3) as in version 1, except with modified atmospheric physics in the form of an alternative shallow convection parameterization. Each model configuration is run with the 11 different initial conditions to provide a 33-member ensemble. The coupled breeding of initial states uses the first version. The climate drift and seasonal prediction skill of each model configuration are discussed in Lim et al. (2009, 2010).

The above description of the P2-M system applies to both the hindcasts (i.e., forecast runs that are started using initial states from previous times) as well as forecasts run in real time. In this work we analyze the skill of the hindcasts only. However, given the same configuration

of the hindcast and real-time systems, we expect that the skill of the hindcasts should be comparable to a suitably large sample of real-time forecasts, assuming a relatively stable climate. The hindcasts we analyze have start times on the first, 11th, and 21st days of each month of the year. To match the period of available global daily precipitation observations (see section 2b), we analyze the period between 1996 and 2009 only.

b. Observations

The observational dataset for verification in this paper is the Global Precipitation Climatology Project (GPCP) daily precipitation with 1° resolution (Huffman et al. 2001). The GPCP data are a blended product derived from both station observations and satellite measurements. The satellite data are sourced from both geostationary and polar-orbiting platforms. When this work commenced, the available daily GPCP data (version 1.1) extended from October 1996 to August 2009, which is the period we have chosen to evaluate the model hindcasts. We map the GPCP data to the model grid by first interpolating the GPCP data to a 0.5° grid, and then averaging in the zonal and meridional directions to match the POAMA grid spacing. Our analysis therefore concentrates on precipitation that is area averaged over a scale of about 250 km, providing a reasonable representation of most synoptic-scale weather. Known problems exist in the GPCP data at high latitudes (Bolvin et al. 2009); however, our results and survey of the literature give us enough confidence to show the skill calculations to a latitude of 80°.

c. Measures of prediction skill

We assess skill by comparing the P2-M forecasts with the verifying GPCP observations. We computed a number of different verification measures, each having different strengths and weaknesses (not shown), and verified that the conclusions are not sensitive to which ones we use. We therefore choose to show the simplest measures for this paper: the correlation of the ensemble mean total precipitation with the observed verification data (hereafter CORT) and the correlation of the ensemble mean precipitation anomalies with the observed anomalies (hereafter CORa). These correlations are computed over time (i.e., using data from many different verification windows), separately for each grid point and each lead time. In the case of CORT, this measure is affected by both the model's ability to accurately represent the climatological seasonal cycle in its forecasts, and the variability. In the case of CORa, the seasonal cycles are removed from both the observations and forecasts by removing their respective climatologies. For the model this is the hindcast climatology, which is

a function of both lead time and start day and month. CORa therefore is affected only by the model's ability to forecast the variability about that climatology.

Computing and showing CORt is a more usual practice for the weather prediction community (e.g., Ebert 2001), whereas concentrating on anomalies (i.e., CORa) is more usual for the seasonal prediction community (e.g., Cottrill et al. 2013). This is partly because users of weather information are more interested in total precipitation, whereas users of climate information are more interested in whether future conditions may be wetter or drier than normal (i.e., anomalies). Another reason is that the numerical weather prediction community tend not to produce large hindcast datasets (which are necessary for computing a model climatology) whereas seasonal prediction systems require hindcasts to assess and remove the climate drift that becomes noticeable at longer lead times (Stockdale 1997). Perhaps the main disadvantage of these two verification measures is that they ignore the probabilistic nature of the ensemble. Their other disadvantage is that the correlation is insensitive to mean bias. However, noting that in this work we are more interested in the relative skill between regions and lead times, we feel that their simplicity outweighs these disadvantages.

For both CORt and CORa the correlation is calculated at each grid point as

$$\text{COR} = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}$$

where x is the ensemble mean forecast precipitation (using totals for CORt and anomalies for CORa), y is the observed precipitation value (totals or anomalies), and n is the number of verification times; each sum is calculated over n values.

To study the differences in skill between different seasons, we show computations of CORt and CORa for the contrasting seasons of December–February (DJF) and June–August (JJA) for which n is 117 and 108 respectively (13 or 12 years \times 3 months per season \times 3 forecast starts per month). When computing the correlation for a particular season, one may at first think that CORt will be equal to CORa since the correlation automatically removes the respective time mean values from the two fields that are being correlated. However, the seasonal cycle is not constant across a 3-month season, so in practice CORt and CORa are not the same.

Further details on the calculation of the climatological season cycles are as follows: For both the observed and model forecast precipitation, the exact same years are used to compute the climatology (i.e., October 1996–August

2009). We are also careful to use the exact same days of the year from the verifying observations as from the model. For example, consider the forecast of the second week from the initial condition of 11 December 2001. The dates of the second week are 19–25 December 2001. The observed climatology for this forecast is computed by averaging the precipitation data for 19–25 December for all 13 years (i.e., 7×13 days of observed data). The model climatology for this forecast is computed by averaging the model precipitation for 19–25 December from all 33 ensemble members of all 13 years of forecasts that were initialized from 11 December (i.e., $7 \times 33 \times 13$ days of model data). Note that unlike Hudson et al. (2013) we do not need to compute different hindcast climatologies for each of the three different model configurations because the resulting ensemble mean anomaly is the same with our use of a multimodel climatology.

d. Forecast time window definition

As stated in the introduction, we take the approach of widening the time averaging window of the forecasts and verifying observations when looking at progressively longer lead times. For example, for a forecast lead time of 1 day we use an averaging window of 1 day, and for a lead time of 1 week we use a window of 1 week. A schematic of this approach and the terminology we use to label it is provided in Fig. 1. Our intention is to provide a seamless transition from weather to climate prediction in this analysis of skill. Note that “1d1d” is what is usually called “day 2” in other papers, and “1w1w” is what is usually called “week 2.” The longest window and lead time combination we consider is 4 weeks (i.e., 4w4w); 4w4w is roughly equivalent to “month 2” in other papers, noting that a month is roughly 4 weeks long. We also study the intermediate window/lead times of 2d2d, 4d4d, and 2w2w, providing a total of six different time scales. Later in the paper we also evaluate forecasts using the more traditional approach of varying the lead time but with a fixed verification window of 1 day. Using the terminology discussed above, this latter analysis focusses on forecasts for 1d0d to 1d2w, where 1d0d is equivalent to the first 24 h of the forecast (see Fig. 9).

e. Seasonal definition

We show our computations of CORt and CORa for the seasons of DJF and JJA only (as described above). Note that we use the starting date (i.e., initial condition) of the model forecasts to determine the season rather than the verifying time. This means that for the 4w4w calculations the verification times extend up to ~ 7 weeks after the end of each season (noting that the latest hindcast each season is initialized on the 21st of the month).

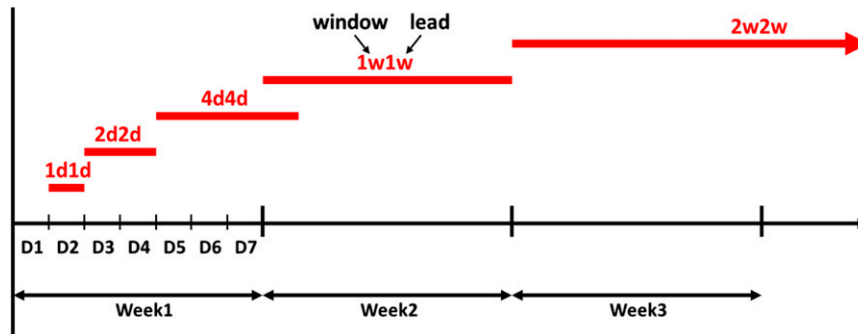


FIG. 1. Schematic of the time window and lead time definitions used in this analysis. The horizontal axis represents forecast time from the initial condition. The expression “1d1d” refers to an averaging window of 1 day at a lead time of 1 day. Similarly, “2d2d” represents an averaging window of 2 days at a lead time of 2 days, and so on. Note that 1d1d is what is usually called “day 2” in other papers, and 1w1w is what is usually called “week 2.”

For example, the 4w4w calculations for JJA will include verifying data from 30 June to 16 October.

3. Results

a. COR_t —Correlation with the ensemble mean totals

Maps of COR_t for the contrasting seasons of DJF and JJA and for the window/lead time combinations of 1d1d, 1w1w, and 4w4w are displayed in Fig. 2. Positive values indicate positive skill in the sense that there is an in-phase relationship between the forecast and observed values. For 1d1d, a positive skill is achieved everywhere except over the subtropical dry zones over Africa, and the eastern Atlantic and eastern Pacific Oceans. In DJF the highest large-scale 1d1d COR_t (>0.5) is achieved over the North Pacific and North Atlantic, whereas in JJA the region of highest large-scale COR_t is over the midlatitudes of the Southern Hemisphere. This is consistent with previous work (Ebert et al. 2003) that shows that extratropical precipitation is generally easier to predict for short lead times in winter when it is associated mainly with synoptic-scale systems such as fronts, whereas in summer it is more often associated with convective systems such as thunderstorms that are harder to predict. (This short-range seasonality in the extratropics will become more apparent in the zonally averaged skill plots in Figs. 3, 6, and 10.)

Interestingly, the 1d1d COR_t maps (Fig. 2) also indicate some patches of very high skill in the equatorial zone, especially over the Indian and Pacific Ocean sectors in DJF. This was not initially anticipated given our review of published papers as discussed in the introduction. We did not expect such high skill in the tropics at short lead times.

At the longer window/lead time scales of 1w1w and 4w4w, the COR_t maps of Fig. 2 indicate greatest skill

($COR_t > 0.7$) over the tropical Pacific, especially in DJF. This appears to be the result of the predictability provided by ENSO. Greatest precipitation skill ($COR_t \geq 0.9$) is achieved over the central-eastern equatorial Pacific because this is where precipitation is most strongly related to the SST variations of ENSO (Weare 1987). Indeed, these maps look much like the maps of SST skill for POAMA provided in Cottrill et al. (2013). Further, DJF is when ENSO events typically reach their peak SST anomaly, so greater precipitation prediction skill from forecasts initialized in DJF is somewhat expected. Elsewhere in the tropics there is moderate skill ($COR_t > 0.5$) over the Indian Ocean and just to the north of the Maritime Continent, especially in DJF, which appears to be at least partially a result of the MJO (cf. Fig. 8 of Marshall et al. 2011).

A further interesting feature from Fig. 2 is the band of $COR_t > 0.3$ extending around the globe at the latitudes of 50° – 65° S for 4w4w in DJF. Our initial thought was that this may be related to the southern annular mode and its relationship with ENSO (L’Heureux and Thompson 2006). This relationship is known to be strongest in DJF. However, as we will show later, this signal mostly disappears when the skill associated with the climatological seasonal cycle is removed (using COR_a), indicating that it stems from a pronounced seasonal cycle that is well represented by the model during DJF for those latitudes.

Information from the intermediate window/lead times is presented in Fig. 3, which shows the zonally averaged COR_t for the model forecasts for six different lead times and averaging windows, extending from 1d1d to 4w4w. In the extratropics at short lead times, greater skill in winter than summer, as discussed above, is readily apparent. In both the Northern and Southern Hemisphere the zonally averaged COR_t is greater than 0.5 in winter and less than 0.5 in summer (1d1d window/lead time).

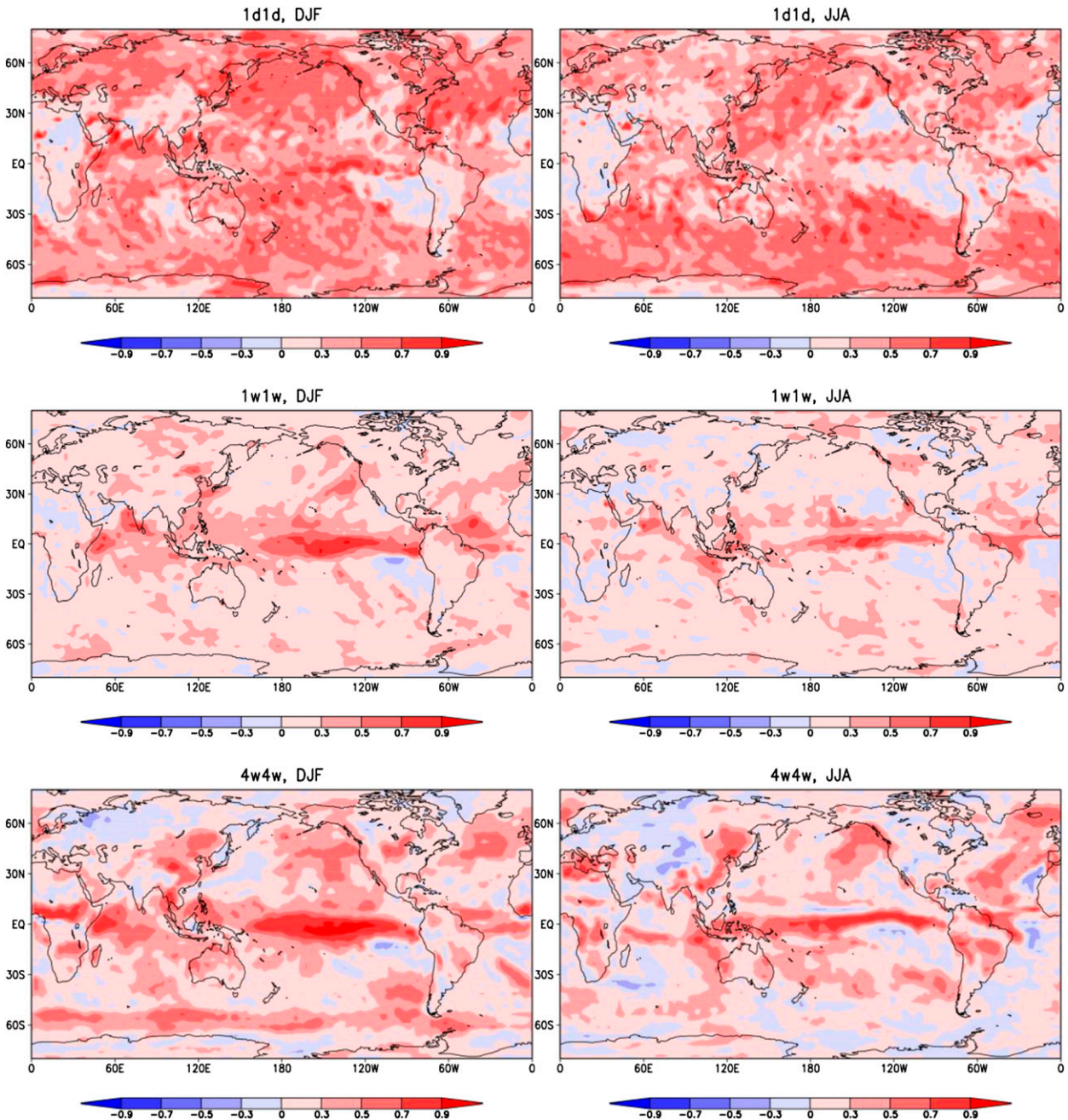


FIG. 2. Maps of CORt for model forecasts at (top) 1d1d, (middle) 1w1w, and (bottom) 4w4w, for (left) DJF and (right) JJA.

As the window/lead time increases, the CORt skill in the extratropics generally decreases until 2w2w, at which point it appears to approximately level off such that the 4w4w CORt is on average somewhat higher. Interestingly, at the 4w4w time scale, the CORt in DJF is on average higher than that in JJA in both hemispheres (and in the tropics).

Turning now to the deep tropics (i.e., within about 10° of the equator), the variation of skill with increasing

window/lead times is much different from that described above for the extratropics. Indeed, Fig. 3 nicely shows how the skill remains remarkably constant with increasing window/lead time in the tropics. In fact, the skill increases somewhat with window/lead time during DJF.

Another way to look at the variation of skill in the tropics versus extratropics with increasing window/lead time is presented in Fig. 4. In this figure we space the

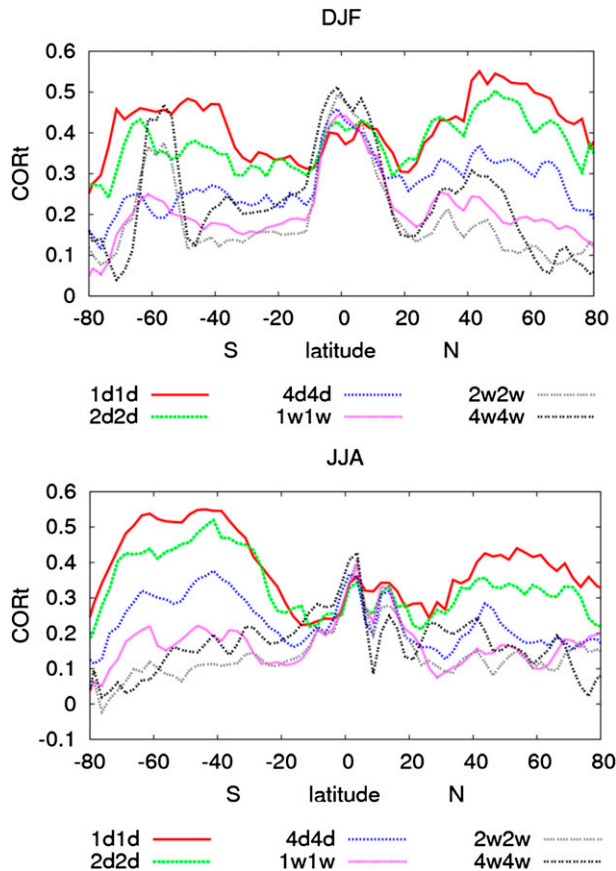


FIG. 3. Zonally averaged CORt for model forecasts at different time window/lead combinations for (top) DJF and (bottom) JJA.

time scales along the x axis according to their logarithm. We can now see more clearly that for the extratropics in both hemispheres there tends to be a minimum in CORt for the 2w2w time scale in all latitude bands in both seasons, except for the 70°–50°S band in DJF (which has its minimum at 1w1w). This indicates that the second half of the first month (or equivalently weeks 3 and 4 together) are the most unpredictable when evaluated this way. In contrast, the tropical latitudes show very little variation of CORt with time scale.

b. CORa—Correlation with the ensemble mean anomalies

As we described in section 2c, CORt may be influenced by the ability of the model forecasts to represent the observed seasonal cycle. If there is a strong seasonal cycle that is accurately represented by the model then CORt will be higher, but if the model gets the seasonal cycle reversed, CORt will be lower. CORa, on the other hand, removes the effects of the climatological seasonal cycle, and it is the more usual way of showing the correlation skill in seasonal prediction studies.

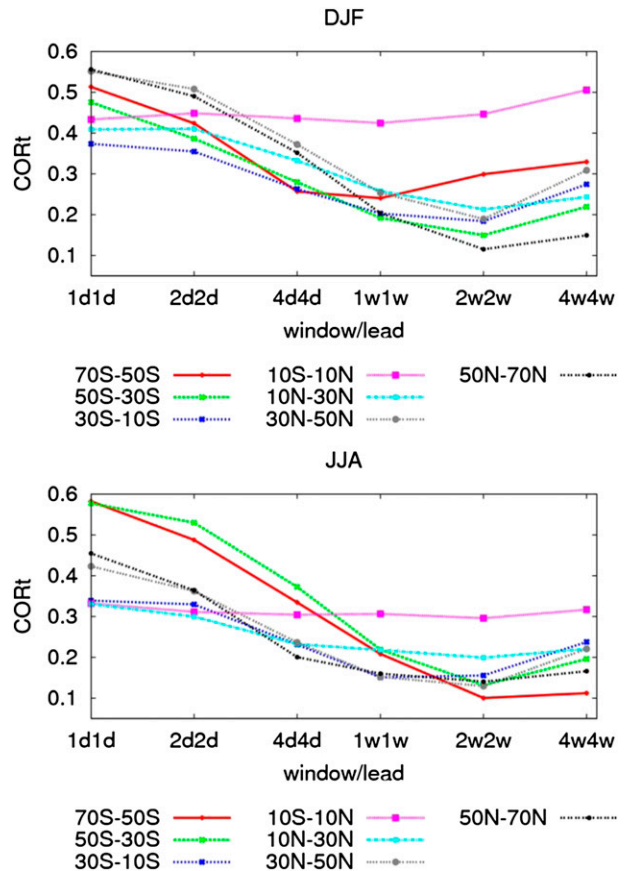


FIG. 4. Zonally averaged CORt over specified latitude ranges vs forecast window/lead times from 1d1d to 4w4w for (top) DJF and (bottom) JJA. Note that the spacing of the time scale is based on the logarithm of the lead time.

Comparing the CORa maps in Fig. 5 with the CORt maps in Fig. 2, the most obvious difference is generally lower values for CORa for 4w4w, but with very little change for 1d1d. The reason for this difference is because a longer averaging window gets a greater contribution to its total variance from the seasonal cycle. Removing the contribution from the seasonal cycle makes the model performance look worse for the longer averaging windows, especially in regions away from the ENSO-dominated tropical Pacific. The most obvious location for this apparently lower skill (when looking at CORa compared to CORt) is over the Southern Ocean around 55°S in DJF for 4w4w. As discussed in the previous section, we initially thought high CORt in this region may be associated with the southern annular mode (see also Lim et al. 2013). However, given the absence of this signal in CORa, it appears to instead be associated with an accurate representation of the seasonal cycle. This reduction in apparent skill when measured with CORa is similar to the effect described by Hamill and Juras (2006).

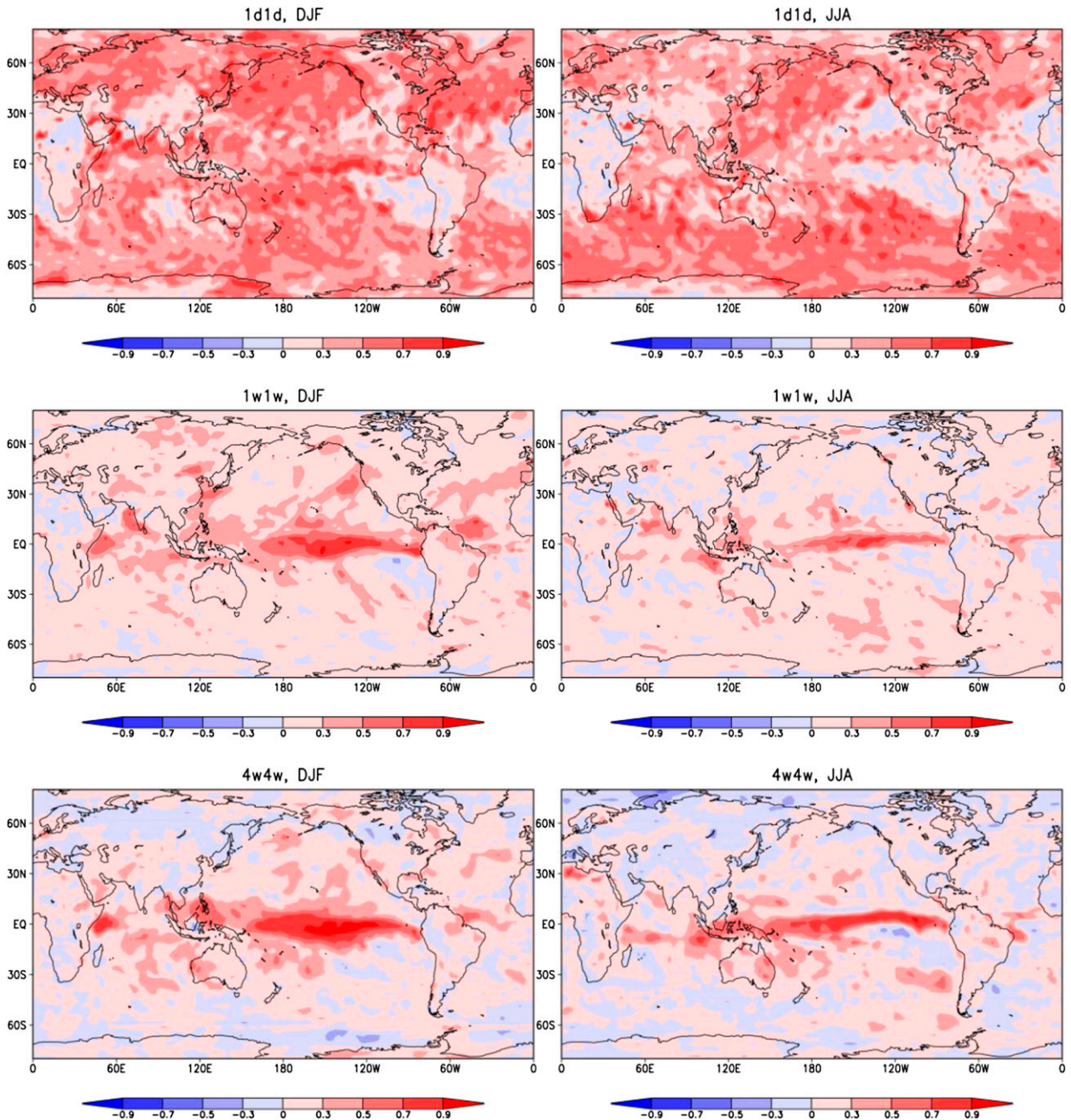


FIG. 5. As in Fig. 2, but for CORa.

Looking at the maps of Fig. 5 in more detail, there are a few regions of relatively high 1w1w and 4w4w skill that stand out. In the tropics for 4w4w, the ENSO-dominated signal in the equatorial Pacific extends westward into the islands of Indonesia and Papua New Guinea in JJA, and more toward the Philippines to the north in DJF, consistent with the empirical findings of McBride et al. (2003). In the Northern Hemisphere there are patches of relatively high 1w1w and 4w4w skill in the North Pacific

and western United States in DJF, consistent with our expectation from knowledge of the Pacific–North American (PNA) pattern (Kumar and Hoerling 1998). In the Southern Hemisphere there is 4w4w skill in the south Indian Ocean and Western Australia in DJF, and eastern Australia in JJA. The latter is expected given the known influence of ENSO in Australia (McBride and Nicholls 1983). Importantly, the abovementioned regions have greater skill than what is achievable from persistence

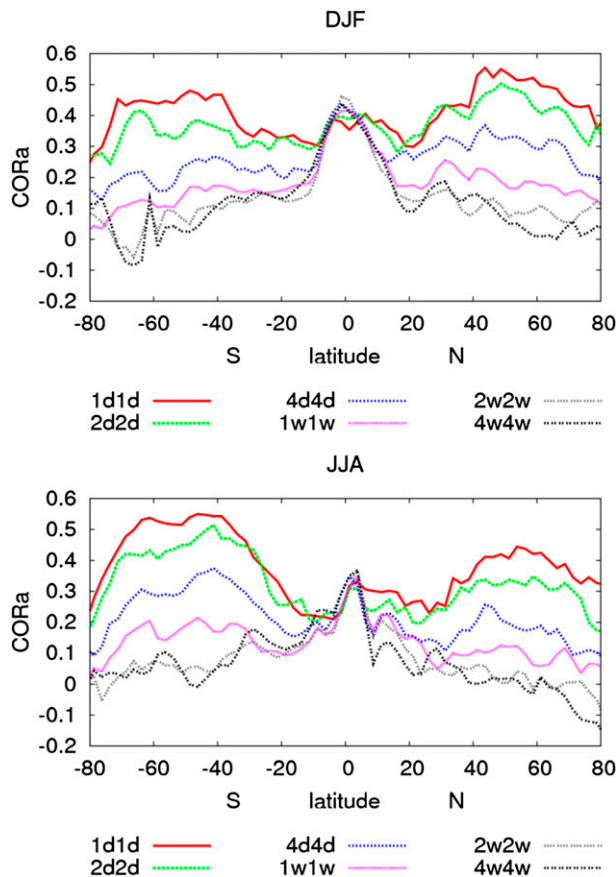


FIG. 6. As in Fig. 3, but for CORa.

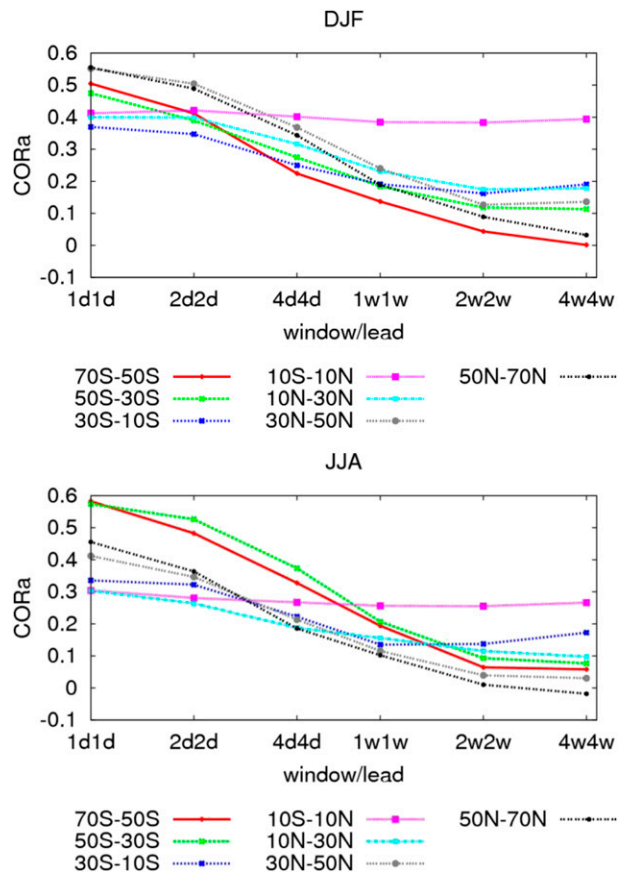


FIG. 7. As in Fig. 4, but for CORa.

(Simmonds and Hope 1997; see also the next subsection). Other interesting patches of high CORa are in northern Africa and the southeast Pacific for 4w4w in JJA, and the western equatorial Indian Ocean for both 1w1w and 4w4w in DJF.

When viewing the zonally averaged CORa values (as a function of latitude and window/lead time) in Fig. 6, a conclusion is reached that is very similar to what we obtained when looking at CORT. That is, that prediction skill decreases with window/lead time in the extratropics (outside of about 10° of the equator) but stays much the same in the tropics. Similarly, when looking at the alternative display of Fig. 7 we can see this variation with window/lead time clearly. We can also see at what point the skill in the tropics (when taken as a whole) begins to exceed that in the extratropics: in DJF it first occurs for 4d4d, and in JJA it first occurs for 1w1w.

c. Comparison with persistence

An important component of predictability is the prediction skill that can come from persistence, so it is of interest to see how these results compare. Figure 8 presents the correlation skill for persistence forecasts for

four different time scales (labeled as P1d1d, P4d4d, P2w2w, and P4w4w) and also shows the correlations for the 1d1d and 4w4w model forecasts for comparison. These persistence calculations used precipitation anomalies (i.e., CORa), and like for the model forecasts an averaging window equal in length to the lead time was used. For example, the P1d1d calculation uses the observed precipitation anomaly on the day before the model initial condition as the forecast, whereas the P4w4w calculation used the precipitation anomaly observed for the 4 weeks leading up to the initial condition.

In general, it can be seen in Fig. 8 (and with comparison to Fig. 6) that the zonally averaged CORa from the model tends to be higher than that for persistence, especially for the shorter time scales. Viewing maps of the persistence skill (not shown) confirms that this is generally the case for individual locations as well. Even at the longer 4w4w time scale, the model CORa exceeds or approximately equals the persistence skill (i.e., for P4w4w) for most latitudes equatorward of 50° . This is an encouraging result for the model because persistence has historically been difficult to beat at this range, as discussed by Vitart (2004).

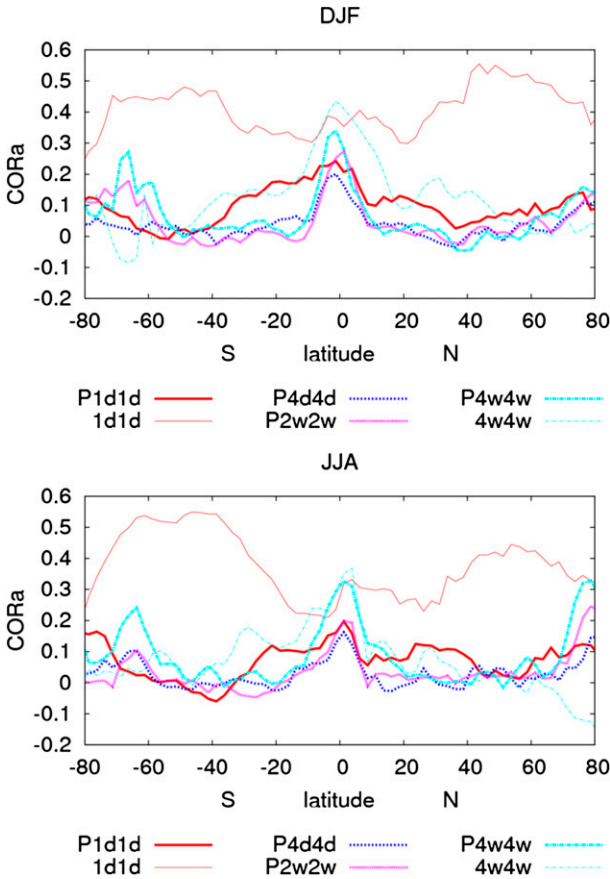


FIG. 8. As in Fig. 6, but showing zonally averaged CORa for persistence forecasts (labeled as P1d1d, P4d4d, P2w2w, P4w4w) and the model forecasts of 1d1d and 4w4w for comparison (as in Fig. 6).

Poleward of 50°, however, there are some notable peaks in P4w4w that are not replicated in the model forecasts, located around 70°S in DJF, and 65°S and 75°N in JJA. The maps of persistence skill (not shown) indicate that these peaks correspond to regions where large and persistent anomalies in sea ice cover occur

(Parkinson and Cavalieri 2008; Wheeler 2008), and an influence of sea ice on precipitation appears quite possible (Weatherly 2004). We note that POAMA-2 uses prescribed sea ice from a multiyear climatology, so is not able to reproduce this persistence skill, but it is something that may be improved by the incorporation of varying sea ice and sea ice anomalies in the initial condition in future versions.

d. Fixed time-averaging window of 1 day

Instead of increasing the time window at the same rate as the lead time, we now present the prediction skill as a function of lead time and latitude for a fixed time window of 1 day (see schematic of the new window and lead time definitions in Fig. 9) in Figs. 10 and 11. In this analysis we show COR_t only. As expected, the skill drops off much more rapidly (and monotonically) with lead time with a fixed window than it does when the window is increased. Importantly, however, the rate at which the COR_t skill drops is much less in the tropics than the extratropics providing the same general conclusion as before, that is, that there is a general transfer of skill from the extratropics to tropics as lead time is increased. The lead time at which the skill in the tropics tends to surpass the skill in the extratropics is shown to be at about 4 days in DJF and about 2 weeks in JJA. These values are respectively similar to and a little longer than the values found when the window length was varied as well (Figs. 4 and 7). Having a slightly longer estimate from this 1-day window calculation makes sense given the window/lead definitions used (cf. Figs. 1 and 9). For example, 1d2w is equivalent to day 15 whereas 2w2w is equivalent to days 15–28.

4. Conclusions

We have analyzed the skill with which an operational forecast system is able to predict precipitation over

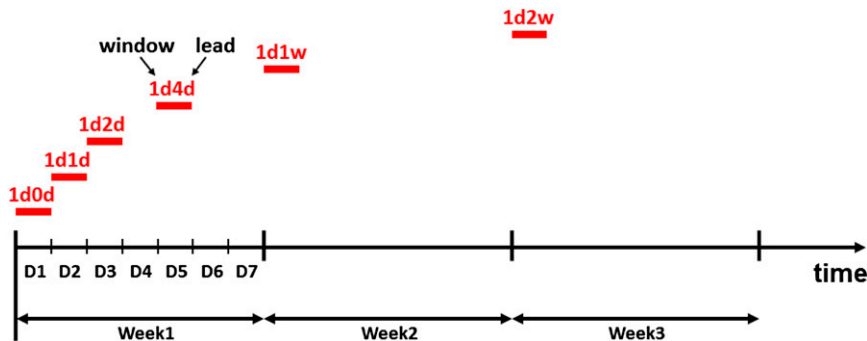


FIG. 9. As in Fig. 1, but showing the window and lead time definitions used for the calculations in Figs. 10 and 11 (i.e., with a fixed 1-day averaging window).

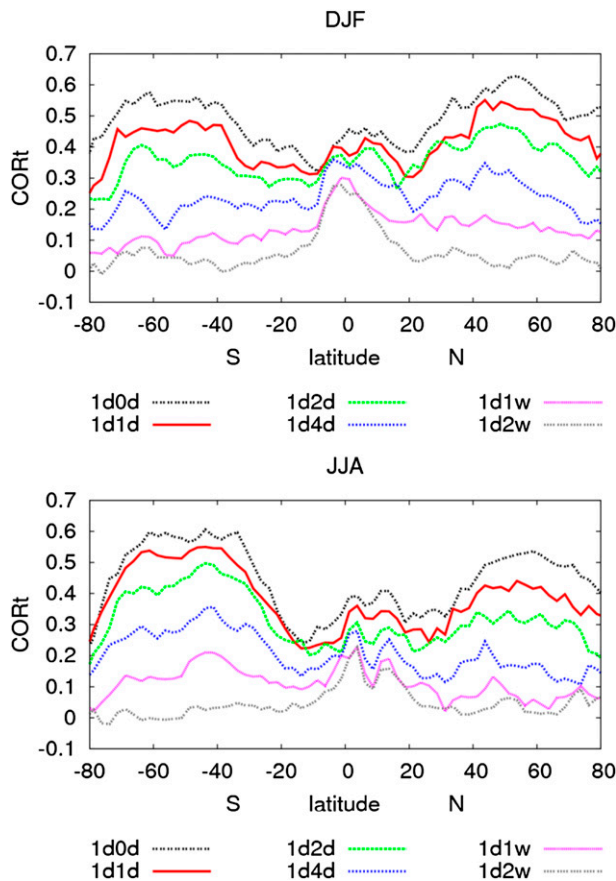


FIG. 10. As in Fig. 3, but using a constant 1-day averaging window as defined in Fig. 9.

a range of time scales from a day to months. We focus on the contrasting results obtained for different latitude bands and at different lead times. To emphasize the seamless transition between weather and climate, we have verified the model predictions after averaging both the forecasts and observed verification data over a time window equal to the forecast lead time. We performed skill calculations both on the total fields, and on anomalies computed by removing the appropriate climatological seasonal cycles from both the forecasts and the verification data. The skill measures we present are based on correlations between the forecasts and observations computed over time for each grid point. Calculations are made for the contrasting DJF and JJA seasons with ~ 13 years of model hindcasts.

At a lead time of 1 day, prediction skill is greatest in the extratropics around 40° – 60° latitude and lowest around 20° latitude and poleward of 70° , and has a secondary local maximum close to the equator. The extratropical skill at this short range is highest in the winter hemisphere, presumably due to the high day-to-day predictability of winter baroclinic weather systems and

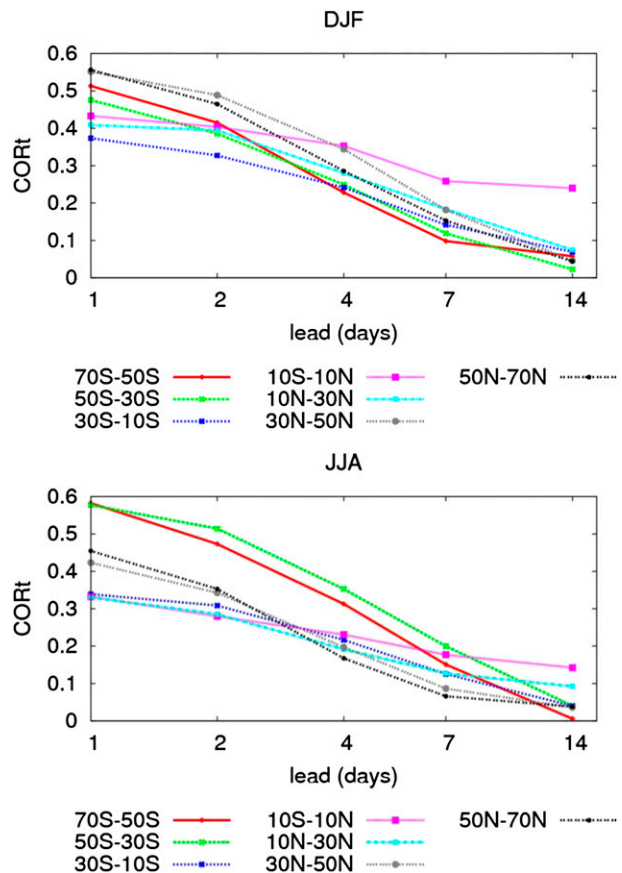


FIG. 11. As in Fig. 4, but using a constant 1-day averaging window as defined in Fig. 9. Note that in this figure a lead of 7 days is equivalent to 1d1w and 14 days to 1d2w.

associated fronts. In the summer hemisphere extratropics it is less, evidently due to the greater difficulty in predicting summer thunderstorms and the weaker summer baroclinic systems, but it still exceeds the 1-day prediction skill near the equator. The local equatorial maximum in the zonal mean is derived from the central and eastern Pacific, and thus appears (even at 1-day lead time) to be related to ENSO.

As both lead time and averaging window are simultaneously increased, the extratropical skill drops rapidly for short to medium lead times, while the equatorial maximum decreases much more slowly or stays approximately constant. The near-equatorial skill becomes equal to or greater than that at any other latitude band at around a 4-day time scale in DJF, and 1 week in JJA. At longer lead times, the extratropical correlations eventually flatten out or increase with lead time, but remain well below the near-equatorial values.

Importantly, the model prediction skill exceeds the skill of a persistence (of anomalies) forecast in most locations, especially at shorter lead times. For predictions

of a 4-week average at a lead of 4 weeks (i.e., 4w4w) the model skill remains better than persistence equatorward of about 50°, but is dramatically worse than persistence in a few locations near the sea ice edges in the Arctic and Antarctic.

To compare with our method of using an increasing window size with increasing forecast lead, we also computed the skill for varying lead times but a fixed averaging window of 1 day, a calculation more similar to the typical practice in weather forecast verification. The correlations at longer leads are smaller than those computed at the same leads but with longer averaging windows, as expected. However, perhaps more surprisingly, the slower decay of equatorial skill found with variable averaging windows is also found with the fixed 1-day averaging window, so that at sufficiently long leads, of between 4 and 14 days depending on season, the equatorial skill still exceeds that in the other latitude bands.

The broad picture we are left with is that on time scales of a few days or less, extratropical precipitation is more predictable than tropical, while at time scales of a week or longer, tropical precipitation, within about 10° of the equator, is more predictable than extratropical. This broad picture is remarkably robust to the details of how one does the calculations. While the absolute values of the skill depends on season, and on whether the averaging window is fixed or increasing with lead time, in all cases the near-equatorial zone eventually becomes more predictable than the extratropics at the lead times we consider.

This picture appears consistent with the view that extratropical predictability is mostly derived from the model's ability to simulate synoptic-scale atmospheric dynamics with rapid growth of initial state error (Lorenz 1969), while predictability in the deep tropics is mostly derived from the response of moist convection to slowly varying forcing such as from sea surface temperature (Charney and Shukla 1981) or the large-scale convergence of tropical waves (Hendon and Salby 1994). If there is any surprise here, it is that tropical influences can provide greater predictability than extratropical atmospheric dynamics at time scales as short as 4 days.

Finally, we advocate the usefulness of computing and displaying forecast skill globally across a large range of time scales as we have done here. Using precipitation as the verifying variable provides what we think is a fair comparison between the tropics and extratropics and the technique of increasing the averaging window size at the same rate as increasing the lead time provides the fairest comparison between different time scales. Recently, the need for seamless verification approaches has been promoted by Ebert et al. (2013), and while other

approaches do exist (DeSole and Tippett 2009), we feel the simplicity of our approach is an important advantage. Future work is planned to analyze other forecast systems (especially those employing a model with higher resolution) and to further investigate the skill as measured by verification measures that take into account the probabilistic nature of the ensemble.

Acknowledgments. We thank Beth Ebert and Huqiang Zhang for advice on verification methods, Harry Hendon for advice on predictability and POAMA, Frederic Vitart for discussions of tropical versus extratropical prediction skill, Eunpa Lim for suggestions on the title and discussions about the southern annular mode, and Phil Reid for advice on sea ice. Beth Ebert, John McBride, and Mike Tippett kindly read earlier versions of this paper. HZ receives funding support from the Australian Climate Change Science Program, AHS was supported by the U.S. Office of Naval Research (N00014-415 12-1-0911), and the Managing Climate Variability Program is acknowledged for their support of POAMA and its products.

REFERENCES

- Alves, O., and Coauthors, 2003: POAMA: Bureau of Meteorology operational coupled model seasonal forecast system. *Science for Drought—Proceedings of the National Drought Forum*, R. Stone and I. Partridge, Eds. Department of Primary Industries, 49–56.
- Baldwin, M. P., and T. J. Dunkerton, 1999: Propagation of the Arctic Oscillation from the stratosphere to the troposphere. *J. Geophys. Res.*, **104** (D24), 30 937–30 946.
- , D. B. Stephenson, D. W. J. Thompson, T. J. Dunkerton, A. J. Charlton, and A. O'Neill, 2003: Stratospheric memory and skill of extended-range weather forecasts. *Science*, **301**, 636–640.
- Bluestein, H. B., 1993: *Observations and Theory of Weather Systems*. Vol. II, *Synoptic–Dynamic Meteorology in Midlatitudes*, Oxford University Press, 606 pp.
- Boer, G. J., 1995: Analyzed and forecast large-scale tropical divergent flow. *Mon. Wea. Rev.*, **123**, 3539–3553.
- Bolvin, D. T., R. F. Adler, G. J. Huffman, E. J. Nelkin, and J. P. Poutiainen, 2009: Comparison of GPCP monthly and daily precipitation estimates with high-latitude gauge observations. *J. Appl. Meteor. Climatol.*, **48**, 1843–1857.
- Bretherton, C. S., P. N. Blossey, and M. Khairoutdinov, 2005: An energy-balance analysis of deep convective self-aggregation above uniform SST. *J. Atmos. Sci.*, **62**, 4273–4292.
- Charney, J. G., 1947: The dynamics of long waves in a baroclinic westerly current. *J. Meteor.*, **4**, 136–162.
- , 1963: A note on large-scale motions in the tropics. *J. Atmos. Sci.*, **20**, 607–609.
- , 1969: A further note on large-scale motions in the tropics. *J. Atmos. Sci.*, **26**, 182–185.
- , and J. Shukla, 1981: Predictability of monsoons. *Monsoon Dynamics*, J. Lighthill and R. P. Pearce, Eds., Cambridge University Press, 99–109.
- Cohen, J., and D. Entekhabi, 1999: Eurasian snow cover variability and Northern Hemisphere climate variability. *Geophys. Res. Lett.*, **26**, 345–348.

- Cottrill, A., and Coauthors, 2013: Seasonal forecasting in the Pacific using the coupled model POAMA-2. *Wea. Forecasting*, **28**, 668–680.
- DelSole, T., and M. K. Tippett, 2009: Average predictability time. Part II: Seamless diagnoses of predictability on multiple time scales. *J. Atmos. Sci.*, **66**, 1188–1204.
- DeWeaver, E., and S. Nigam, 2004: On the forcing of ENSO teleconnections by anomalous heating and cooling. *J. Climate*, **17**, 3225–3235.
- Ebert, E. E., 2001: Ability of a poor man's ensemble to predict the probability and distribution of precipitation. *Mon. Wea. Rev.*, **129**, 2461–2480.
- , U. Damrath, W. Ergen, and M. E. Baldwin, 2003: The WGENE assessment of short-term quantitative precipitation forecasts. *Bull. Amer. Meteor. Soc.*, **84**, 481–492.
- , and Coauthors, 2013: Progress and challenges in forecast verification. *Meteor. Appl.*, **20**, 130–139.
- Goddard, L., S. J. Mason, S. E. Zebiak, C. F. Ropelewski, R. Basher, and M. A. Cane, 2001: Current approaches to seasonal-to-interannual climate predictions. *Int. J. Climatol.*, **21**, 1111–1152.
- Hamill, T. M., and J. Juras, 2006: Measuring forecast skill: Is it real skill or is it the varying climatology? *Quart. J. Roy. Meteor. Soc.*, **132**, 2905–2923.
- Hendon, H. H., and M. L. Salby, 1994: The life cycle of the Madden-Julian oscillation. *J. Atmos. Sci.*, **51**, 2225–2237.
- Hoerling, M. P., and A. Kumar, 2002: Atmospheric response patterns associated with tropical forcing. *J. Climate*, **15**, 2184–2203.
- Holland, M. M., E. Blanchard-Wrigglesworth, J. Kay, and S. Vavrus, 2013: Initial-value predictability of Antarctic sea ice in the Community Climate System Model 3. *Geophys. Res. Lett.*, **40**, 2121–2124, doi:10.1002/grl.50410.
- Hudson, D., O. Alves, H. H. Hendon, and G. Wang, 2011: The impact of atmospheric initialisation on seasonal prediction of tropical Pacific SST. *Climate Dyn.*, **36**, 1155–1171.
- , A. G. Marshall, Y. Yin, O. Alves, and H. H. Hendon, 2013: Improving intraseasonal prediction with a new ensemble generation strategy. *Mon. Wea. Rev.*, **141**, 4429–4449.
- Huffman, G. J., R. F. Adler, M. Morrissey, D. T. Bolvin, S. Curtis, R. Joyce, B. McGavock, and J. Susskind, 2001: Global precipitation at one-degree daily resolution from multi-satellite observations. *J. Hydrometeorol.*, **2**, 36–50.
- Kiladis, G. N., and H. F. Diaz, 1989: Global climatic anomalies associated with extremes in the Southern Oscillation. *J. Climate*, **2**, 1069–1090.
- Koster, R. D., and M. J. Suarez, 2003: Impact of land surface initialization on seasonal precipitation and temperature prediction. *J. Hydrometeorol.*, **4**, 408–423.
- Kumar, A., and M. P. Hoerling, 1998: Annual cycle of Pacific–North American seasonal predictability associated with different phases of ENSO. *J. Climate*, **11**, 3295–3308.
- L'Heureux, M. L., and D. W. J. Thompson, 2006: Observed relationships between the El Niño–Southern Oscillation and the extratropical zonal-mean circulation. *J. Climate*, **19**, 276–287.
- Lim, E.-P., H. H. Hendon, O. Alves, Y. Yin, M. Zhao, G. Wang, D. Hudson, and G. Liu, 2009: Impact of SST bias correction on prediction of ENSO and Australian winter rainfall. *CAWCR Research Letters*, No. 3, BoM, Melbourne, Victoria, Australia, 22–29. [Available online at <http://www.cawcr.gov.au/publications/researchletters.php>.]
- , —, —, —, G. Wang, D. Hudson, M. Zhao, and L. Shi, 2010: Dynamical seasonal prediction of tropical Indo-Pacific SST and Australian rainfall with improved ocean initial conditions. CAWCR Tech. Rep. 32, 26 pp. [Available online at <http://cawcr.gov.au/publications/technicalreports.php>.]
- , —, and H. Rashid, 2013: Seasonal predictability of the southern annular mode due to its association with ENSO. *J. Climate*, **26**, 8037–8054.
- Lorenz, E. N., 1955: Available potential energy and the maintenance of the general circulation. *Tellus*, **2**, 157–167.
- , 1969: Three approaches to atmospheric predictability. *Bull. Amer. Meteor. Soc.*, **50**, 345–349.
- Mapes, B. E., 1993: Gregarious tropical convection. *J. Atmos. Sci.*, **50**, 2026–2037.
- Marshall, A. G., D. Hudson, M. C. Wheeler, H. H. Hendon, and O. Alves, 2011: Assessing the simulation and prediction of rainfall associated with the MJO in the POAMA seasonal forecast system. *Climate Dyn.*, **37**, 2129–2141, doi:10.1007/s00382-010-0948-2.
- , —, —, —, and —, 2012: Simulation and prediction of the southern annular mode and its influence on Australian intra-seasonal climate in POAMA. *Climate Dyn.*, **38**, 2483–2502, doi:10.1007/s00382-011-1140-z.
- , —, H. H. Hendon, M. J. Pook, O. Alves, and M. C. Wheeler, 2013: Simulation and prediction of blocking in the Australian region and its influence on intra-seasonal rainfall in POAMA-2. *Climate Dyn.*, doi:10.1007/s00382-013-1974-7, in press.
- McBride, J. L., and N. Nicholls, 1983: Seasonal relationships between Australian rainfall and the Southern Oscillation. *Mon. Wea. Rev.*, **111**, 1998–2004.
- , M. R. Haylock, and N. Nicholls, 2003: Relationships between the Maritime Continent heat source and the El Niño–Southern Oscillation phenomenon. *J. Climate*, **16**, 2905–2914.
- Parkinson, C. L., and D. J. Cavalieri, 2008: Arctic sea ice variability and trends, 1979–2006. *J. Geophys. Res.*, **113**, C07003, doi:10.1029/2007JC004558.
- Polvani, L. M., and P. J. Kushner, 2002: Tropospheric response to stratospheric perturbations in a relatively simple general circulation model. *Geophys. Res. Lett.*, **29** (7), doi:10.1029/2001GL014284.
- Roff, G., D. W. J. Thompson, and H. H. Hendon, 2011: Does increasing model stratospheric resolution improve extended-range forecast skill? *Geophys. Res. Lett.*, **38**, L05809, doi:10.1029/2010GL046515.
- Schumacher, C., and R. A. Houze Jr., 2003: Stratiform rain in the tropics as seen by the TRMM Precipitation Radar. *J. Climate*, **16**, 1739–1756.
- Shukla, J., 1989: Tropical forecasting: Predictability perspective. *Aust. Meteor. Mag.*, **37**, 141–153.
- Simmonds, I., and P. Hope, 1997: Persistence characteristics of Australian rainfall anomalies. *Int. J. Climatol.*, **17**, 597–613.
- Sobel, A. H., 2012: Tropical weather. *Nat. Educ. Knowl.*, **3** (12), 2.
- Stockdale, T. N., 1997: Coupled ocean–atmosphere forecasts in the presence of climate drift. *Mon. Wea. Rev.*, **125**, 809–818.
- Uppala, S. M., and Coauthors, 2005: The ERA-40 Re-Analysis. *Quart. J. Roy. Meteor. Soc.*, **131**, 2961–3012.
- Vitart, F., 2004: Monthly forecasting at ECMWF. *Mon. Wea. Rev.*, **132**, 2761–2779.
- Waliser, D., and Coauthors, 2006: The Experimental MJO Prediction Project. *Bull. Amer. Meteor. Soc.*, **87**, 425–431.
- Wang, G., D. Hudson, Y. Yin, O. Alves, H. Hendon, S. Langford, G. Liu, and F. Tseitkin, 2011: POAMA-2

- SST skill assessment and beyond. *CAWCR Research Letters*, No. 6, BoM, Melbourne, Victoria, Australia, 40–46. [Available online at <http://www.cawcr.gov.au/publications/researchletters.php>.]
- Weare, B. C., 1987: Relationships between monthly precipitation and SST variations in the tropical Pacific region. *Mon. Wea. Rev.*, **115**, 2687–2698.
- Weatherly, J. W., 2004: Sensitivity of Antarctic precipitation to sea ice concentrations in a general circulation model. *J. Climate*, **17**, 3214–3223.
- Wheeler, M. C., 2008: Seasonal climate summary Southern Hemisphere (summer 2007-08): Mature La Niña, an active MJO, strongly positive SAM, and highly anomalous sea-ice. *Aust. Meteor. Mag.*, **57**, 379–393.
- , and G. N. Kiladis, 1999: Convectively coupled equatorial waves: Analysis of clouds and temperature in the wave-number–frequency domain. *J. Atmos. Sci.*, **56**, 374–399.
- Yin, Y., O. Alves, and P. R. Oke, 2011: An ensemble ocean data assimilation system for seasonal prediction. *Mon. Wea. Rev.*, **139**, 786–808.