Skillful All-Season S2S Prediction of U.S. Precipitation Using the MJO and QBO

KYLE M. NARDI
Department of Atmospheric Science, Colorado State University, Fort Collins, Colorado

CORY F. BAGGETT
National Oceanic and Atmospheric Administration, NWS/NCEP/CPC, College Park, and Innovim, LLC, Greenbelt, Maryland

ELIZABETH A. BARNES AND ERIC D. MALONEY
Department of Atmospheric Science, Colorado State University, Fort Collins, Colorado

DANIEL S. HARNOS
National Oceanic and Atmospheric Administration, NWS/NCEP/CPC, College Park, Maryland

LAURA M. CIASTO
National Oceanic and Atmospheric Administration, NWS/NCEP/CPC, College Park, and Innovim, LLC, Greenbelt, Maryland

(Manuscript received 22 November 2019, in final form 1 August 2020)

ABSTRACT: Although useful at short and medium ranges, current dynamical models provide little additional skill for precipitation forecasts beyond week 2 (14 days). However, recent studies have demonstrated that downstream forcing by the Madden–Julian oscillation (MJO) and quasi-biennial oscillation (QBO) influences subseasonal variability, and predictability, of sensible weather across North America. Building on prior studies evaluating the influence of the MJO and QBO on the subseasonal prediction of North American weather, we apply an empirical model that uses the MJO and QBO as predictors to forecast anomalous (i.e., categorical above- or below-normal) pentadal precipitation at weeks 3–6 (15–42 days). A novel aspect of our study is the application and evaluation of the model for subseasonal prediction of precipitation across the entire contiguous United States and Alaska during all seasons. In almost all regions and seasons, the model provides “skillful forecasts of opportunity” for 20%–50% of all forecasts valid weeks 3–6. We also find that this model skill is correlated with historical responses of precipitation, and related synoptic quantities, to the MJO and QBO. Finally, we show that the inclusion of the QBO as a predictor increases the frequency of skillful forecasts of opportunity over most of the contiguous United States and Alaska during all seasons. These findings will provide guidance to forecasters regarding the utility of the MJO and QBO for subseasonal precipitation outlooks.

KEYWORDS: Teleconnections; Precipitation; Climate prediction; Forecast verification/skill; Numerical weather prediction/forecasting

1. Introduction

The impacts of precipitation events on a variety of stakeholders are well documented. Intense precipitation associated with landfalling atmospheric rivers (ARs) has triggered high streamflows and flash flooding in California (Dettinger et al. 2011; Neiman et al. 2002; Ralph et al. 2006, 2010, 2013; Ralph and Dettinger 2012; Waliser and Guan 2017), the U.S. Pacific Northwest (Neiman et al. 2008b,a, 2011), the U.S. Southwest (Ralph and Galarneau 2017; Rivera et al. 2014), the U.S. Midwest (Dirmeyer and Kinter 2009, 2010), and the U.S. Southeast (Lackmann 2013; Lavers and Villarini 2013; Mahoney et al. 2016; Moore et al. 2012, 2015). In addition, intense precipitation from convective storms (Dirmeyer and Kinter 2010; Hitchens et al. 2010; Lackmann 2013; Lapenta et al. 1995; Schumacher and Johnson 2006, 2008) and tropical cyclones (Atallah et al. 2007; Nogueira and Keim 2010; Prat and Nelson 2013; Wood and Ritchie 2013) have documented links to flooding.

Aside from short-term public safety hazards such as flooding, precipitation events have long-term impacts through contributions to the regional water supply. For instance, wintertime precipitation in the western United States is crucial for maintenance and enhancement of the snowpack in mountainous regions (Guan et al. 2010). In turn, this enhanced snowpack...

Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/WAF-D-19-0232.s1.

* Current affiliation: Department of Meteorology and Atmospheric Science, The Pennsylvania State University, University Park, Pennsylvania.

** Corresponding author: Kyle M. Nardi, kmn182@psu.edu

DOI: 10.1175/WAF-D-19-0232.1

© 2020 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).
bolsters the regional water supply and helps alleviate or prevent drought at lower elevations (Luo et al. 2017; Margulis et al. 2016; Mote et al. 2016). Precipitation during the warm season is also critical for agricultural interests elsewhere in the United States, where the amount of annual precipitation influences crop yields (Brown et al. 1986; Schiraldi and Roundy 2017).

Given the range of impacts from precipitation events of all magnitudes, it is vital to provide skillful precipitation forecasts valid as far into the future as possible. This is one of the principal missions of the National Oceanic and Atmospheric Administration (NOAA) National Centers for Environmental Prediction’s (NCEP) Climate Prediction Center (CPC). The CPC currently offers subseasonal precipitation outlooks across the contiguous United States and Alaska at leads of 3–4 weeks, known as their week 3–4 precipitation outlook. To create the outlook, the CPC employs a suite of forecast tools that includes both dynamical and empirical models. However, the dynamical models, the empirical models, and the CPC’s official week 3–4 precipitation outlook only offer modest skill when verified against observations. Therefore, the week 3–4 precipitation outlook remains experimental, while new forecasting tools, such as the incorporation of stratospheric predictors, are being researched and developed to improve the outlook.

Consistent with the CPC’s forecast tools, Pan et al. (2019) found that, while potential methods of extending this time frame may exist, dynamical model hindcasts from the Subseasonal-to-Seasonal (S2S) Prediction Project do not provide additional skill for wintertime precipitation forecasts in the western United States beyond 14 days. Meanwhile, Sharma et al. (2017) demonstrated that precipitation hindcasts for the eastern United States from the NCEP Global Ensemble Forecast System Reforecast (GEFSRv2) provide little skill beyond 7 days. These predictive limitations of dynamical models are consistent with those found for specific drivers of precipitation such as ARs (Mundhenk et al. 2018; Nardi et al. 2018; Nayak et al. 2014; Wick et al. 2013), blocking associated with Rossby wave patterns (Martínez-Alvarado et al. 2018; Matsueda et al. 2011; Vigaud et al. 2018), and tropical cyclones (Zhang et al. 2014; Zhong et al. 2018).

Although dynamical models struggle to forecast precipitation beyond 2 weeks, empirical models based on sources of climate variability such as the Madden–Julian oscillation (MJO) and the stratospheric quasi-biennial oscillation (QBO) have the potential to add predictive skill beyond the limits of dynamical models. The MJO, a 30–90-day cycle of deep convection over the tropical Pacific (Zhang 2005), modulates downstream Rossby wave patterns (Gibson et al. 2020; Henderson et al. 2016; Riddle et al. 2013; Tseng et al. 2018) and, in turn, precipitation-inducing phenomena such as ARs (DeFlorio et al. 2019a; Mundhenk et al. 2016), tropical cyclones (Maloney and Hartmann 2000a,b; Slade and Maloney 2013), severe convection (Baggett et al. 2018), and the North American Monsoon (Lorenz and Hartmann 2006). This modulation by the MJO extends to observed precipitation over North America (Barrett et al. 2015; Baxter et al. 2014; Dong et al. 2018; Jones 2000; Jones and Carvalho 2012; Lin et al. 2010; Moon et al. 2012; Zhou et al. 2012). The QBO manifests as a 2–3-yr periodic cycle of downward propagating easterly and westerly wind regimes in the equatorial stratosphere (Baldwin et al. 2001). Like the MJO, the QBO also has demonstrated synoptic teleconnection patterns, specifically a modulation of the storm track over the North Pacific (Wang et al. 2018a).

Recent studies have demonstrated dynamical links between the MJO and QBO. Son et al. (2017) showed that the QBO accounts for about 40% of the variation in MJO activity during the boreal winter, while Yoo and Son (2016) found that the robust relationship between the MJO and QBO is most notable under easterly QBO conditions, with the strongest relationship manifesting during the boreal winter and fall. As a result of this connection, the MJO has been found to be more predictable during the easterly phase of the QBO (Marshall et al. 2017). Hendon and Abhik (2018) explained this relationship using static stability arguments, arguing that QBO-induced temperature anomalies, and a related decrease in stability, near the tropopause are more closely aligned with MJO-related convection under easterly QBO conditions. This decrease in stability makes the MJO more sensitive to stratospheric cooling associated with the easterly QBO. Meanwhile, Zhang and Zhang (2018) offered an alternative explanation, showing that under easterly QBO conditions more MJO events form in the Indian Ocean, which mitigates the barrier effects of the Maritime Continent, thus promoting longer MJO event duration and more MJO days overall. Klotzbach et al. (2019) demonstrated that the relationship between the MJO and QBO has been bolstered by increased temperature anomalies in the upper troposphere and lower stratosphere that manifested in the early 1980s, thus making the connection a relatively new phenomenon.

As noted above, the physical mechanisms via which the QBO impacts the MJO and its teleconnections are still a subject of ongoing research, as the relative impacts of the QBO on the background state, which influences the MJO and its teleconnections, are not fully understood. For the purposes of this study, however, we are interested in observed effects that the MJO and QBO produce with respect to midlatitude teleconnections. For example, Feng and Lin (2019) found that the QBO’s modulation of the subtropical westerly jet impacts the propagation of MJO-induced Rossby waves over the North Pacific. The combined MJO–QBO state has also been shown to modulate downstream sensible weather, specifically the frequency of precipitation-inducing events over North America (Baggett et al. 2017; Wang et al. 2018b).

As a result of this dynamical relationship between the MJO and QBO, prior studies have exploited the MJO, or the combined MJO and QBO state, to improve predictions of precipitation. For instance, the MJO alone has been shown to enhance the skill of dynamical (DeFlorio et al. 2019a,b; Pan et al. 2019) and simple decision (Jones et al. 2011) models in forecasting precipitation or precipitation-inducing events (ARs). Based on the influence of the MJO and QBO on ARs in the western United States, Baggett et al. (2017) and Mundhenk et al. (2018) demonstrated that an empirical model based on the current state of the MJO and QBO
provides skillful forecasts of AR activity in the western United States at leads of 2–5 weeks. Baggett et al. (2018) also found that a similar empirical model based on the MJO provides skill in forecasting severe convection over the U.S. Southeast and Great Plains at leads of 3–5 weeks. Building on these verification studies, as well as the observed modulation of sensible weather and the dynamical links between the two indices, we apply an empirical model based on the MJO and QBO to predict anomalous (categorical above- or below-normal) pentadal precipitation at S2S leads, which we define as weeks 3–6 (15–42 days). While most studies restrict their analyses of the MJO or QBO influence to a particular region or season, we run the empirical model using the MJO and QBO during all seasons across the contiguous United States and Alaska. The results of our analysis will directly inform the CPC about the potential utility of incorporating the QBO as a predictor within its empirical models in order to improve its week 3–4 precipitation outlook.

We note that El Niño–Southern Oscillation (ENSO) provides another potential source of predictability. ENSO manifests itself as a 3–6-yr cycle of ocean–atmosphere interactions associated with anomalous sea surface temperatures in the equatorial Pacific Ocean (Trenberth 1997). ENSO is known to influence the global circulation and resultant seasonal precipitation patterns across North America (Chiodi and Harrison 2015; Higgins et al. 2000; Ropelewski and Halpert 1986, 1987). This observed influence on precipitation patterns translates to enhanced model skill in forecasting seasonal precipitation in North America (Chen et al. 2017; Wood et al. 2002). Even so, we choose not to add ENSO as an additional predictor for two principal reasons: 1) analysis of seasonal data parsed by MJO, QBO, and ENSO conditions is difficult due to sample size considerations and 2) prior studies have found that the seasonal impacts of ENSO do not strongly influence the sub-seasonal relationship between the MJO and QBO (Baggett et al. 2017; Son et al. 2017; Yoo and Son 2016).

In this study, we first summarize the utility of the empirical model in predicting precipitation based on the number of “skillful forecasts of opportunity,” which we will define later as forecasts, made during certain MJO and QBO phase combinations, that are significantly better than a random forecast. Then, we analyze the influence of the QBO as a predictor by comparing results to an alternative version of the model with the QBO removed (i.e., MJO only).

2. Methods

a. Empirical model

To produce precipitation forecasts, we employ the empirical model introduced by Mundhenk et al. (2018) and later modified by Baggett et al. (2018). This model uses the current state of the MJO and QBO to create categorical forecasts of above- or below-normal pentadal precipitation for all lead times through week 6. Here, “pentadal” refers to overlapping 5-day forward running means. Henceforth, we reference these pentads by the lead time of the starting point of their period. For example, day 6 refers to the pentad spanning the period 6–10 days after initialization, while day 7 refers to the pentad spanning the period 7–11 days after initialization. Throughout this study, we define weeks such that, for example, week 6 refers to pentads starting 36–42 days after initialization. Unlike prior precipitation forecast verification studies, we extend our analysis through week 6, so we opt to examine 5-day forecast windows since it is currently unrealistic to expect a model to accurately predict the occurrence of precipitation on a given day 5–6 weeks in advance. We note that the model is capable of predicting precipitation anomalies using longer averaging windows. However, analysis (not shown) of model performance using slightly longer 7-day averaging windows indicates that overall model skill is quite similar across commonly used averaging windows.

The model makes a binary precipitation anomaly forecast, for either an “above-normal” or “below-normal” anomaly, based on the historical response of the forecast variable to the current state of the MJO and QBO. To accomplish this, the model compares the conditional distribution of precipitation anomalies to the full historical distribution of precipitation anomalies for that season for the years 1979–2017. The conditional distributions are based on season, lead time, and MJO and QBO characteristics at initialization. We calculate anomalies by removing the seasonal cycle, which we define as the mean and first two harmonics of the particular time series. Since we wish to assess pentadal precipitation anomalies, both the conditional and full historical distributions are smoothed using 5-day forward averaging. If the median of the conditional distribution is greater than the median of the full distribution, then the model forecasts above-normal precipitation. If the median of the conditional distribution is less than the median of the full distribution, then the model forecasts below-normal precipitation. We use medians rather than means in order to better account for climatological precipitation distributions that are not normal.

We apply the model for initializations during all moving three-month seasons such as January–March (JFM), February–April (FMA), etc. However, since the model makes forecasts valid through week 6, it is often the case that a forecast initialized at the end of a given three-month season will be valid for days well outside of the season. For example, a 36-day forecast initialized on 30 November (at the end of the SON season) would be valid for the period spanning from 5 to 9 January. In this situation, it does not make sense to forecast precipitation responses for this valid time period using climatology from 1 September to 30 November. Therefore, we follow Baggett et al. (2018) and allow the seasonal distributions used to make forecasts to shift as a function of lead time such that the start and end dates of each season are shifted by the number of lead days. For example, JFM for 14-day forecasts implies distributions drawn from dates spanning 15 January–14 April, while JFM for 21-day forecasts implies dates spanning 22 January–21 April. By implementing this method, we guarantee that all forecasts will use anomaly distributions that include all valid dates.

b. Predictors and predictands

To create the conditional and full distributions, we examine historical precipitation (the predictand) and MJO and QBO
(the predictors) data spanning from 1 January 1979 to 31 December 2017. In characterizing the state of the MJO, we use the real-time multivariate MJO (RMM) index (Wheeler and Hendon 2004). The RMM index assigns a unique MJO phase (1–8) and amplitude for each day. Similar to Mundhenk et al. (2018) and Baggett et al. (2018), we employ an MJO amplitude threshold at initialization such that forecasts are only made under “active” MJO conditions. Specifically, an MJO amplitude greater than or equal to 1.0 is required for the model to make a prediction on a given day. Following this protocol, forecasts are initialized on all active MJO days during the time period, thus accounting for approximately 61% of all days from 1979 to 2017.

To characterize the state of the QBO, we use the reanalysis-based CPC QBO index of monthly averaged, standardized anomalies of zonal wind at 50 hPa, calculated over the equator. The QBO index assigns a monthly standardized value based on the prior month’s QBO conditions. In real time, the QBO index for the prior month only becomes available around the beginning of each month. Therefore, we shift the index forward one month in time in order to simulate the QBO index as a real-time predictor. The empirical forecast model is conditioned on active MJO days with easterly QBO conditions and on active MJO days with westerly QBO conditions, where easterly QBO conditions occur if the QBO index is less than zero and westerly QBO conditions occur otherwise. For the time period 1979–2017, we find that easterly QBO months occur approximately 44% of the time, while westerly QBO months occur approximately 56% of the time. We do not apply a QBO intensity threshold, as we do with the MJO. Using a nonzero QBO index threshold to define easterly and westerly periods would reduce the number of forecast initializations available for our analysis and would also preclude a direct comparison of the empirical model with and without the QBO included. Furthermore, the QBO is largely a continuous phenomenon, unlike the MJO, with active anomalies always existing somewhere in the equatorial stratosphere.

To characterize precipitation, we use daily total precipitation from the CPC Global Unified Gauge-Based Analysis of Daily Precipitation (Chen et al. 2008; Xie et al. 2007). This dataset has a spatial resolution of 0.5° × 0.5°. To remain consistent with the MJO and QBO datasets, where easterly QBO conditions occur if the QBO index is less than zero and westerly QBO conditions occur otherwise. For the time period 1979–2017, we find that easterly QBO months occur approximately 44% of the time, while westerly QBO months occur approximately 56% of the time. We do not apply a QBO intensity threshold, as we do with the MJO. Using a nonzero QBO index threshold to define easterly and westerly periods would reduce the number of forecast initializations available for our analysis and would also preclude a direct comparison of the empirical model with and without the QBO included. Furthermore, the QBO is largely a continuous phenomenon, unlike the MJO, with active anomalies always existing somewhere in the equatorial stratosphere.

To characterize precipitation, we use daily total precipitation from the CPC Global Unified Gauge-Based Analysis of Daily Precipitation (Chen et al. 2008; Xie et al. 2007). This dataset has a spatial resolution of 0.5° × 0.5°. To remain consistent with the MJO and QBO datasets, we subset the precipitation data to cover the time period from 1 January 1979 to 31 December 2017. We note that, while global in nature, the dataset only covers land areas. That is, daily total precipitation is averaged over all grid cells within the contiguous United States and Alaska. We apply regions of this size because they are large enough to capture the typical spatial scale of midlatitude cyclones that often have embedded finer-scale features such as fronts, vorticity maxima, and orographic lift. These regions are also small enough to reduce the occurrence of forecasting total precipitation over domains with vastly different characteristics. Moreover, this region size is consistent with a similar analysis in Baggett et al. (2018). However, additional analysis (not shown) indicates that overall model skill is similar for larger forecast regions.

Prior to calculating anomalies in each region, full daily total precipitation is averaged over all grid cells within the region. The location of each region is based on the latitude and longitude coordinates of the central grid cell, with the central grid cells spaced every 2.5° from one another. As mentioned above, based on the precipitation dataset used here, these regions cover a maximum of 225 precipitation grid cells (0.5° × 0.5° in size), although the true number of averaged grid cells varies by region since the precipitation dataset only includes grid cells over land. That is, daily total precipitation in coastal regions is only calculated using land grid cells. While the high spatial resolution of the precipitation dataset allows for additional overlapping forecast regions, we limit the number of overlapping regions analyzed to reduce computational costs.

To maintain consistency with the precipitation data, we analyze anomalies of 500-hPa geopotential height and IVT over the same 262 overlapping 7.5° × 7.5° regions. However, since the gridded datasets of these two quantities have a lower resolution (i.e., 2.5° × 2.5°), these forecast regions have only nine grid cells. As with precipitation, we average geopotential height and IVT over the grid cells within each region prior to calculating anomalies.

d. Model skill metric

We assess the model’s skill using a leave-one-year-out cross-validation approach (Baggett et al. 2018; Johnson et al. 2014; Mundhenk et al. 2018). Using this approach, we make forecasts for a particular year (specifically, a 3-month season during the year) based on seasonal data from the rest of the historical period with that year left out. We then compare the forecasts for the left-out year to what was actually observed. Since we condition on easterly and westerly QBO phases

\[ \text{IVT} = \sqrt{\left(\frac{1}{g} \int_{1000}^{500} qu \, dp \right)^2 + \left(\frac{1}{g} \int_{1000}^{500} qv \, dp \right)^2} \]
separately, we can initially assess the model’s skill under each QBO phase. In later analyses, we opt to combine verification results into one “all QBO” pool to isolate the influence of adding the QBO as a predictor.

Skill is defined using a variation of the Heidke skill score (HSS) with a baseline random forecast (Baggett et al. 2018; Mundhenk et al. 2018):

$$\text{HSS} = \left( \frac{C - E}{T - E} \right) \times 100,$$

where $C$ is the number of correct forecasts; $T$ is the total number of forecasts; and $E$ is the expected number of correct forecasts. Since the model uses the median of the full historical distribution to make forecasts of above- or below-normal precipitation, a random forecast is expected to be correct 50% of the time. This implies that $E$ is always equal to $T/2$.

Under this definition, HSS ranges from $-100$ (all incorrect) to 100 (all correct). An HSS greater than 0 implies that the model provides more skill than a random, equal-chances forecast. An HSS of 30 corresponds to a model that is correct 65 out of 100 forecasts made (roughly twice as many correct as incorrect). For each combination of MJO phase $P$ (1–8) and lead time $L$ (0–42 days), the HSS should be interpreted as the skill of a forecast for the 5-day period $L$ to $L + 4$ days after MJO phase $P$ when accounting for QBO phase.

e. Statistical significance

We assess statistical significance of model skill based on the following null hypothesis: MJO and QBO conditions at the time of forecast initialization do not provide a source of positive skill for precipitation forecasts. To test this hypothesis, we employ the random walks method described in DelSole and Tippett (2016). We determine the number of correct forecasts from the empirical model and compare to the number of correct forecasts expected from a model that randomly forecasts above- or below-normal precipitation. We then compare the difference ($dN$) in correct forecasts between the empirical model and the baseline random model to the $n$th percentile of a Gaussian distribution of random differences centered on 0. If the true $dN$ falls above the $n$th percentile, we conclude that the empirical model is significantly better than the random model at the $n$% confidence level. If the difference between the true empirical model and the random model is the addition of MJO and QBO conditions as predictors, which is the case in this study, then we reject the null hypothesis.

In addition, we assess statistical significance of composite anomaly responses to MJO and QBO conditions based on a slightly modified null hypothesis: MJO and QBO conditions on a given day do not influence subsequent pentadal anomalies. We test this hypothesis using a block bootstrapping approach (Baggett et al. 2018; Mundhenk et al. 2018) to better account for autocorrelation associated with the MJO, whose phases often occur in consecutive blocks of days (i.e., a set of days with the same active MJO phase). We assign each block of days in the time period a random MJO phase (1–8) and start date and calculate composite anomalies for each MJO phase and lead. We perform 1000 such iterations in order to produce a random distribution of composite anomalies for each phase and lead combination. We then compare the observed composite anomalies to the random distribution. If the observed anomaly falls outside of the two-sided $p$% confidence bounds, then we consider the anomaly statistically significant at the $p$% confidence level. We later apply the same block bootstrapping method to test the statistical significance of the percentiles of precipitation anomalies.

3. Results

a. Skill for all regions and seasons

We first examine the empirical model’s skill in predicting above- or below-normal pentadal precipitation anomalies for each combination of MJO phase and lead time during four nonoverlapping seasons: DJF (winter), MAM (spring), JJA (summer), and SON (fall). Figure 1 shows a good example of the model’s positive skill (HSS, see methodology) for precipitation forecasts, shown here for the U.S. Southeast under easterly and westerly QBO conditions. Here, we show results for all forecasts initialized during fall, so this plot would be relevant to a forecaster making a forecast on any day between 1 September and 30 November. In this scenario, the HSS values for a lead of 21 days denote skill for a forecast valid 21–25 days after initialization, and the seasonal distributions in this case range from 22 September to 21 December. We highlight positive HSS by whiting out those phase and lead combinations with negative skill. Phase and lead combinations with positive skill are then colored based on the model’s forecast: above-normal (blue) or below-normal (red) pentadal precipitation anomalies. We note that the same results are not expected for every region and season, so we provide similar HSS plots for all 262 regions in a supplemental online repository (see the appendix).

For both QBO phases in the U.S. Southeast during fall, numerous phase and lead combinations exhibit appreciable positive skill that is significantly better than a random model at 90% confidence (black dots). As seen in Fig. 1 and the online repository, skillful forecasts are generally well distributed by lead time through week 6. Also evident in Fig. 1 (and later figures) are the demonstrated differences in results between easterly and westerly QBO phases, which is consistent with studies such as Yoo and Son (2016) that showed a stronger modulating influence of MJO convection under easterly QBO conditions. We also emphasize here that these results use 5-day averaging windows and 7.5° × 7.5° regions. Though not appreciably affecting overall model skill, the application of a slightly longer 7-day averaging window or a larger forecasting domain may result in changes in skill for individual phase and lead combinations.

Nonetheless, a distinct propagating MJO signal (marked by stripes of high HSS values) appears in the U.S. Southeast during fall. As discussed in Baggett et al. (2018), such significant skill at long lead times can be explained by the propagating nature of the MJO and its teleconnections, whereby a teleconnection forced by one MJO phase may occur under a subsequent MJO phase even though the teleconnection was forced by the original MJO phase. In this way, the presence of
skill at longer leads (e.g., weeks 3–6) is consistent with prior studies (Baggett et al. 2018; Johnson et al. 2014; Mundhenk et al. 2018) that have applied this type of empirical model that relies on the propagation of the MJO.

While Fig. 1 highlights phase and lead combinations that have positive skill in the U.S. Southeast associated with the propagation of the MJO, it is not realistic to expect such a strong propagating MJO signal (and associated skill) to appear for all regions and seasons. It is often the case that the model’s skill is negative (i.e., the model is less skillful than a random forecast) for certain phase and lead combinations. Figure 2 shows the average HSS for each region and season calculated over all phase and lead combinations. For this and all similar figures, the value for each of the 262 overlapping forecast regions, with the average HSS for these regions indicated by the shading of the grid cells at the centers of the boxes. Figure 2 indicates that in some regions and seasons the aggregated negative skill over all MJO phases for weeks 3–6 actually outweighs the positive skill. However, examination of individual regions and seasons indicates that it is still possible for regions and seasons with high aggregated negative skill to feature some phase and lead combinations with positive skill, even if the pattern of high skill is not as coherent as in Fig. 1. For example, Fig. 2 shows that under the easterly QBO, negative skill outweighs positive skill over weeks 3–6 over the West Coast during winter and the southern Great Plains during summer. However, analysis of these regions (Figs. S1 and S2 in the online supplemental material) shows that positive skill is still available for some phase and lead combinations. Over the West Coast during winter, positive skill
for below-normal precipitation forecasts is especially prevalent during weeks 4 and 5 following MJO phases 3 and 4 under easterly QBO conditions. Over the southern Great Plains during summer, positive skill manifests during weeks 4 and 5 following MJO phases 7 and 8 under easterly QBO conditions.

Therefore, the key to successfully applying this model is to take advantage of those conditions under which a skillful
forecast can be made. To better isolate such instances of positive skill, we analyze all phase and lead combinations in each region and season to identify “skillful forecasts of opportunity.” We define a skillful forecast of opportunity as a particular phase and lead combination for which the model’s HSS is significantly better than a random forecast at the 90% confidence level using random walks. This characterization of skillful forecasts of opportunity is consistent with Baggett et al. (2018) and Mundhenk et al. (2018). In Fig. 1, for example, those phase and lead combinations considered skillful forecasts of opportunity are marked with black dots.

Figure 3 displays the frequency of skillful forecasts of opportunity across the contiguous United States and Alaska during each of the four seasons. For each season, we weight the frequencies based on how often each active MJO phase occurs under each QBO phase. Given an active MJO (61% of the time) and a particular phase of the QBO, skillful forecasts of opportunity are available approximately 20%–50% of the time during weeks 3–6 in most regions and seasons. While we show results parsed by QBO phase (easterly QBO occurs 44% of the time versus 56% for westerly QBO), we note that similar frequencies occur when combining verification metrics for both QBO phases (not shown). In some regions and seasons, such as the Midwest during spring under both QBO phases, the model only provides skillful forecasts of opportunity at a rate of about 10%–20%. However, in other regions and seasons, like the Pacific Northwest during summer under westerly QBO or the southern Great Plains during winter under easterly QBO, skillful forecasts of opportunity are present for greater than 50% of forecasts, given active MJO conditions.

How much skill do these skillful forecasts of opportunity actually provide? To answer this question, we calculate the average HSS among skillful forecasts of opportunity. As with frequency, we weight the average HSS based on how often each active MJO phase occurs. We find that, on average, skillful forecasts of opportunity provide appreciable skill (HSS between 20 and 30) in most regions and seasons (Fig. 4). In some regions and seasons skill scores exceed 33, implying that the model is correct more than twice as often as incorrect. Therefore, although skillful forecasts of opportunity are not always available, they provide a valuable source of skill when they do arise.

b. Drivers of empirical model skill

We next seek to provide context to the empirical model’s skill at select phase and lead combinations. Since the empirical model bases its forecasts on historical responses of pentadal precipitation anomalies to MJO and QBO conditions, we expect that the model’s skill is a result of the strength of the historical response to MJO and QBO conditions at initialization. We first examine the strength of the historical response for each phase and lead combination in a sample region and season, the U.S. Southeast during fall (the same region and season as in Fig. 1). We use the percentile of the median of the conditional distribution of precipitation anomalies, with respect to the full historical distribution of precipitation anomalies, as a proxy for the strength of the historical response for the given phase and lead combination. Figure 5 shows these percentiles for the U.S. Southeast during fall. Open dots indicate that the actual composite precipitation anomalies at that phase and lead combination are significantly different than those expected by random chance based on a block-bootstrapping method with 1000 iterations at 80% confidence, while filled dots denote 90% confidence.

Comparison of Figs. 1 and 5 indicates a high correlation between the HSS and the percentiles of the median conditional precipitation anomalies in the U.S. Southeast during fall. To better understand this relationship, we examine the strength of response of precipitation anomalies for each region while parsing by MJO phase, QBO phase, and season. Figures 6 and 7 show spatial fields of the percentiles of precipitation anomalies for the period 15–19 days after initialization for each MJO and QBO phase during winter (DJF) and fall (SON), respectively. We choose to highlight this lead time period because it falls at the beginning of the CPC’s week 3–4 precipitation outlook. For a forecaster at the CPC, these plots could be used as tool to inform the outlook, as they provide the dominant precipitation response at a lead of 3 weeks based on the current state of the MJO and QBO.

We note that strong precipitation responses often occur in regional clusters that are, in most instances, statistically significant with at least 80% confidence using block bootstrapping with 1000 iterations (significance denoted by white dots). One such cluster can be seen in the U.S. Pacific Northwest during winter under westerly QBO conditions (Fig. 6). In this region and season, strong positive precipitation anomaly responses occur 15–19 days after active MJO phases 1–3. In contrast, strong negative precipitation anomaly responses occur 15–19 days after active MJO phases 4–6. In general, the strength, location, and spatial extent of these clusters depends on season, though we note that strong responses manifest during all seasons (precipitation anomaly responses for other seasons are shown in Figs. S3 and S4). Nonetheless, we find that precipitation anomalies are modulated by the phases of the MJO and QBO. We also find that this modulation by MJO and QBO phase occurs for anomalies of precipitation into Week 6 (Figs. S5–S8), as well as for related synoptic quantities like 500-hPa geopotential height (Figs. S9–S16) and IVT (Figs. S17–S24). Figures 6 and 7 also highlight, for a given MJO phase, differences in historical precipitation anomaly responses between easterly and westerly QBO phases. For example, Fig. 6 indicates a cluster of statistically significant positive precipitation anomalies in the western United States 15–19 days following MJO phase 4 under easterly QBO conditions. However, there is a similarly strong cluster of negative precipitation anomalies in the western United States 15–19 days following MJO phase 4 under westerly QBO conditions. Again, visual inspection of all similar plots shows that the precipitation responses vary as a function of QBO phase, which is consistent with the findings of past studies showing that the strength of the MJO and QBO relationship differs by QBO phase (see introduction). These differences merit future study.
This regional perspective further highlights the parallels that exist between the clusters of strong historical responses seen in Figs. 6 and 7 and skill seen, for example, the U.S. Southeast during fall (Fig. 1). Under westerly QBO conditions, for example, skillful forecasts of opportunity for below-normal precipitation exist in the U.S. Southeast 15–19 days after MJO phases 2 and 3 during the fall (Fig. 7). Meanwhile, Fig. 7 indicates a cluster of relatively strong
negative precipitation anomaly responses over the region 15–19 days after MJO phases 2 and 3 under westerly QBO conditions. Figure 1 also highlights skillful forecasts of opportunity for above-normal precipitation in the U.S. Southeast 15–19 days after MJO phases 5, 6, and 8 under westerly QBO conditions. As expected, Fig. 7 shows broad areas of strong positive precipitation anomaly responses over the eastern half of the United States, including the Southeast.
15–19 days after MJO phases 5, 6, and 8. Analysis over all regions and seasons (not shown) indicates that the percentile of the conditional precipitation anomaly for each region and season is highly correlated with the empirical model’s skill.

Since model skill strongly varies with the historical precipitation response to MJO and QBO conditions, we next examine the relationship between historical precipitation response and the responses of other synoptic variables to MJO and QBO conditions. Figure 8 shows the Spearman correlation between the percentiles of the conditional distribution median 500-hPa geopotential height and precipitation anomalies. Again, percentiles are calculated in terms of the full historical distribution. With a few exceptions, especially over the U.S. Midwest, precipitation anomaly responses are negatively correlated with geopotential height anomaly responses (i.e., positive precipitation anomalies with anomalous troughing) at 500 hPa (red regions in Fig. 8). This correlation is most pronounced over the western half of the United States, where instability and other impacts from overhead troughs interact with the mountainous terrain to produce enhanced precipitation. Since upper-level divergence (convergence) theoretically occurs downstream of 500-hPa troughs (ridges), we also test the correlation of precipitation anomalies with geopotential height anomalies in regions upstream, though we find that correlations are very similar (not shown). This may occur because we use a combination of spatial ($7.5^\circ \times 7.5^\circ$ forecast regions) and temporal (5-day averaging) scales that is sufficient to capture both the upper-level height feature and the downstream precipitation response.

Figure 9 shows the Spearman correlation between the percentiles of the conditional distribution median IVT and...
precipitation anomalies. We find that IVT and precipitation responses are highly positively correlated in the western and southeast United States. These high correlations are likely related to the climatological distribution of landfalling ARs and tropical cyclones, both of which can produce extreme rainfall associated with anomalously high IVT. We find slightly negative correlations between IVT and precipitation responses in the U.S. northern Great Plains, potentially due in part to occasional downsloping of westerly plumes of high winds aloft (likely associated with plumes of anomalously high IVT) in the lee of the Rockies. Possible evidence of this is an associated cluster of high positive correlations that occur on the western slopes of the Rockies, especially during winter, spring, and fall. Nonetheless, we note that further study is required to definitively attribute any of these correlations to specific phenomena.

In the supplemental material, we also examine the influence of underlying synoptic conditions on the prediction skill for above- or below-normal pentadal precipitation anomalies. We find that the magnitudes of geopotential height anomaly responses are positively correlated with HSS for precipitation (Fig. S25), mostly in the western United States. We also find that the magnitudes of historical IVT responses are positively correlated with HSS for precipitation (Fig. S26), largely along the western and Gulf coasts, regions most influenced by focused areas of high IVT. However, since the correlations are not equal to 1, these synoptic quantities alone do not fully account for the variance (i.e., the correlations squared) of the model’s skill in predicting pentadal precipitation anomalies for weeks 3–6. Certainly, geopotential height and IVT fields alone do not provide enough physical information about other synoptic drivers of precipitation, such as surface temperature and moisture fields. In addition, these synoptic quantities do not resolve important mesoscale circulations (e.g., gust fronts, sea and land breezes, mountain and valley circulations), terrain features, or land-atmosphere interactions (e.g., soil moisture, vegetation fraction, etc.).

Overall, our findings showing that the MJO and QBO modulate precipitation, geopotential height, and IVT anomalies agree with past findings from those like Wang et al. (2018a), who in their Fig. 11 schematically demonstrate the influence of the combined MJO–QBO state on the North Pacific storm track. However, further study is required to isolate the most important drivers at all spatial and temporal scales, of the strong precipitation responses to MJO and QBO conditions. Specifically, future studies should leverage findings of past studies (see introduction) demonstrating the modulation of sensible weather over North America in order to place the results shown here into proper dynamical context.

c. Influence of the QBO

In the preceding sections, we established the ability of the empirical model to use the MJO and QBO as predictors to produce skillful forecasts of opportunity for above- or

![Percentile of Total Precipitation Anomaly for Days 15 through 19 using the MJO and QBO during Winter](image-url)
below-normal pentadal precipitation anomalies. We now seek to isolate the importance of the QBO as a predictor. To characterize the influence of the QBO phase on the predictions, we run the model using only the MJO (i.e., no conditioning on QBO phase) and compare regional and seasonal skill to the skill expected when incorporating the QBO. To maintain a constant sample size between cases, we combine the verification metrics for easterly and westerly QBO conditions and calculate skill scores. While prior studies (see introduction) have demonstrated the QBO’s modulation of the MJO, the nature of the modulation varies by QBO phase (see Figs. 6, 7, for example). Therefore, we expect model performance to differ when adding the specific QBO phase as a predictor.

Figure 10 demonstrates that, in all seasons, adding the QBO increases the frequency of skillful forecasts of opportunity in a majority of regions over all MJO phases. When adding the QBO as a predictor, skillful forecasts of opportunity are more frequent in over 70% of the regions, with the exception of the fall (about 61%). Table 1 lists, by season, the exact percentage of regions with an increase in skillful forecasts of opportunity. Figure 11 shows that adding the QBO also increases the average HSS of skillful forecasts of opportunity. Figure 10 demonstrates that, in all seasons, adding the QBO increases the frequency of skillful forecasts of opportunity in a majority of regions over all MJO phases. When adding the QBO as a predictor, skillful forecasts of opportunity are more frequent in over 70% of the regions, with the exception of the fall (about 61%). Table 1 lists, by season, the exact percentage of regions with an increase in skillful forecasts of opportunity. Figure 11 shows that adding the QBO also increases the average HSS of skillful forecasts of opportunity.
Table 1 shows that during the summer, the percentage of regions with higher frequency of skillful forecasts of opportunity increases from 72.9% to 88.2% to 93.9% as confidence level increases. Meanwhile, the percentage of regions with higher average HSS for skillful forecasts of opportunity decreases from 92.0% to 77.1% to 69.1% as confidence level increases. We note from Table 1 that the majority of regions and seasons see an enhancement of skillful forecasts of opportunity.

Fig. 8. The Spearman correlation between the percentiles, with respect to their respective full distributions, of the conditional distribution median 500-hPa geopotential height (“Z500”) and precipitation anomalies. We show the correlation for each QBO phase. Reds indicate increasing precipitation anomalies with decreasing height anomalies (anomalous troughs), while blues indicate increasing precipitation anomalies with increasing height anomalies (anomalous ridges). Correlation is calculated over MJO phase and lead combinations from week 3 through 6 (lead days 15–42).
for all three commonly used statistical significance thresholds. Therefore, we are confident that the QBO’s influence is robust across varying definitions of skillful forecasts of opportunity.

4. Conclusions
We apply an empirical model based on the MJO and QBO to predict above- or below-normal pentad precipitation anomalies for weeks 3–6 (15–42 days) throughout the contiguous United States.
States and Alaska during all seasons. First, in each region and season we analyze the occurrence of skillful forecasts of opportunity, phase and lead combinations at which the empirical model’s skill is significantly better than a random model at 90% confidence. We find that such opportunities are available for approximately 20%–50% of all forecasts in most regions and seasons, while some regions exhibit even higher frequencies. We also find that, on average, skillful forecasts of opportunity provide HSS values between 15 and 30.

We next put the empirical model’s skill at weeks 3–6 into context by exploring the relationship between the historical precipitation response (anomalies) to the MJO and QBO and the HSS for precipitation for each phase and lead combination. We find that historical precipitation responses during weeks 3–6 are modulated by the MJO and QBO during all seasons throughout the contiguous United States and Alaska. Furthermore, responses of other synoptic drivers of precipitation, like geopotential height and IVT, are also modulated by the MJO and QBO during weeks 3–6, a finding that is consistent with prior studies (Baggett et al. 2017; Feng and Lin 2019; Wang et al. 2018b). Overall, we find that historical precipitation responses to the MJO and QBO are strongly correlated with empirical model skill. In addition, we show that historical 500-hPa geopotential height and IVT responses are also correlated with precipitation responses and associated precipitation forecast skill, though neither quantity fully explains the skill or the historical response to MJO and QBO conditions.

Last, we assess the influence of adding the QBO as a predictor by comparing the empirical model to an alternative version that

<table>
<thead>
<tr>
<th>Season</th>
<th>Confidence level</th>
<th>Percent FOO frequency increased (%)</th>
<th>Percent avg FOO HSS increased (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Winter</td>
<td>90</td>
<td>88.2</td>
<td>95.4</td>
</tr>
<tr>
<td></td>
<td>95</td>
<td>96.9</td>
<td>85.9</td>
</tr>
<tr>
<td></td>
<td>97.5</td>
<td>96.6</td>
<td>75.2</td>
</tr>
<tr>
<td>Spring</td>
<td>90</td>
<td>88.9</td>
<td>93.9</td>
</tr>
<tr>
<td></td>
<td>95</td>
<td>95.8</td>
<td>83.6</td>
</tr>
<tr>
<td></td>
<td>97.5</td>
<td>96.6</td>
<td>73.7</td>
</tr>
<tr>
<td>Summer</td>
<td>90</td>
<td>72.9</td>
<td>92.0</td>
</tr>
<tr>
<td></td>
<td>95</td>
<td>88.2</td>
<td>77.1</td>
</tr>
<tr>
<td></td>
<td>97.5</td>
<td>93.9</td>
<td>69.1</td>
</tr>
<tr>
<td>Fall</td>
<td>90</td>
<td>60.7</td>
<td>96.6</td>
</tr>
<tr>
<td></td>
<td>95</td>
<td>74.0</td>
<td>82.1</td>
</tr>
<tr>
<td></td>
<td>97.5</td>
<td>82.1</td>
<td>69.8</td>
</tr>
</tbody>
</table>
solely uses the MJO as a predictor. We show that while combining the MJO and QBO tends to further decrease skill when it is already modest or low, adding the QBO increases the frequency of skillful forecasts of opportunity in a majority of regions and seasons. We also find that adding the QBO increases the average HSS of skillful forecasts of opportunity in most regions and seasons. The exact fraction of regions enhanced varies by season and confidence level. Thus, we conclude that adding the QBO preferentially enhances skill where it is already high.

These results demonstrate the utility of the MJO and QBO as predictors for anomalous pentadal precipitation in a variety of regions and seasons. We build on the findings of Baggett et al. (2018) and Mundhenk et al. (2018), who applied this type of model to S2S forecasts of severe convection and ARs, two phenomena associated with precipitation. We also build on the findings of Johnson et al. (2014), who first demonstrated the utility of an empirical model based on historical MJO and ENSO conditions to predict temperatures across the United States at S2S leads. Given past studies that have demonstrated the modulating influence of the MJO and QBO on precipitation-inducing phenomena, it makes sense that the MJO and QBO would be useful predictors of precipitation.

While the empirical model’s utility is restricted to select “opportunities,” we demonstrate that these opportunities present a significant improvement over current dynamical models at S2S leads. This empirical model provides an additional forecasting tool that could help inform the CPC’s week 3–4 precipitation outlook, thereby supplementing dynamical models while also highlighting sources of predictability that dynamical models have yet to fully exploit.

Acknowledgments. The CPC Global Unified Precipitation data are provided by NOAA/OAR/ESRL PSD, Boulder, Colorado, from their website (https://www.esrl.noaa.gov/psd/). NCEP–NCAR reanalysis data provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, from their website at https://www.esrl.noaa.gov/psd/. QBO data are provided by NOAA/NWS/CPC, Camp Springs, Maryland, from their website (http://www.cpc.ncep.noaa.gov/data/indices/). Please visit https://www.cpc.ncep.noaa.gov/data/indices/Readme.index.shtml for additional information about the QBO index. RMM MJO data are provided by the Australian Bureau of Meteorology from their website (http://www.bom.gov.au/climate/mjo/). IVT calculations were performed by C.F.B. Empirical model code was originally developed by Bryan Mundhenk using Python 2.7 and later modified into its current form by C.F.B. and K.M.N. using NCL 6.4. K.M.N., E.A.B., E.D.M., C.F.B., D.S.H., and L.M.C. are supported by NOAA Climate Test Bed Grant NA18OAR4310296. E.D.M. is also supported by the NSF Climate and Large-Scale Dynamics Grant AGS-1841754. E.A.B. and E.D.M. are also funded, in part, by NOAA OWAQ Grant NA19OAR4590151.
APPENDIX

Online Repository
Since most of the analyses herein and in the main text are summaries over all phase and lead combinations, we provide an online repository of HSS calculations by region and season. The contents of this online repository (https://hdl.handle.net/10217/195747) depict HSS for each region and season in an MJO phase by lead time space (similar to Fig. 1 for the U.S. Southeast during fall). The repository also contains maps of HSS for all MJO/QBO phase and lead combinations. In this repository, stakeholders can select a region and season of interest and identify the exact MJO and QBO phases for which skillful forecasts of opportunity are available. The analyses performed here, for precipitation using the MJO and QBO, are also included in an online web application project (http://barnes.atmos.colostate.edu/S2SPredictionModel/).

REFERENCES


Ralph, F., and M. Dettinger, 2012: Historical and national perspectives on extreme West Coast precipitation associated with