Seasonal-to-Subseasonal Model Forecast Performance during Agricultural Drought Transition Periods in the U.S. Corn Belt

NICHOLAS J. SCHIRALDI AND PAUL E. ROUNDY

Department of Atmospheric and Environmental Sciences, University at Albany, State University of New York, Albany, New York

(Manuscript received 6 February 2017, in final form 31 May 2017)

ABSTRACT

The prediction of drought onset and decay in the U.S. Corn Belt region (CBR) on seasonal-to-subseasonal time scales has not been well studied. This study utilizes the subseasonal-to-seasonal prediction archive to assess model errors in large-scale circulation patterns associated with agricultural drought transition periods, targeting models used by the European Centre for Medium-Range Forecasts, National Centers for Environmental Prediction, and Australian Bureau of Meteorology. An analysis of the seasonal cycle of bias for geopotential anomalies at 200 hPa and net radiation at the top of the atmosphere in each model is presented and used to subtract the long-term bias from each model. Model fields are decomposed into three spectral bands—low frequency (periods $>$ 100 days), intraseasonal (periods 20–100 days), and synoptic (periods $<$ 20 days)—to demonstrate each model's ability to predict patterns associated with agricultural drought transition periods in each band. Results demonstrate that ECMWF and NCEP struggle in predicting the large-scale circulation patterns associated with 20-day agricultural drought and onset transitions, but are more skillful in predicting the patterns associated with 60-day agricultural drought onset and decay events at reforecast hour lead window 360–480 (F360–F480). BoM was not skillful in predicting the circulation patterns associated with either type of drought transition. Results also demonstrate that the errors associated with these events are no worse than historical errors for the target study period.

1. Introduction

The prediction of drought onset and decay over the U.S. Corn Belt region (CBR) on seasonal-to-subseasonal time scales has not been well studied. Since drought in the CBR is one of the most costly natural disasters (Wilhite 2000; Adonizio et al. 2012; Hoerling et al. 2014), accurate prediction of the atmospheric phenomena that contribute to transition periods toward or away from drought is critical to understanding and communicating the risks of economic loss. Traditionally, assessments of the prediction of drought over the CBR have focused on the seasonal prediction (with leads of 1–6 months) of quantities such as the standardized precipitation index (SPI) and soil moisture (Quan et al. 2012; Mo and Lyon 2015; Luo and Wood 2007). Recent studies have demonstrated skill in predicting drought over the CBR (Quan et al. 2012; Mo and Lyon 2015; Luo and Wood 2007), while others have shown poor forecast performance (Wu and Kinter 2009; Hoerling et al. 2014), highlighting that the prediction of drought varies greatly from event to event. Overall, the probability of detection of drought onset (using SPI as an indicator) in the CBR for a multimodel seasonal forecast ensemble is less than 45% (Yuan and Wood 2013). While providing accurate seasonal forecasts of drought would yield tremendous socioeconomic gains, doing so may be beyond the limits of predictability (Lorenz 1963; Guo et al. 2011). On the other hand, gradually varying signals in the climate system might provide some predictability (Mo and Lyon 2015). As such, it is important to understand the ability of models to predict agricultural drought transitions at shorter leads of 1–4 weeks.

Short- ($<$1 week) and medium-range (2–3 week) precipitation forecast skill over the CBR have been assessed in the literature (Li and Robertson 2015), and many statistical postprocessing methods have been developed to improve such forecasts (Hamill et al. 2004, 2008; Scheuerer and Hamill 2015). However, a comprehensive evaluation of model performance during drought transition periods at such lead times has not been presented. Recent access to reforecast data

Corresponding author: Nicholas Schiraldi, nschiraldi@albany.edu

DOI: 10.1175/MWR-D-17-0026.1

TABLE 1. For each model, the time range (forecast lead time, in days), resolution, ensemble size, frequency of initialization, reforecast period, and whether or not the model has ocean coupling and sea ice coupling is presented for reference. This study focuses on forecasts verifying during May–August (1995–2013). Adapted from Vitart et al. (2016) to show only the models presented in this study.

Model	Time range (days)	Resolution	Ensemble size	Frequency of initialization	Reforecast period	Ocean coupling	Sea ice coupling
BoM	0–62	~2° × 2°; L17	33	Twice weekly	1981–2013	Yes	No
ECMWF	0–32	0.25° × 0.25° for days 0–10, 0.5° × 0.5° after day 10; L91	51	Twice weekly	1995–2015	Yes	No
NCEP	0–44	~1° × 1°; L64	4	6 hourly	1981–2015	Yes	Yes

initialized from near-real-time versions of numerical weather models employed by 11 modeling centers, via the subseasonal-to-seasonal prediction archive (S2S; Vitart et al. 2016), allows for such an assessment.

Schiraldi and Roundy (2017) demonstrated that 20- and 60-day agricultural drought transition onset and decay events are forced primarily by intraseasonal atmospheric Rossby wave trains that occur in conjunction with westward-propagating intraseasonal convection in the Northern Hemisphere subtropics. Schiraldi and Roundy (2017) further demonstrate a low-frequency wave train that increases the likelihood of drought onset or decay twofold. Understanding how well represented these patterns are by numerical weather prediction may yield a better understanding of the predictability of such events. Rather than focusing solely on the skill of precipitation forecasts during such transition periods, this study assesses the ability of models to predict correctly the synoptic, intraseasonal, and low-frequency patterns relevant to 20- and 60-day agricultural transitions over the CBR.

2. Data

Reforecasts from the Australian Bureau of Meteorology (BoM), European Centre for Medium-Range Weather Forecasts (ECMWF), and National Centers for Environmental Prediction (NCEP) are evaluated in this study. As a result of the nature of the reforecast release schedule of the S2S dataset (Vitart et al. 2016), these are the only models that have sufficient data (in a continuous set of reforecasts) at the time of this study. The data are output on varying horizontal grids, with varying initialization frequencies, ensemble sizes, reforecast periods, and lead ranges (Table 1). For each model's own reforecast period, 24-hourly ensemble mean forecasts of geopotential height at 200 hPa (Z200) and net radiation at the top of the atmosphere (NRAD TOA) from reforecast hour 0000 (F0000) to reforecast hour 1440 (F1440) initialized at 0000 UTC are obtained.

a. ECMWF

To achieve a continuous reforecast of ECMWF, reforecasts from different versions of the model are utilized. Reforecasts from ECMWF initialized between 1 January and 31 December 2015 are used in this study, yielding a reforecast period from 1 January 1995 through 31 December 2014. Dates initialized between 1 January and 12 May 2015 use cycle 40r1, while reforecasts after 12 May 2015 use cycle 41r1, both being different than the operational model (cycle 41r2) at the time of this paper. The main differences between cycle 40r1 and 41r1 are that in 41r1, the reforecasts began to be initialized twice weekly, the reforecast was extended out to 46 days, and there was an upgrade to the resolution of the data assimilation (ECMWF 2016). Since forecasts out to 46 days are not available for the entire period, only forecasts out to 32 days are considered, on a 1.5° × 1.5° grid. Because of horizontal resolution upgrades implemented in cycle 41r2 (ECMWF 2016), the systematic bias and reforecast skill presented in this paper may not be applicable to the current version of the model, but can be used to benchmark cycle 41r2 and future versions of the model.

b. BoM

The reforecast period obtained for BoM covers from 1 January 1981 through 31 December 2013. Reforecasts are initialized every 5 days starting from 1 January, at 0000 UTC, output onto an approximately 2.5° × 2.5° grid and run with 32 ensemble members. Reforecasts out to 62 days are available, but for consistency with ECMWF, only reforecasts out to 32 days are analyzed. The reforecasts are initialized from the operational model at the time of this study (Vitart et al. 2016).

c. NCEP

The climate forecast system reforecast data from NCEP are obtained from 1 January 1982 through 31 December 2014. Since the NCEP reforecast dataset ends on 31 December 2010, the dataset was extended

through 2014 using archived forecasts from operational runs that are consistent with the reforecast model. Unlike BoM and ECMWF, the CFS is deterministic (Saha et al. 2014). A four-member ensemble is created using overlapping lead windows from the 0000, 0600, 1200, and 1800 UTC initialized forecasts. Data are output onto a 2.5° × 2.5° grid, and run to exactly 45 days, although forecasts only through 32 days are analyzed. Previous literature has noted a change in the tropical climatology of the initial conditions (CFSR) and subsequent CFS reforecasts after 1999 (Saha et al. 2014). Anomalies are created with respect to the outgoing longwave radiation (OLR) dataset described in section 2d and this shift in the data should not have a large impact on the results presented.

d. About the reforecasts

While a longer time period of reforecasts is available for NCEP and BoM compared to ECMWF, this study focuses only on agricultural drought transition events occurring in the common time period between all three models: 1995–2013. This decreases the sample size of CBR drought transition events to forty-nine 20-day drought decay events, sixty 20-day drought onset events, twenty-one 60-day drought decay events, and twenty-five 60-day drought onset events. These events remain reasonably scattered through the May–August climatology, as discussed by Schiraldi and Roundy (2017). The results presented are representative of the full NCEP and BoM datasets.

Using a consistent climatology between all three models introduces another caveat: that the models are not initialized on the same days throughout that climatology. As such, the model reforecasts were averaged to six 120-h lead windows: F000–F120, F120–F240, F240–F360, F360–F480, F480–F600, and F600–F720. Reforecasts are analyzed if the transition event occurs at some point during the target window. This leads to reforecasts across modeling centers being slightly shifted in initialization, but capturing the same events across lead windows. At longer leads (>120 h), these slight shifts are irrelevant. Only the ensemble mean is discussed for each model.

e. Reanalysis data

Analyses at 0000 UTC from the European Centre for Medium-Range Weather Forecasts interim reanalysis (ERA-I; Berrisford et al. 2011) are used to verify the 5-day mean reforecasts for Z200. This approach may give ECMWF a small advantage over the other models, but it is likely to have little effect on longer-lead reforecasts. The NRAD TOA reforecasts are verified using the National Oceanic and Atmospheric Administration OLR daily climate (Lee 2014). The OLR dataset is available on a 1.0° × 1.0° grid from 1979 to the present. As in Schiraldi and Roundy (2017), NCEP's Unified Gauge-Based Precipitation Dataset (UPD; Chen et al. 2008), which contains a 0.25° × 0.25° mesh of daily (1200–1200 UTC) precipitation observations over a 59-yr base period (1 January 1948–31 December 2006) for the region bounded by 20.125°–49.875°N, 230.125°–304.875°E, is used to identify drought transition periods. The UPD have been extended through the present, on a slightly shifted grid, and are updated daily.

3. Methods

a. Identifying drought transition events

The primary focus of this study is to assess how well the BoM, ECMWF, and NCEP models predict patterns associated with agricultural drought transition periods in the CBR during May–August. Following Schiraldi and Roundy (2017), a CBR-based precipitation index is created by averaging grid points over Illinois, Indiana, Iowa, Minnesota, Missouri, Nebraska, Ohio, and South Dakota. Precipitation anomalies are created for a forward- and a backward-looking n-day accumulated precipitation time series (i.e., on 1 June there is a time series summing 1 June + n day and 1 June − n day), by first applying a power transform to each forward- and backward-looking time series to make them approximately normal, then removing the first five seasonal harmonics (1979–2014) from each time series, and finally standardizing the resulting anomalies by the standard deviation of daily data (Schiraldi and Roundy 2017). As described by Schiraldi and Roundy (2017), drought decay is defined as the previous n-day accumulated precipitation anomalies ≤ -0.5 standard deviation followed by at least n days of accumulated precipitation anomalies ≥ 0.5 standard deviations, where drought onset is the opposite scenario. This study focuses on n day = 10 and 30, equating to transitions occurring over 20 and 60 days, where n day = 10 will be referred to as 20-day drought onset or decay transition and n day = 30 is referred to as 60-day drought onset or decay transition. Schiraldi and Roundy (2017) also analyzed transitions of other lengths and found that they tended to cluster about these two sets.

b. Computing reforecast anomalies

The Z200 reforecast anomalies are created by subtracting a 35-yr climatology (1979–2014) calculated from 0000 UTC analyses from the ERA-I, to be consistent with the reforecasts verifying only at 0000 UTC. This climatology is calculated by regressing the Z200

analyses onto the first five seasonal harmonics. The results have been tested using between four and six harmonics, with no significant impact reported between five and six. Similarly, NRAD TOA forecast anomalies are calculated from the OLR dataset. The 0000 UTC anomalies are computed from daily mean data. Since reforecasts are only reported at 0000 UTC, variations in the diurnal cycle that are in the daily mean OLR dataset are likely not well represented in the reforecasts. Any bias that arises from this inconsistency is calculated and subtracted prior to the analysis described in section 3d.

c. Filtering the reforecast

The main goal of this study is to analyze how well the low-frequency (periods >100 days), intraseasonal (periods >20 and <100 days), and synoptic (periods <20 days) circulation patterns associated with drought onset and decay (Schiraldi and Roundy 2017) are predicted by S2S model reforecasts. To do so, Z200 and NRAD TOA must be linearly decomposed into the low-frequency, intraseasonal, and synoptic temporal bands. Since the reforecasts are not initialized at a consistent frequency, a traditional Lanczos filter cannot be used. Instead, Fourier regression is utilized to filter each 5-day-mean lead window. A matrix $\mathbf{x}$ is created with n sine and cosine waves harmonic to the total length of time the dataset spans, including only the periods corresponding to the filter band of interest, ordered in the columns of $\mathbf{x}$ with time t increasing down the rows. For each lead window, a second matrix $\mathbf{Y}$ is filled with the Z200 or NRAD TOA anomalies, with t increasing down the rows. Ordinary least squares regression is then performed as

$$c = \mathrm{inv}(\mathbf{x}^{\mathrm{T}} \cdot \mathbf{x}) \cdot \mathbf{x}^{\mathrm{T}} \cdot \mathbf{Y} \tag{1}$$

to extract the linear relationship between the harmonics associated with the filtered periods $\mathbf{x}$ and the reforecast data $\mathbf{Y}$. Since $\mathbf{Y}$ is not linearly spaced in time, only the rows associated with the Julian days of year of $\mathbf{Y}$ are taken from $\mathbf{x}$. The filtered fields $\mathbf{F}$ are reconstructed as

$$\mathbf{F} = \mathbf{x} \cdot c, \tag{2}$$

retaining only the time steps included in $\mathbf{Y}$. This method is used to calculate the low-frequency and intraseasonal bands, while the synoptic filtered data are calculated by subtracting the sum of the low-frequency and intraseasonal anomaly from the total anomaly. A caveat of this approach is that Fourier regression filtering is subject to Gibbs ringing at the edges of the filter band, but the breadth of the bands reduces such impacts.

d. Calculating the long-term bias and bias-corrected reforecast

Fourier regression is also used to calculate the seasonal cycle of the systematic bias in Z200 and NRAD TOA. Bias is calculated by subtracting the 5-day-mean reforecasts from their associated 5-day-mean verifications at each grid point. Then, the first five harmonics of the seasonal cycle are fit to the bias time series at each grid point. The results are not sensitive to additional harmonics beyond five. The seasonal cycle of bias, based on the initialized Julian day of year and reforecast lead window the total field, is used to create a bias-corrected reforecast. This approach removes the long-term systematic bias but retains any short-term regime-dependent biases. The seasonal cycle of bias is calculated for the full reforecast period available for each model, not a unified time period.

e. Skill metrics

Both error and anomaly correlation coefficient (ACC) are used to measure the skill of the reforecasts for each modeling center. Error is defined as

$$\mathrm{error}_{i,l,x,y} = f_{i,l,x,y} - v_{i,l,x,y}, \tag{3}$$

where f is the reforecast, v is the verification, i is the initialization, l is the lead window, x is the longitude index, and y is the latitude index. The anomaly correlation coefficient is defined as the correlation between the reforecast and verification, and is computed as

$$\mathrm{ACC}_{i,l} = \frac{\sum(\mathbf{Z}_{f(i,l)} - \overline{\mathbf{Z}_{f(i,l)}})(\mathbf{Z}_{v(i,l)} - \overline{\mathbf{Z}_{v(i,l)}})}{\sqrt{\sum(\mathbf{Z}_{f(i,l)} - \overline{\mathbf{Z}_{f(i,l)}})^2 \sum(\mathbf{Z}_{v(i,l)} - \overline{\mathbf{Z}_{v(i,l)}})^2}}, \tag{4}$$

where $\mathbf{Z}_f$ is the reforecast anomaly and $\mathbf{Z}_v$ is the verification anomaly, while the overbar denotes the area mean. Gridded data are weighted by the square root of the cosine of latitude before calculating the ACC.

f. Statistical significance and performance benchmark

A Monte Carlo test is applied to the composite data to test for statistically significant reforecast anomalies and to provide a benchmark against historical performance in the reforecast dataset. Events from the population of each reforecast composite are drawn at random 1000 times with replacement to generate 1000 new reforecast composites. The data from the 1000 new reforecast composites are then tested for statistically different signals from zero at the 95th percentile. An estimate of the expected performance is calculated by drawing at random with replacement 1000 reforecasts initialized in the same month as the transition event of

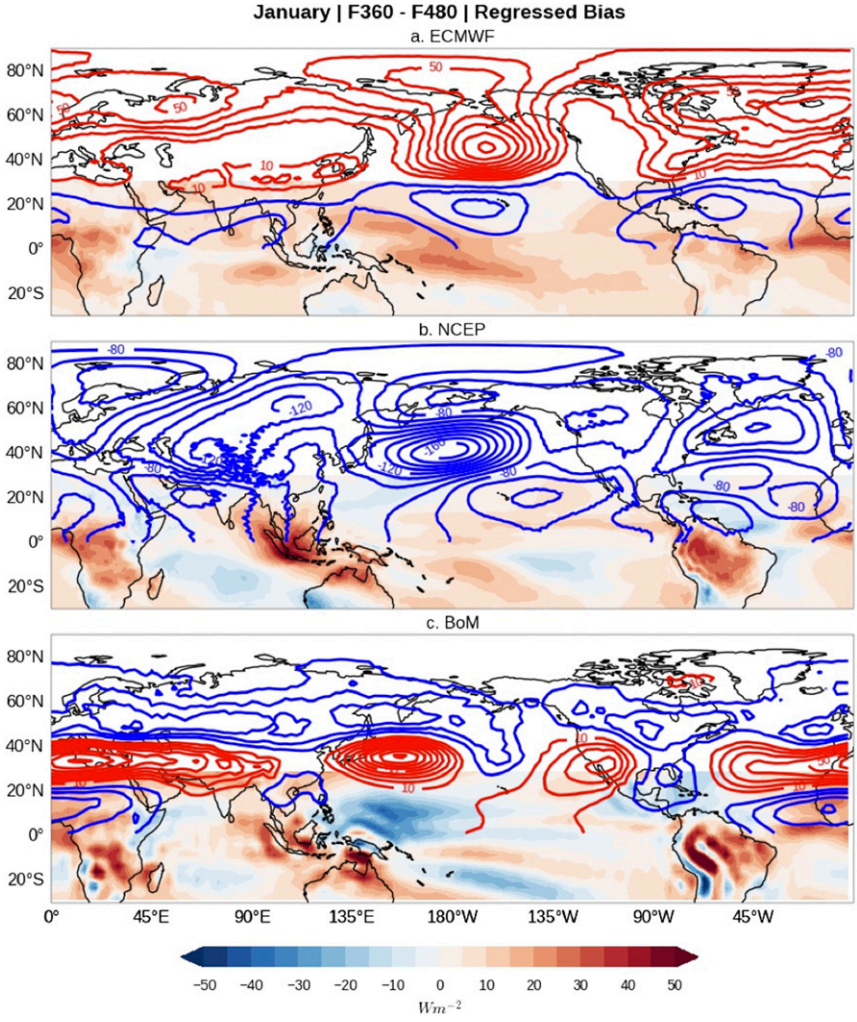

FIG. 1. The regressed bias for Z200 (contoured) and NRAD TOA (shaded) is presented for 15–20-day reforecasts initialized in January (representative of DJF and all leads) in (a) ECMWF, (b) NCEP, and (c) BoM. The Z200 contours start at ±10 m and are contoured every 10 m.

interest after first removing the transition from the available population. From there, the 95% confidence band of expected skill is taken by sorting the skill metric, and taking the 25th and 975th values from the random draws.

g. *Fourier decomposition for sparse data*

Wavenumber frequency spectra are calculated using a two-dimensional Fourier decomposition in time and space. Since the reforecast data are intermittently initialized, the Fourier decomposition in time is calculated with respect to daily initialized data. This decomposition is computed by creating a discrete Fourier transform matrix containing coefficients for all daily frequencies from

$$\mathbf{A}_k = \sum_{n=0}^{n-1} e^{-i2\pi nk/N} \quad k = 0, \ldots, n-1, \tag{5}$$

where N is the length of time, n is the irregular time stepping of the model initialization dates, and k is the daily time stepping. The Fourier decomposition in time then follows as

$$\mathcal{F}(\mathbf{A}_k) = \mathbf{x} \cdot \mathbf{A}_k, \tag{6}$$

where $\mathbf{x}$ is a matrix of the data with time going down the rows and space across the columns. The ends of $\mathbf{x}$ are tapered by a cosine bell to minimize spectral leakage. The highest frequencies are distorted because of the sparsity of the data and are not plotted. Tests on similar data with imposed artificial gaps but with data available

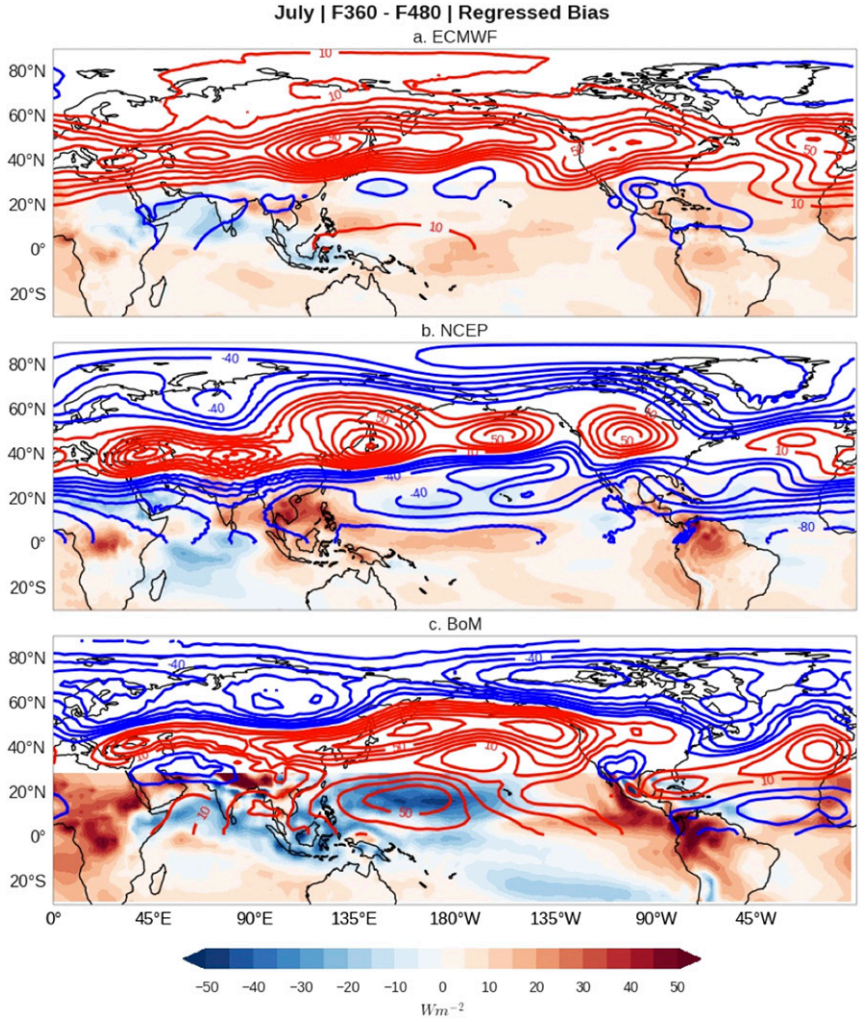

FIG. 2. The regressed bias for Z200 and NRAD TOA is presented for 15–20-day reforecasts initialized in July (representative of JJA and all leads) in (a) ECMWF, (b) NCEP, and (c) BoM. The Z200 contours start at ±10 m and are contoured every 10 m.

for comparison without gaps show that intraseasonal and lower frequencies are not distorted.

4. Seasonal cycle of bias

The seasonal cycle of model bias varies with initialized day of year, lead window, and model. For each of the BoM, ECMWF, and NCEP models, the bias in Z200 and NRAD TOA grows in amplitude with lead window, while the bias patterns remain largely in phase across lead windows for a given day of year. The exact bias patterns themselves differ from model to model, but there are some common overarching themes, especially in Z200. The seasonal cycle of bias of ECMWF is discussed in section 4a, NCEP in section 4b, and BoM in section 4c.

a. ECMWF

During the winter, the ECMWF Z200 bias is characterized by anomalously high heights over the central Pacific Ocean, eastern North America, the North Atlantic Ocean, and northern Eurasia, with anomalously low heights equatorward of 30°N and over the west coast of the United States (Fig. 1a). The anomalously low heights in the tropics may be in response to a bias favoring reduced convection globally, as evidenced by the dry bias throughout much of the global tropics (Fig. 1a). The dry bias in the global tropics may lead to a decrease in wave activity across the extratropical waveguide, allowing the model to quickly return to its internal climate. This effect is compounded by a weaker background extratropical potential vorticity gradient, resulting in a

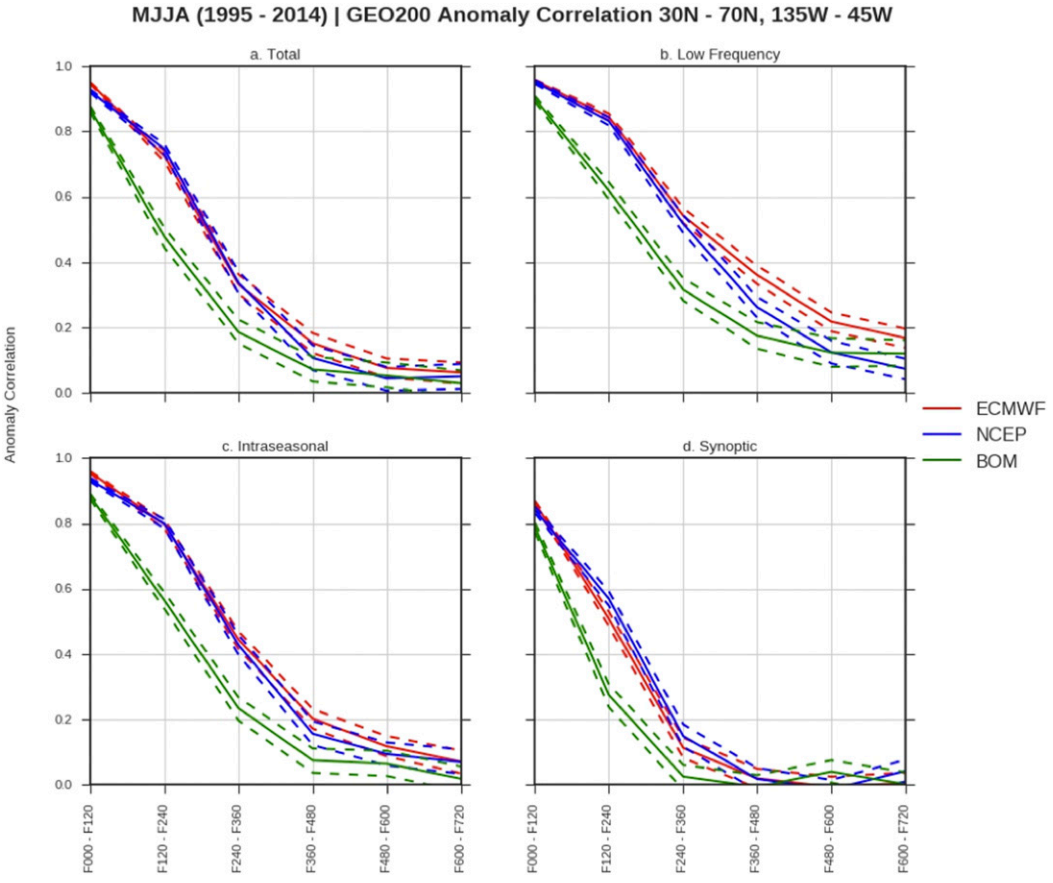

FIG. 3. The average anomaly correlation coefficient for 200-hPa geopotential heights during May–August (1995–2013) averaged over 30°–70°N and 135°–45°W. The solid line is the mean ACC for ECMWF (red), NCEP (blue), and BoM (green), while the dashed line approximates the 95% confidence interval of the mean for the (a) total, (b) low-frequency, (c) intraseasonal, and (d) synoptic fields.

poleward-shifted extratropical Rossby waveguide and an enhanced wave-breaking frequency.

The positive height bias in Z200 grows throughout the Northern Hemisphere spring (not shown), leading to a positive height bias between roughly 25° and 60°N across the globe, throughout the summer (Fig. 2a). The exception is over Greenland, where a weak negative height anomaly bias exists in Z200 (Fig. 2a). The ECMWF dataset used in this paper does not couple sea ice; therefore, it may not be appropriately melting sea ice in this region, which could explain some of this negative height bias, although testing this hypothesis is beyond the scope of this study. The tropics are characterized by a negative height bias in Z200, although it is not as uniform across the globe as in the winter (Fig. 2a). The dry bias in NRAD TOA throughout the tropics exists during the summer, except for over the Indian Ocean and Maritime Continent, where a weak wet bias exists (Fig. 2a). These patterns suggest that ECMWF has a bias favoring drought development in the central plains due to less large-scale support for rainfall over the region at longer lead times.

b. NCEP

NCEP has a similar summertime bias compared to ECMWF in Z200, but a much different wintertime bias. During the winter, NCEP has a cool bias in the Northern Hemisphere (Fig. 1b). There are three main centers of negative height biases, over the Pacific Ocean, over Eurasia, and over the North Atlantic. It is unclear why NCEP has such a strong negative height bias at Z200 during the winter, but the negative height bias would have strong implications on winter Rossby wave characteristics in the model and should be investigated further in future work.

NCEP's winter bias in NRAD TOA indicates suppressed convection over the Maritime Continent, active convection in the Indian Ocean, suppressed convection in the equatorial Pacific north of the equator, and a center of active convection south of the equator near

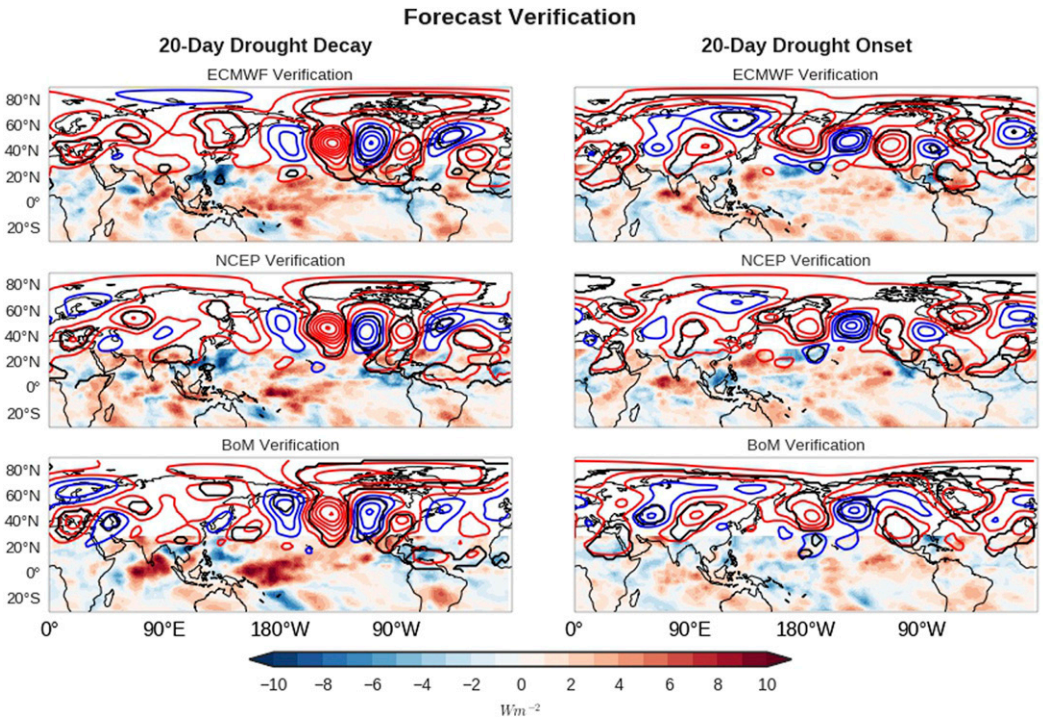

FIG. 4. Composite mean verification 200-hPa geopotential height anomalies (contours) and OLR anomalies (filled contours) for 20-day drought (left) decay and (right) onset events. The solid black line indicates 200-hPa geopotential height anomalies that are statistically different from zero at the 95% confidence level. The Z200 contour intervals start at +5 m (red) and −5 m (blue) with an interval of 10 m. OLR anomalies tend to be significant at $\pm 2\ \mathrm{W\,m^{-2}}$.

135°W (Fig. 1b). There does not appear to be a clear pathway whereby these NRAD TOA biases could be associated with the net cool bias in Z200. The Z200 bias changes drastically in the summer, where much of the globe from approximately 40° to 60°N is characterized by a positive Z200 height anomaly, with negative Z200 height anomalies poleward and equatorward (Fig. 2b). NCEP's Z200 bias favors positive heights over the central United States, implying a bias toward drought during the summer (Fig. 2b).

c. *BoM*

BoM has a much different seasonal cycle of bias compared to ECMWF. During the winter, BoM has an expansive wet bias from the equator to 20°N and from 135°E to 180° (Fig. 2c). This wet bias feeds a positive Z200 bias on its poleward side via anomalous adiabatic heating and the redistribution of potential vorticity by the divergent wind (not shown). This results in a stronger potential vorticity gradient and extended western Pacific jet, substantially modulating the extratropical waveguide. Farther downstream, there is a positive Z200 height bias over the western half of the United States, centered over northern Mexico, and a negative Z200 height bias over the rest of North America (Fig. 2c). There are interesting bias patterns in NRAD TOA over South America and Africa, but discussing them in detail is beyond the scope of this paper (Fig. 2c).

During the summer, the negative NRAD TOA bias in the tropical west Pacific strengthens, shifting poleward and extending well beyond 180° (Fig. 2c). This shift leads to a poleward adjustment in the positive Z200 height bias over the Pacific Ocean, tightening the geopotential gradient across the Pacific (Fig. 2c). The strongest positive Z200 height bias across the Pacific Ocean occurs near 45°N, close to its location in ECMWF; however, BoM has a negative Z200 bias across much of the globe poleward of 60°N. Farther downstream, the eastern extent of the Pacific Z200 ridge anomaly extends into the west coast of the United States and into the Great Lakes region, with a negative Z200 bias centered over Texas (Fig. 2c). This circulation bias demonstrates that BoM tends to favor anticyclonic wave breaking over the United States. Overall, the Z200 bias suggests that BoM may favor an enhanced frequency of drought transition events over the region during the summer, depending on where the wave tends to break over North America. The positive Z200 bias from 40° to 60°N is

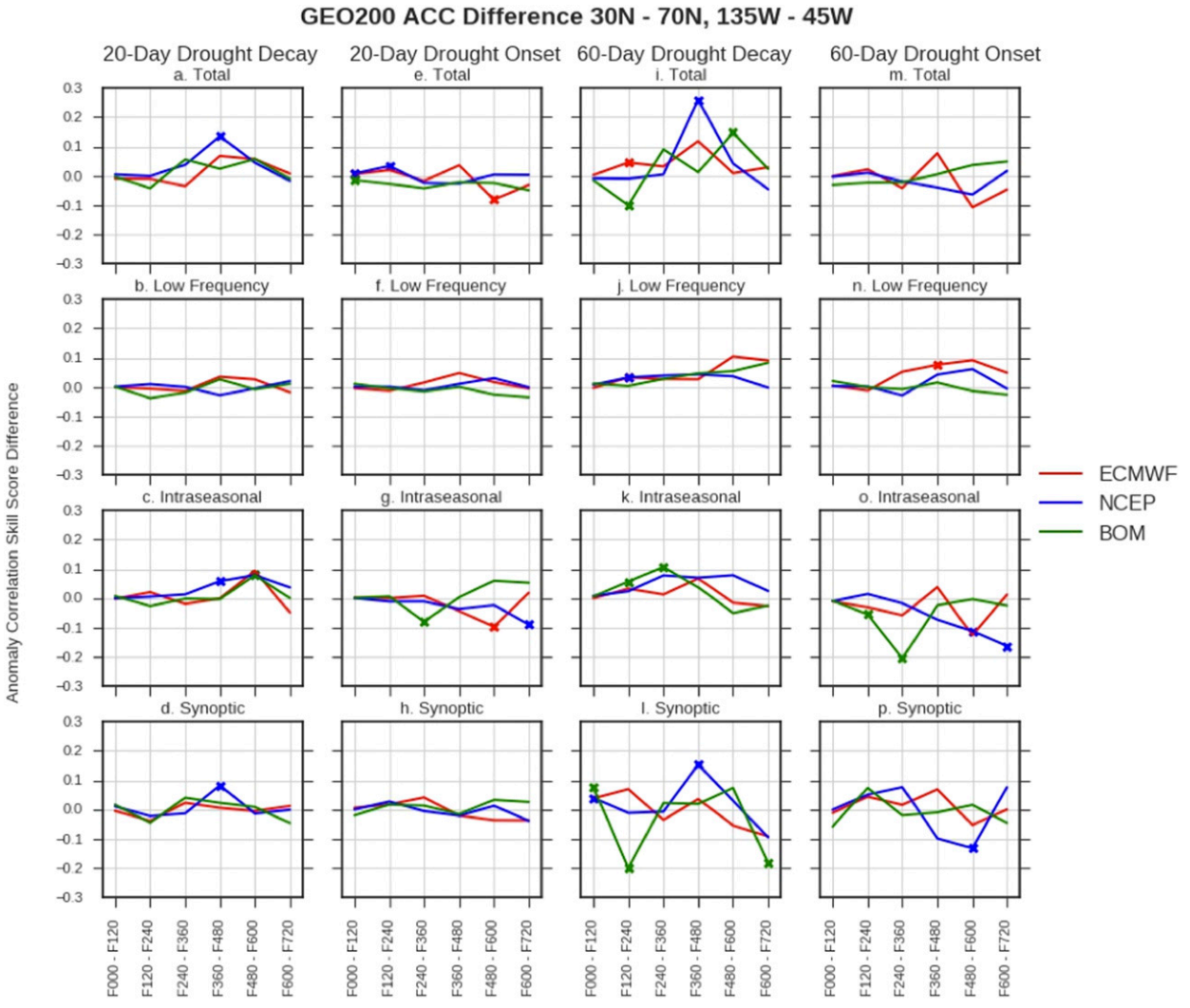

FIG. 5. The average ACC difference between 200-hPa geopotential heights during May–August (1995–2013) averaged over 30°–70°N and 135°–45°W, and the ACC of the events over the same box. Positive values indicate the events have higher ACCs than the 95th percentile confidence interval of climatology, where negative values indicate the events have lower ACCs than climatology. Tick marks along the line indicate the difference is statistically significant at the 95% confidence level.

similar to the ECMWF (Fig. 2a) and NCEP (Fig. 2b) results, suggesting the possibility that errors may arise from common internal forcing mechanisms (e.g., biases associated with physical parameterizations) across all three models.

5. Assessment of model performance during May–August

Schiraldi and Roundy (2017) demonstrated low-frequency, intraseasonal, and synoptic Z200 and NRAD TOA circulation patterns associated with agricultural drought (May–August 1979–2014), noting the largest impacts on the transition periods came from the intraseasonal band. Schiraldi and Roundy (2017) note the amplification of an intraseasonal extratropical Rossby wave that occurs coincident with westward-propagating tropical convection, leading to a center of convection at 135°E on the transition day. This section diagnoses how well these patterns are reforecast in NCEP, ECMWF, and BoM at leads of up to 30 days. Section 5a provides context for the expected model performance during May–August 1995–2013, while section 5b describes errors associated with predicting transition events.

a. Climatological skill

The ACC results between the reforecast and verification provide a demonstration of reforecast skill over a

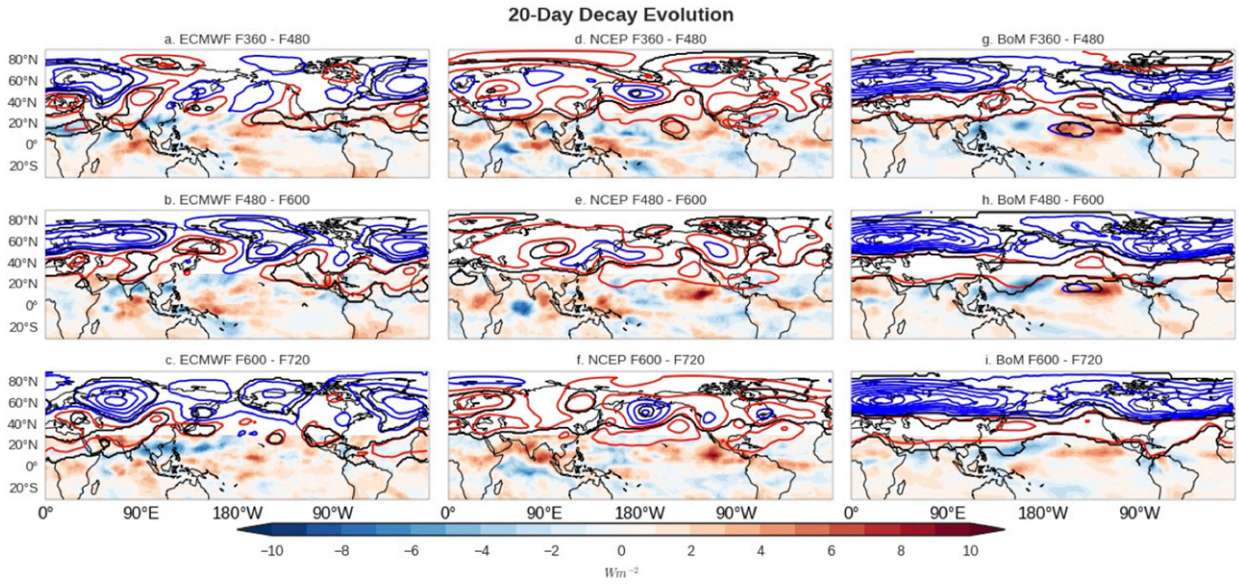

FIG. 6. Composite mean Z200 (contours) and NRAD TOA (filled contours) reforecasts for (a)–(c) ECMWF, (d)–(f) NCEP, and (g)–(i) BoM. The Z200 contour intervals start at +5 m (red) and −5 m (blue). OLR anomalies tend to be significant at ±2 W m^{-2}.

given area. ACC is calculated for Z200 over North America, from 30° to 70°N and from 135° to 45°W for the total field and each filtered band. A mean ACC is then calculated to approximate the model performance in this region. The 95% confidence interval for Z200 ACC is calculated by drawing randomly from initialization dates throughout the period of interest, 1000 times with replacement, sorting, and taking the 25th and 975th values. During this period, ECMWF retains the highest ACC throughout all leads, although NCEP ACC is nearly identical through leads F000–F240 (Fig. 3a). BoM exhibits the lowest ACC throughout the reforecast period (Fig. 3a). Both ECMWF and NCEP drop below 0.2 by lead F360–F480, whereas BoM drops below 0.2 at F240–F360 (Fig. 3a).

The ECMWF low-frequency reforecasts have the highest ACC totals throughout all lead windows, dropping below 0.2 after lead F480–F600, which is a full lead window after the BoM and nearly a full lead window after NCEP (Fig. 3b). NCEP and ECMWF ACCs are however nearly identical through lead F240–F360, after which they diverge (Fig. 3b). BoM low-frequency reforecasts remain a distant third throughout all lead windows (Fig. 3b). For all three models, the low-frequency reforecasts are more skillful than the total field, but the degradation of low-frequency ACC demonstrates the models' struggled to persist the low-frequency background conditions of the varying intraseasonal flow in which the reforecasts occur.

The ACCs in the intraseasonal band are also slightly higher than the total field (Fig. 3c). The ECMWF and NCEP models have nearly identical ACCs in the intraseasonal band, diverging after lead F240–F360, while BoM continues to have a lower ACC (Fig. 3c). Still, the low intraseasonal ACC shows that all models struggle with predicting intraseasonal variability. NCEP has the highest ACC in the synoptic band through F240–F360, although ECMWF is only about 0.10 lower, and the raw scores are low overall (Fig. 3d). As expected, the synoptic band experiences the sharpest decrease in skill, with ACC following below 0.2 in ECMWF and NCEP just prior to F240–F360 (Fig. 3d).

b. Evaluation of model performance during agricultural drought transition periods

Each of the model's composite mean reforecast verification is slightly different (Fig. 4). These differences arise from models having different initialization dates throughout the study period. Overall, the composite mean Z200 and NRAD TOA anomalies are in phase, but the amplitudes of the anomalies are different for both 20-day onset (Fig. 4, left) and 20-day decay events (Fig. 4, right), as well as 60-day transition events, which are roughly consistent with the 20-day transition events (not shown). These differences are not statistically significant and the patterns are consistent with those discussed by Schiraldi and Roundy (2017). To account for these differences, errors discussed are computed from each model's respective verification. This section focuses primarily on leads of F360–F720 where the Z200 ACCs over North America are the lowest for all three models in all three filtered bands (Fig. 3).

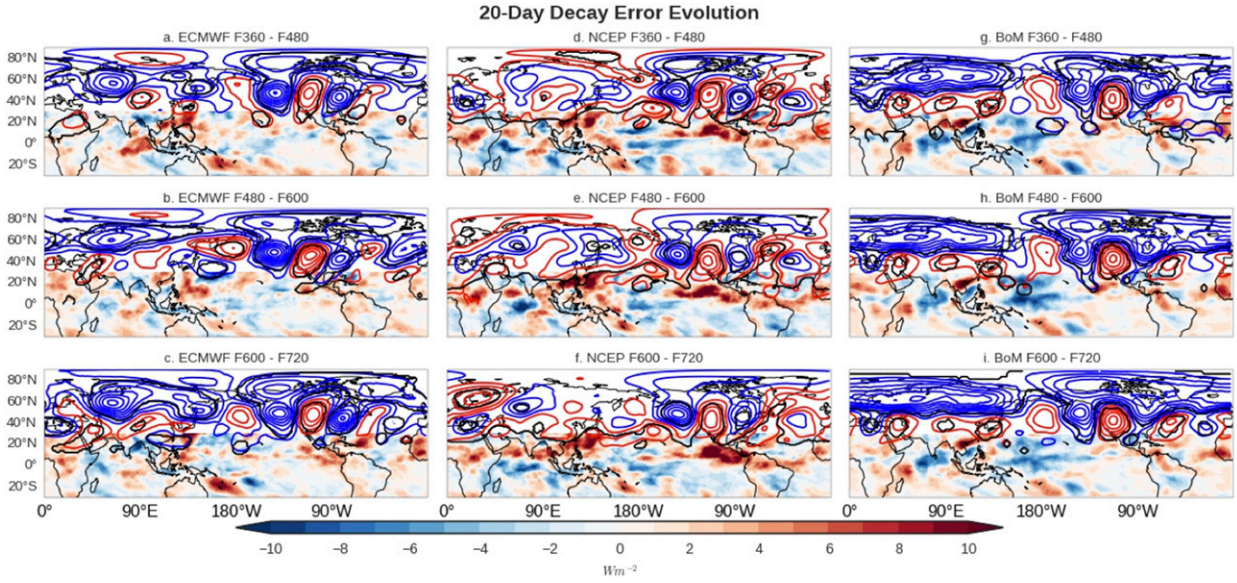

FIG. 7. Composite mean Z200 (contours) and NRAD TOA (filled contours) reforecast errors for (a)–(c) ECMWF, (d)–(f) NCEP, and (g)–(i) BoM. The Z200 contour intervals start at +5 m (red) and −5 m (blue), with an interval of 10 m. OLR anomalies tend to be significant at $\pm 2\ \mathrm{W\,m^{-2}}$.

1) COMPARISON WITH HISTORICAL ERRORS

It is important to consider whether the ACCs associated with agricultural drought transition events over the CBR are statistically different from climatology. As such, ACC reference scores (ACC-RSs) are computed with respect to the ACC of a randomly selected distribution of dates during May–August (MJJA). Positive ACC-RSs indicate better reforecast performance with respect to any given reforecast during MJJA, while negative ACC-RSs represent poorer reforecast performance. The ACC-RSs are considered different if the 95% confidence interval of the average ACCs over the set of events falls outside of the 95% confidence interval

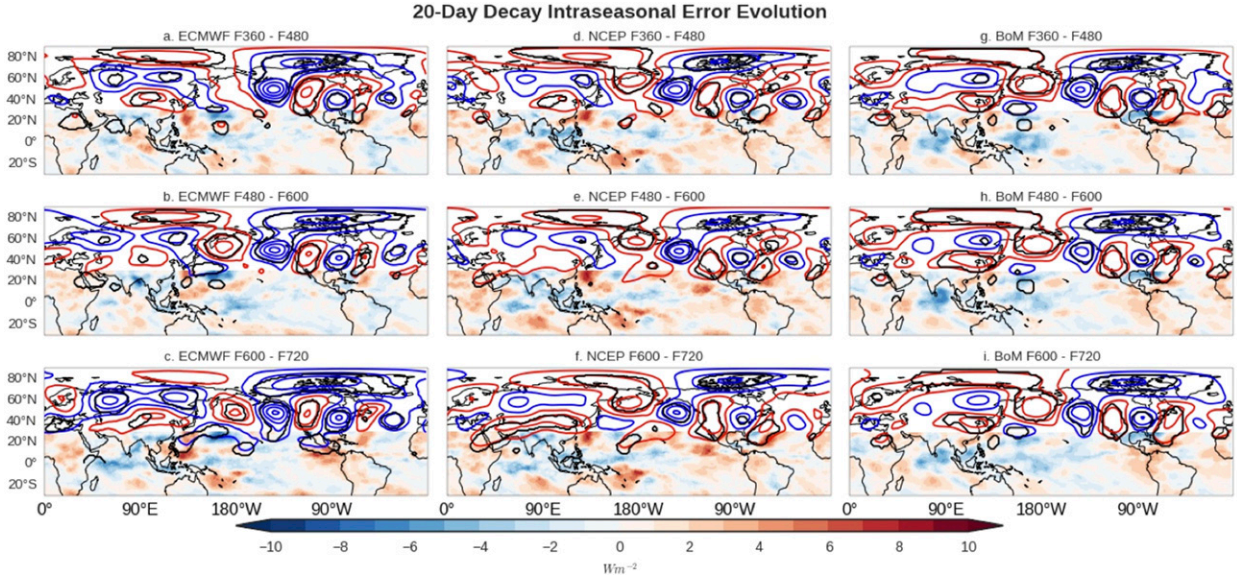

FIG. 8. Composite mean Z200 (contours) and NRAD TOA (filled contours) reforecasts errors in the intraseasonal band for (a)–(c) ECMWF, (d)–(f) NCEP, and (g)–(i) BoM. The Z200 contour intervals start at +5 m (red) and −5 m (blue), with an interval of 10 m. OLR anomalies tend to be significant at $\pm 2\ \mathrm{W\,m^{-2}}$.

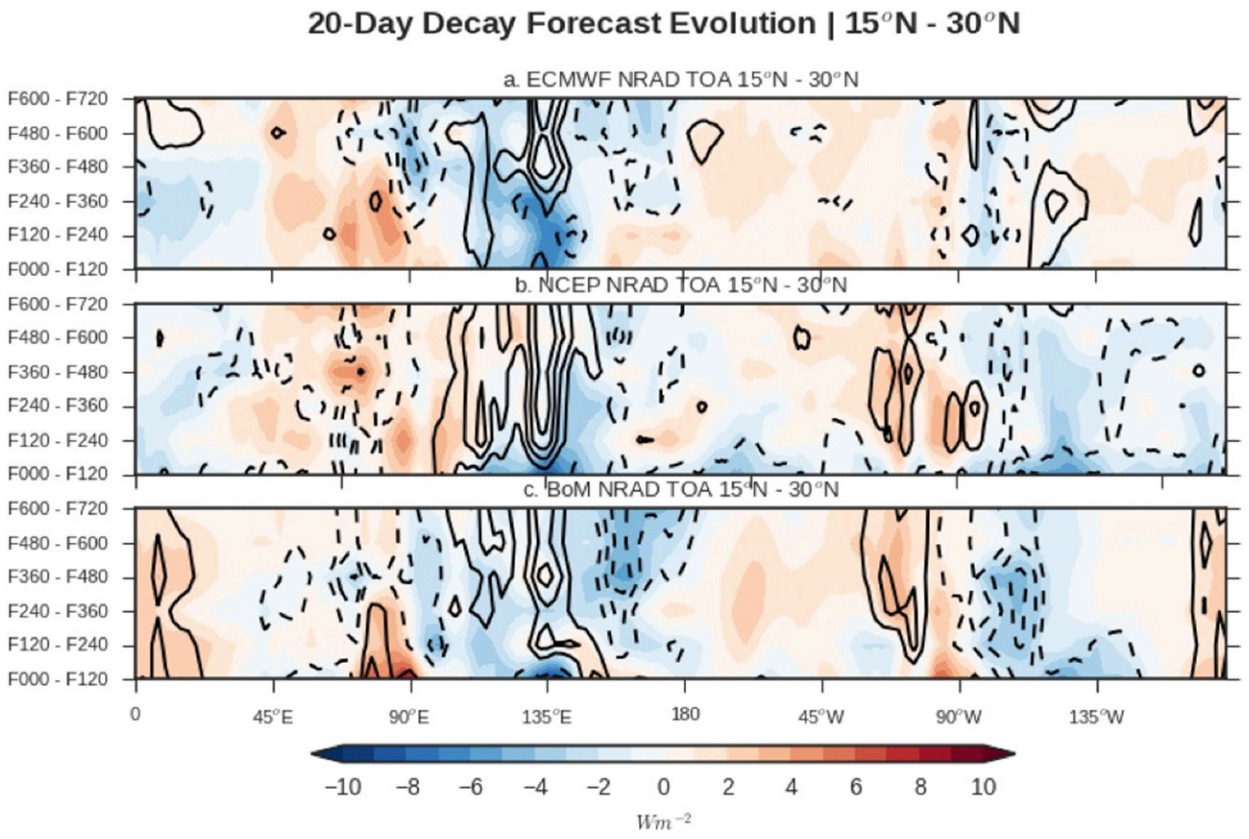

FIG. 9. Time–longitude diagram detailing the composite mean NRAD TOA reforecast (filled contours) for (a) ECMWF, (b) NCEP, and (c) BoM and errors (contours). Positive errors (suppressed bias) are presented as solid lines starting at $+2\,\mathrm{W\,m^{-2}}$ with a contour interval of $2\,\mathrm{W\,m^{-2}}$, where dashed contours indicate negative errors (active bias) starting at $-2\,\mathrm{W\,m^{-2}}$ with a contour interval of $-2\,\mathrm{W\,m^{-2}}$. NRAD TOA values tend to be significant at $\pm 2\,\mathrm{W\,m^{-2}}$.

of the ACCs for climatology. This test is performed on the total field and filtered fields separately to assess whether differences in one band dominate differences in the total ACC. After F240, when the difference in the unfiltered field is significant for NCEP, the synoptic and intraseasonal differences tend to be significant (Figs. 5a,d,i,l), with better performance than average in predicting 20- and 60-day drought decay events, at F360–F480. ECMWF has lower ACC-RSs at F480–F600 during 20-day drought onset events, in the total and intraseasonal fields (Figs. 5e,g), whereas BoM has higher ACC-RSs predicting 60-day drought decay at F480–F600, with no dominant contribution from either filtered field (Fig. 5i).

This analysis shows that enhanced or decaying ACC-RSs at one lead window do not transfer to enhanced ACC-RSs at lead windows closer to verification, and it is difficult to use this analysis to assess whether one filtered band is contributing most to these changes. To further assess this question, composite mean reforecasts and errors for each filtered band are used to assess the errors associated with the patterns Schiraldi and Roundy (2017) determined to have the largest influence on these transition periods.

2) 20-DAY AGRICULTURAL DROUGHT DECAY

Reforecasts of Z200 and NRAD TOA during 20-day drought decay periods differ from model to model, but there are some consistent themes. At F600–F720, the ECMWF composite ensemble mean Z200 shows a negative height anomaly over Alaska, with a positive height anomaly centered over Texas (Fig. 6c). In addition, the enhanced convection at approximately 20°N extends from 100° to approximately 150°E (Fig. 6c). This convection builds a ridge on its poleward side through the advection of potential vorticity by the divergent wind, yielding an extended subtropical Pacific jet (not shown), which is not consistent with the pattern evolution identified by Schiraldi and Roundy (2017). Further, the convection is more expansive than the verification (Fig. 4), and much of the amplitude of this convection is seen in the synoptic band (not shown). During F480–F600, ECMWF starts to decrease the intensity of the

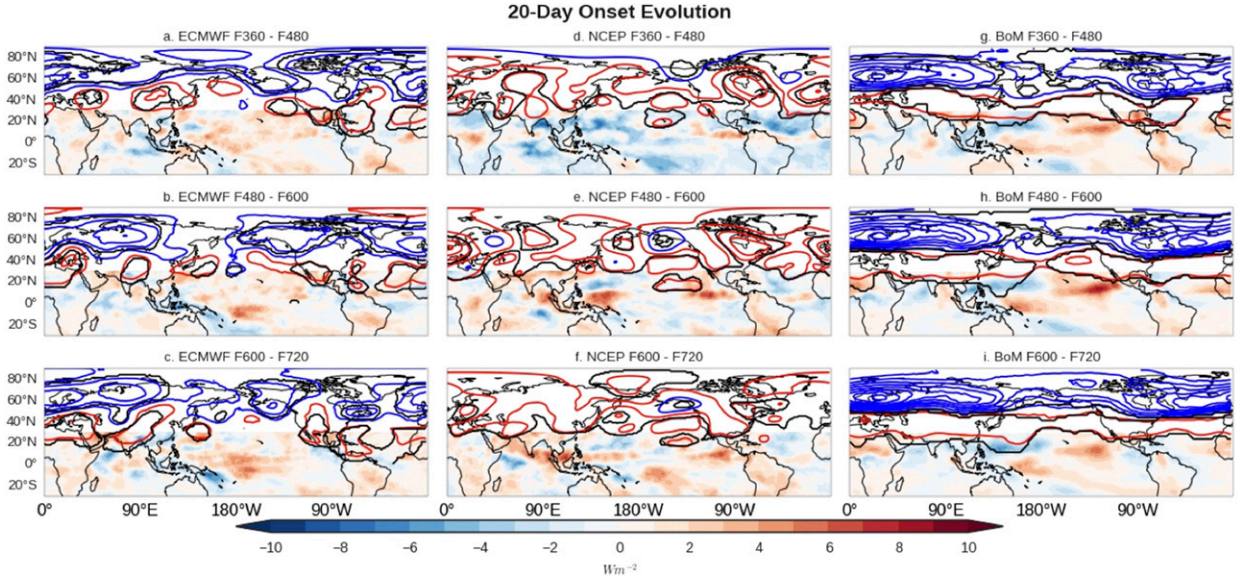

FIG. 10. Composite mean Z200 (contours) and NRAD TOA (filled contours) reforecasts for (a)–(c) ECMWF, (d)–(f) NCEP, and (g)–(i) BoM. The Z200 contour intervals start at +5 m (red) and −5 m (blue). OLR anomalies tend to be significant at ±2 W m^{-2}.

convection in this region. The large-scale circulation pattern over the Pacific and North America during this lead window is 180° out of phase with the verification. The error in the phase of the wave train can be linked to an area of positive errors in NRAD TOA, near 135°W from 0° to 30°N (Fig. 7b).

A bias toward suppressed convection near 135°W from 0° to 30°N yields a negative Z200 bias on its poleward side and positive Z200 height bias farther poleward (Fig. 7b). Much of the amplitude of the NRAD TOA error occurs in the intraseasonal band suggesting that the intraseasonal propagation of the region of convection is not being accurately reforecast (Fig. 8b). A time–longitude diagram of NRAD TOA averaged from 15° to 30°N shows the dry bias near 135°W in ECMWF (Fig. 9a) at leads after F360. Sensitivity testing in the model would be

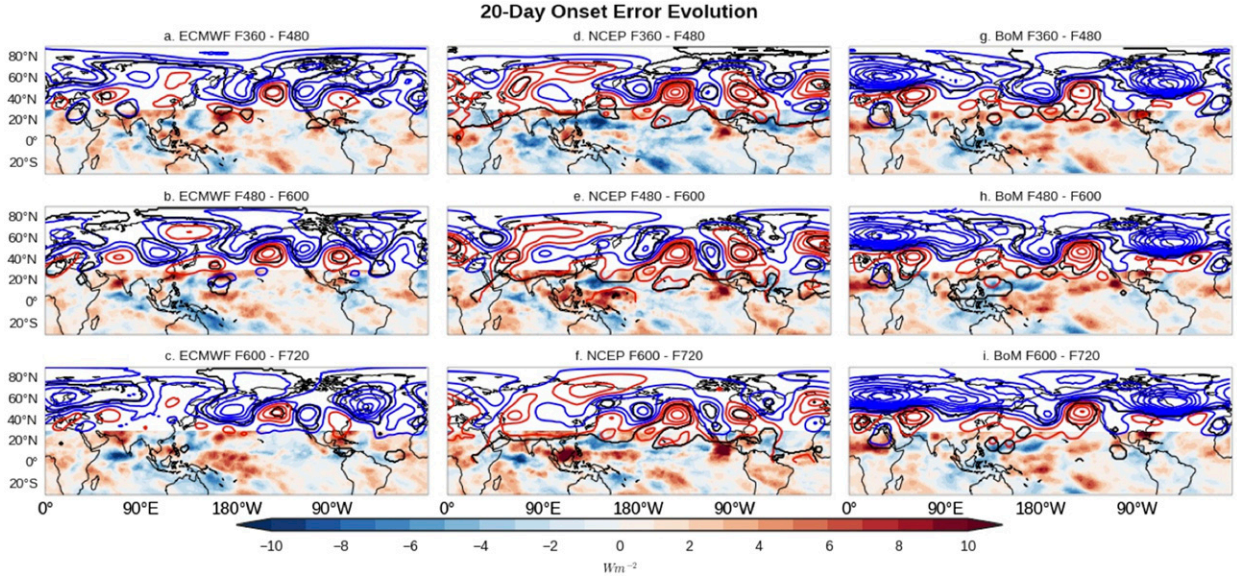

FIG. 11. Composite mean Z200 (contours) and NRAD TOA (filled contours) reforecasts errors for (a)–(c) ECMWF, (d)–(f) NCEP, and (g)–(i) BoM. The Z200 contour intervals start at +5 m (red) and −5 m (blue), with an interval of 10 m. OLR anomalies tend to be significant at ±2 Wm^{-2}.

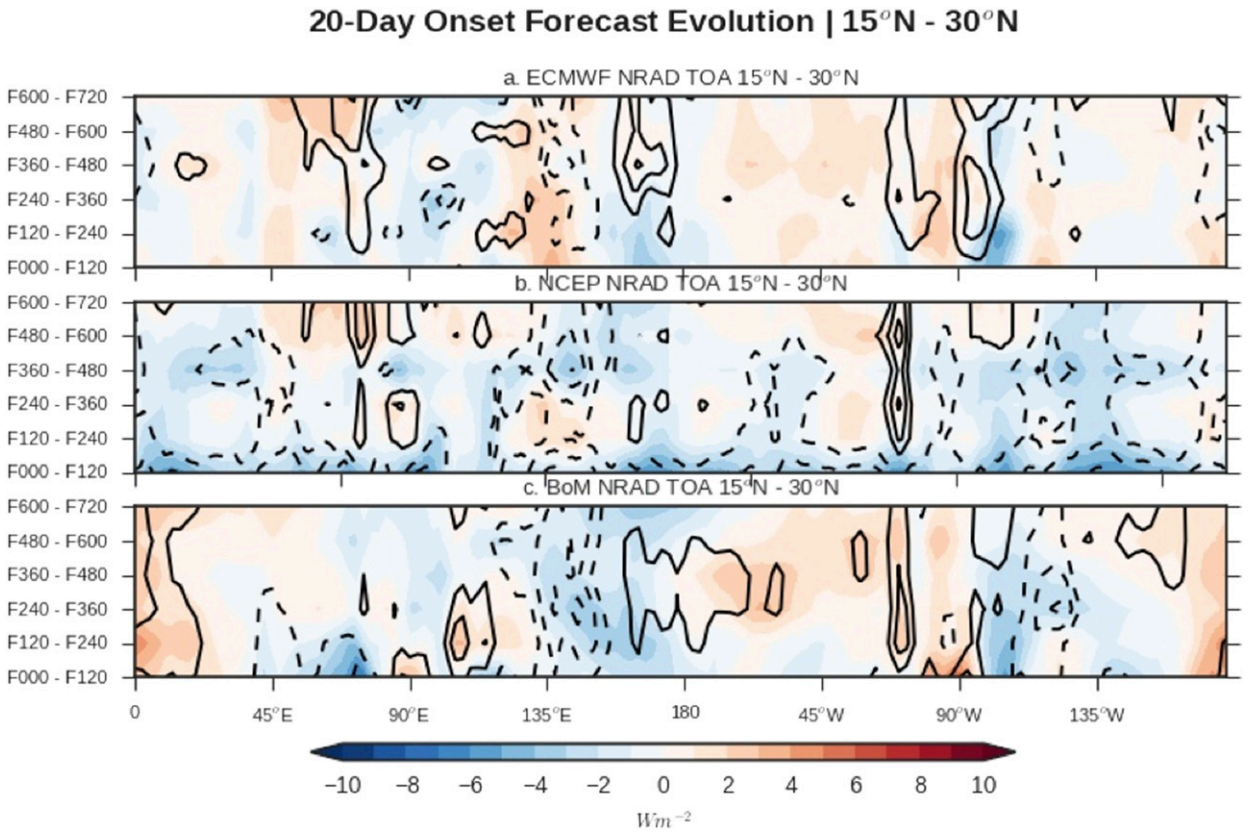

FIG. 12. Time–longitude diagram detailing the composite mean NRAD TOA reforecast (filled contours) for (a) ECMWF, (b) NCEP, and (c) BoM and errors (contours). Positive errors (suppressed bias) are presented as solid lines starting at $+2\,\mathrm{W\,m^{-2}}$ with a contour interval of $2\,\mathrm{W\,m^{-2}}$, where dashed contours indicate negative errors (active bias) starting at $-2\,\mathrm{W\,m^{-2}}$ with a contour interval of $-2\,\mathrm{W\,m^{-2}}$. NRAD TOA values tend to be significant at $\pm 2\,\mathrm{W\,m^{-2}}$.

required to demonstrate the source of the dry bias and whether it is driving errors in the extratropical circulation or is a result of errors in that circulation. Such testing is beyond the scope of this study. NCEP and BoM have a similar dry bias near 135°E (Figs. 9b,c), but their Z200 circulation patterns are different (Figs. 6d–i).

NCEP reforecasts of 20-day drought decay events are generally consistent across lead windows F360–F720. These lead windows are characterized by a statistically significant negative Z200 anomaly centered near 50°N, 180°, with a positive Z200 anomaly on its equatorward and poleward sides (Figs. 6d–f). The signal is not very clear in the composite Z200 reforecasts over North America, indicating large event-to-event differences and no obvious Z200 bias associated with these events (Figs. 6d–f). There are however significant errors in the verifying wave train over the United States (Figs. 7d–f). There are significant biases in NRAD TOA, especially near 135°E, where the convection is largely suppressed in all three lead windows (Figs. 7d–f). Like ECMWF, much of the amplitude of these errors is accounted for by errors in the intraseasonal field (Figs. 8d–f). Unlike ECMWF, the dry bias in NCEP persists from F120–F240 through F720, which may be linked to poor skill at earlier leads compared to ECMWF (Fig. 9b). At F120–F240, the intraseasonal extratropical Rossby wave is out of phase with the verification, which may also be linked to this dry bias (not shown).

The BoM Z200 reforecast is also relatively consistent through leads F360–F720, yielding a negative Z200 height anomaly from 40° to 60°N across the globe, with positive Z200 anomalies on its poleward and equatorward sides (Figs. 6g–i). BoM has a dry bias in NRAD TOA near 135°E in the tropics, but it is characterized by a more pronounced wet bias east and west of 135°E (Fig. 9c). It is difficult to determine what is driving the extratropical circulation in BoM, and the results suggest the need for sensitivity testing in the model.

3) 20-DAY AGRICULTURAL DROUGHT ONSET

During 20-day agricultural drought decay events, intraseasonal NRAD TOA and Z200 patterns associated with 20-day agricultural onset events are not well predicted by any model. ECMWF predicts a positive

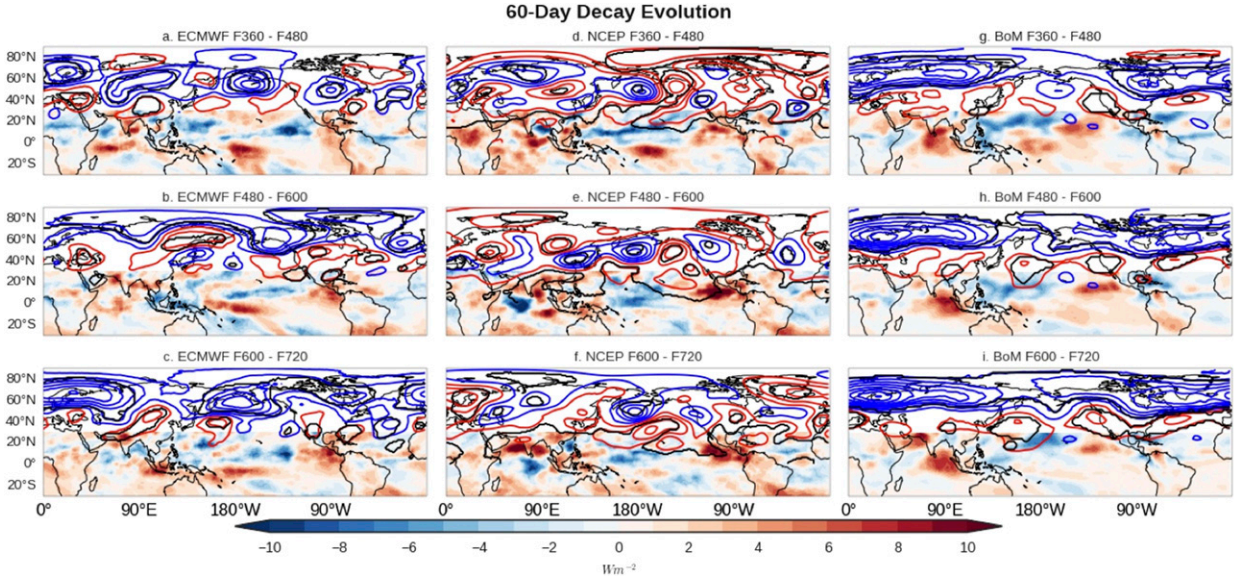

FIG. 13. Composite mean Z200 (contours) and NRAD TOA (filled contours) reforecasts for (a)–(c) ECMWF, (d)–(f) NCEP, and (g)–(i) BoM. The Z200 contour intervals start at +5 m (red) and −5 m (blue). OLR anomalies tend to be significant at ±2 W m^{-2}.

Z200 anomaly over the Rocky Mountains with negative height anomalies downstream centered over the northeastern United States, and upstream centered over Alaska at F600–F720 (Fig. 10c). The large-scale pattern is generally representative of the extratropical Rossby wave state over North America during 20-day drought onset events (Fig. 3a). While Z200 has relatively low errors over the United States (Fig. 11c), the large-scale pattern does not progress through to shorter lead times (Figs. 10a,b), leading to higher errors at leads F360–F600 (Figs. 11a,b). Opposite to the 20-day agricultural decay events, ECMWF now has a relative wet bias near 135°E during lead F240–F600 (Figs. 11a–c). NCEP has no clear favored Z200 reforecast state, and BoM persists a circumglobal trough between 40° and 60°N (Figs. 10d–i). Like ECMWF, NCEP and BoM show a general wet bias near 135°E, which may play some role in the errors associated with these events (Fig. 12). Much of the

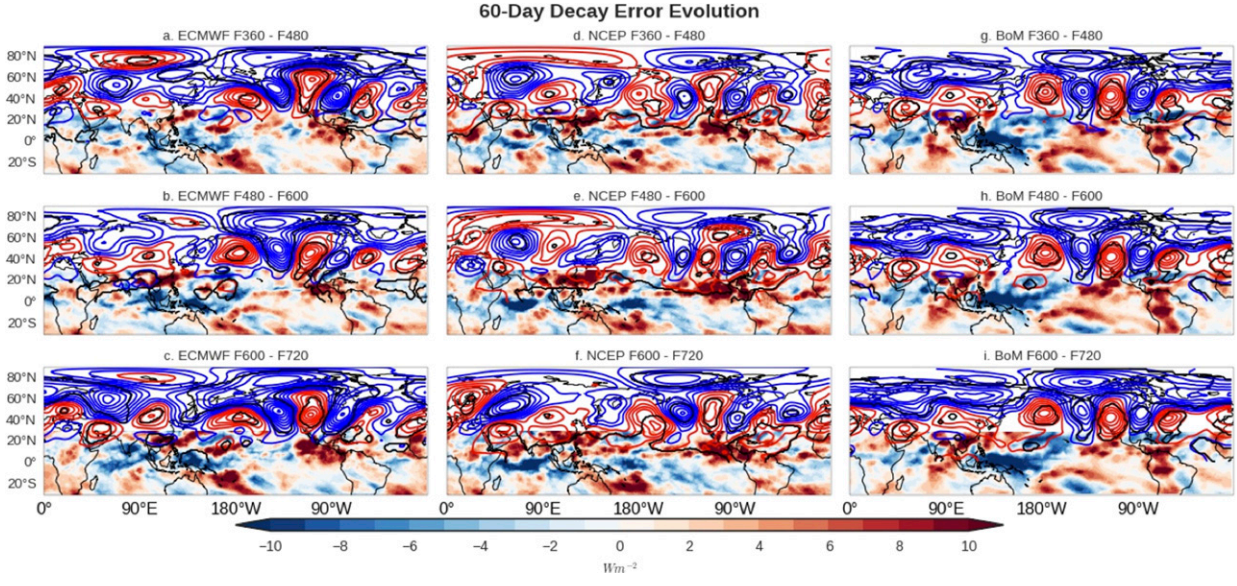

FIG. 14. Composite mean Z200 (contours) and NRAD TOA (filled contours) reforecasts errors for (a)–(c) ECMWF, (d)–(f) NCEP, and (g)–(i) BoM. The Z200 contour intervals start at +5 m (red) and −5 m (blue), with an interval of 10 m. OLR anomalies tend to be significant at ±2 W m^{-2}

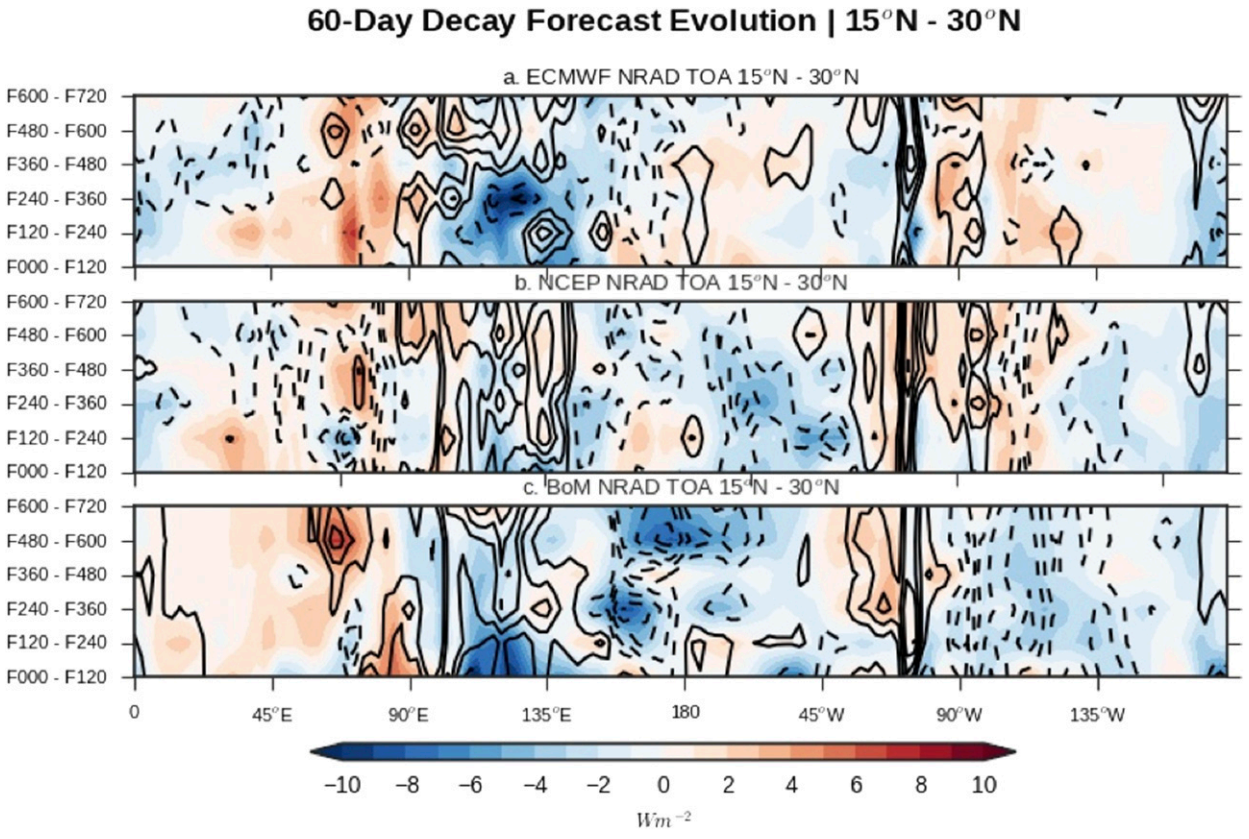

FIG. 15. Time–longitude diagram detailing the composite mean NRAD TOA reforecast (filled contours) for (a) ECMWF, (b) NCEP, and (c) BoM and errors (contours). Positive errors (suppressed bias) are presented as solid lines starting at $+2\,\mathrm{W\,m^{-2}}$ with a contour interval of $2\,\mathrm{W\,m^{-2}}$, where dashed contours indicate negative errors (active bias) starting at $-2\,\mathrm{W\,m^{-2}}$ with a contour interval of $-2\,\mathrm{W\,m^{-2}}$. NRAD TOA values tend to be significant at $\pm 2\,\mathrm{W\,m^{-2}}$.

amplitude of the errors in all three models is accounted for in the intraseasonal field (not shown).

4) 60-DAY AGRICULTURAL DROUGHT DECAY

ECMWF and NCEP do a much better job at predicting the large-scale Z200 pattern associated with 60-day agricultural drought decay transitions compared to 20-day transitions. At lead F600–F720, ECMWF accurately positions a negative Z200 anomaly over Alaska, but does not correctly phase the wave train over the United States (Fig. 13c). The negative Z200 anomaly over Alaska is shifted slightly poleward, and its zonal extent is wider than that of the verification (not shown). Eighty percent of the amplitude of this negative height anomaly is explained by the intraseasonal band (not shown). NCEP reforecasts a negative 200-hPa anomaly in a similar region, and phases the wave correctly over North America, although only the ridge over the eastern half of the United States is statistically significant (Fig. 13f), implying reduced consistency across its ensemble. Similar to ECMWF, much of the amplitude of this signal is accounted for in the intraseasonal band (not shown). Like 20-day decay cases, BoM Z200 reforecasts continue to persist a circumglobal trough poleward of 30°N during *60-day drought decay* (Fig. 13i). ECMWF, NCEP, and BoM continue to have a bias in NRAD TOA toward suppressed convection at 135°E, from 15° to 30°N (Figs. 14c,f,i), with much of the amplitude of these errors accounted for by the intraseasonal band (not shown).

At F480–F600, ECMWF shifts the position of the negative Z200 anomaly over Alaska slightly east (Fig. 13b). ECMWF also builds positive Z200 anomalies over the central United States, which dynamically supports suppressed rainfall over the CBR and continued drought conditions (Fig. 13b). NCEP, however, maintains the negative Z200 anomalies over Alaska in the same region during F480–F600, with positive Z200 anomalies in the Gulf of Alaska, and negative Z200 anomalies over the western half of the United States (Fig. 13e). Thus, NCEP correctly suggests a transition away from drought at lead F480–F600. NCEP correctly

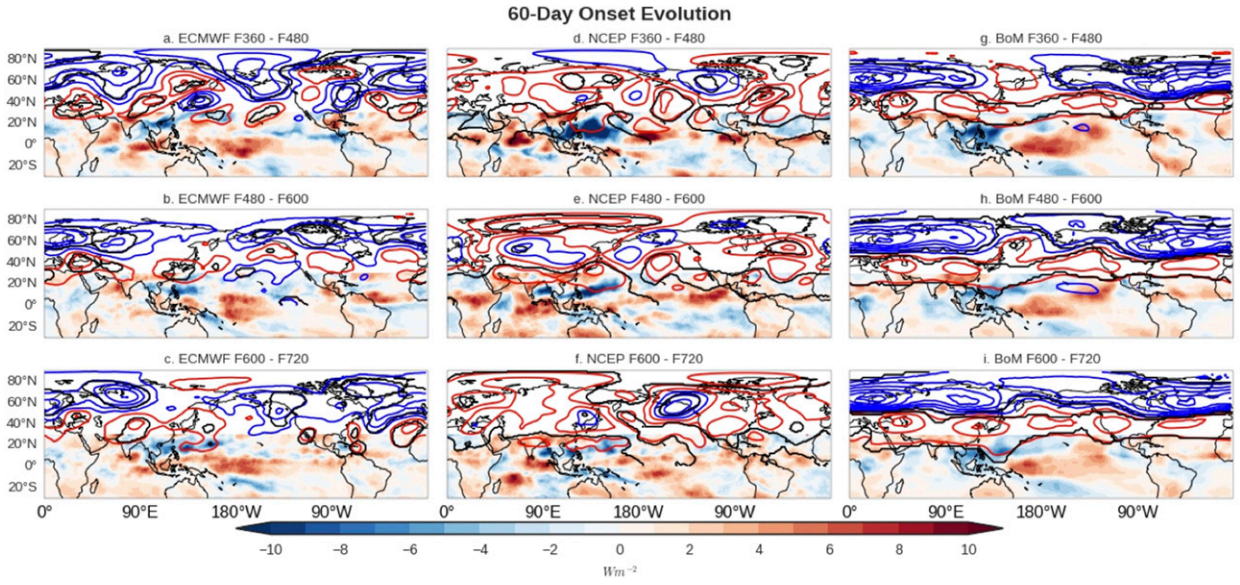

FIG. 16. Composite mean Z200 (contours) and NRAD TOA (filled contours) reforecasts for (a)–(c) ECMWF, (d)–(f) NCEP, and (g)–(i) BoM. The Z200 contour intervals start at +5 m (red) and −5 m (blue), with increments every 10 m. OLR anomalies tend to be significant at $\pm 2\ \mathrm{W\ m^{-2}}$.

phases the extratropical Rossby wave, despite having a dry bias near 135°E in the tropics. Further, much of the extratropical Rossby wave amplitude is in the intraseasonal band (not shown). At F480–F600, BoM reforecasts do not materially change.

By F360–F480, ECMWF and NCEP are consistent with the large-scale Z200 pattern governing the transition. ECMWF shifts the negative Z200 anomaly over Alaska to the west (Fig. 13a). ECMWF also builds a trough over the central plains and a ridge over the southeastern United States (Fig. 13a). While there are still substantial errors with this reforecast (Fig. 14a), in general, the results agree with patterns associated with a shift away from drought over the CBR. ECMWF accurately predicts convection near 135°E in the tropics (Fig. 15a), which appears to play a role in organizing the extratropical Rossby wave train over the United States. The NCEP F360–F480 Z200 circulation pattern does not change substantially from the F480–F600 reforecast over the Pacific (Fig. 13d), and continues to reforecast a pattern that supports a transition away from drought over the CBR. Unlike ECMWF, NCEP continues to have a dry bias near 135°E in the tropics at this lead (Fig. 15b). Once again, the BoM reforecast does not change much during this lead (Fig. 13g).

5) 60-DAY AGRICULTURAL DROUGHT ONSET

ECMWF and NCEP also have lower reforecast errors associated with 60-day agricultural drought onset, relative to 20-day agricultural drought onset. At F600–F720, ECMWF accurately predicts negative Z200 anomalies over Alaska, although the composite reforecast does not have a clear signal over the United States (Fig. 16c). At this lead, NCEP places negative Z200 anomalies in a similar region, but it too does not predict a coherent pattern over the United States (Fig. 16f). At F480–F600, ECMWF establishes largely zonal flow across the United States and Gulf of Alaska, shifting the negative Z200 anomalies west and east, with a minimum over the Bering Strait, and another over northern Canada (Fig. 16b). At F480–F600, NCEP features a positive Z200 anomaly over the central Pacific and another south of Greenland (Fig. 16e). Neither of these features is linked to the circulation anomalies identified by Schiraldi and Roundy (2017). At F360–F480, ECMWF has the correct phase of an extratropical Rossby wave extending from the Pacific to the Atlantic. That is, it reforecasts negative Z200 anomalies over Alaska, positive Z200 anomalies (although they are not statistically significant) over the western half of the United States, and a trough over the eastern half of the United States. The largest difference in this reforecast from the other leads is in the tropics, where there is a region of suppressed convection near 135°E between 15° and 30°N (Fig. 17a). The ECMWF reforecast at F360–F480, is much better than that of NCEP. NCEP features an expansive area of negative NRAD TOA anomalies from 90°E to 180° (Fig. 17b), and an incorrectly phased wave train over the United States (Fig. 16c). The zonal flow over the United States implied by NCEP would

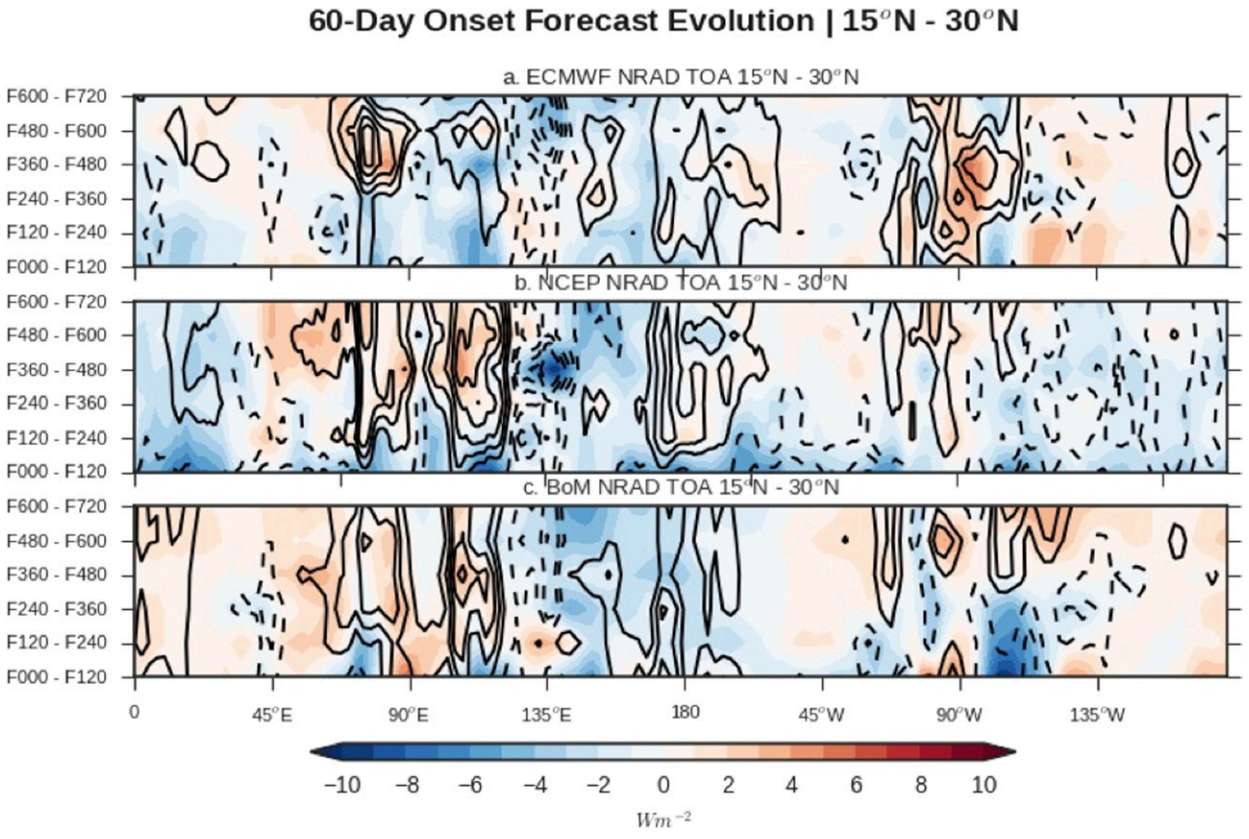

FIG. 17. Time–longitude diagram detailing the composite mean NRAD TOA reforecast (filled contours) for (a) ECMWF, (b) NCEP, and (c) BoM and errors (contours). Positive errors (suppressed bias) are presented as solid lines starting at $+2\,\mathrm{W\,m^{-2}}$ with a contour interval of $2\,\mathrm{W\,m^{-2}}$, where dashed contours indicate negative errors (active bias) starting at $-2\,\mathrm{W\,m^{-2}}$ with a contour interval of $-2\,\mathrm{W\,m^{-2}}$. NRAD TOA values tend to be significant at $\pm 2\,\mathrm{W\,m^{-2}}$.

dynamically favor suppressed rainfall over the region. Throughout the reforecast horizon, BoM is similar to the other transition events, reforecasting circumglobal positive Z200 anomalies in the tropics and negative Z200 anomalies in the extratropics (Figs. 16g–i).

6. Reforecast model wave characteristics

It is apparent that ECMWF, NCEP, and BoM struggle fundamentally with predicting the circulation patterns (both NRAD TOA and Z200) associated with 20- and 60-day agricultural drought transition periods. It remains difficult to assess whether this is due to fundamental flaws in the model physics or just a lack of predictability of such events, without careful model sensitivity testing. Spectrum analysis of reforecast data compared with similar spectra generated from the ERA-I may provide insights into biases in model variability that might explain some of the reforecast errors. This section discusses biases in the wavenumber frequency domain from 30° to 60°N in Z200 and from 0° to 30°N in NRAD TOA. The wavenumber frequency spectra are calculated by applying Eq. (6) in time, and a conventional Fourier decomposition in space. This calculation is repeated on the ERA-I Z200 and satellite OLR fields, with imposed irregular time stepping equivalent to the models' initialization frequency and in the model reforecast fields at F360–F480, for target dates in May through August, in each ensemble member. The average of the spectra of individual ensemble members is presented. The difference between the ERA-I Z200, satellite OLR, and reforecast spectra are presented for NRAD TOA in section 6a and Z200 in section 6b.

a. NRAD TOA

Throughout the target study period, NCEP, ECMWF, and BoM show large differences in spectral power between the reforecast and analyzed OLR results. The largest differences in ECMWF occur in westward-propagating waves, from wavenumbers 5 to 15 and for periods greater than 15 days, where ECMWF exhibits more power than the satellite OLR (Fig. 18a). Outside of those bands, ECMWF has less power relative to the satellite OLR data. ECMWF exhibits a decrease in power in

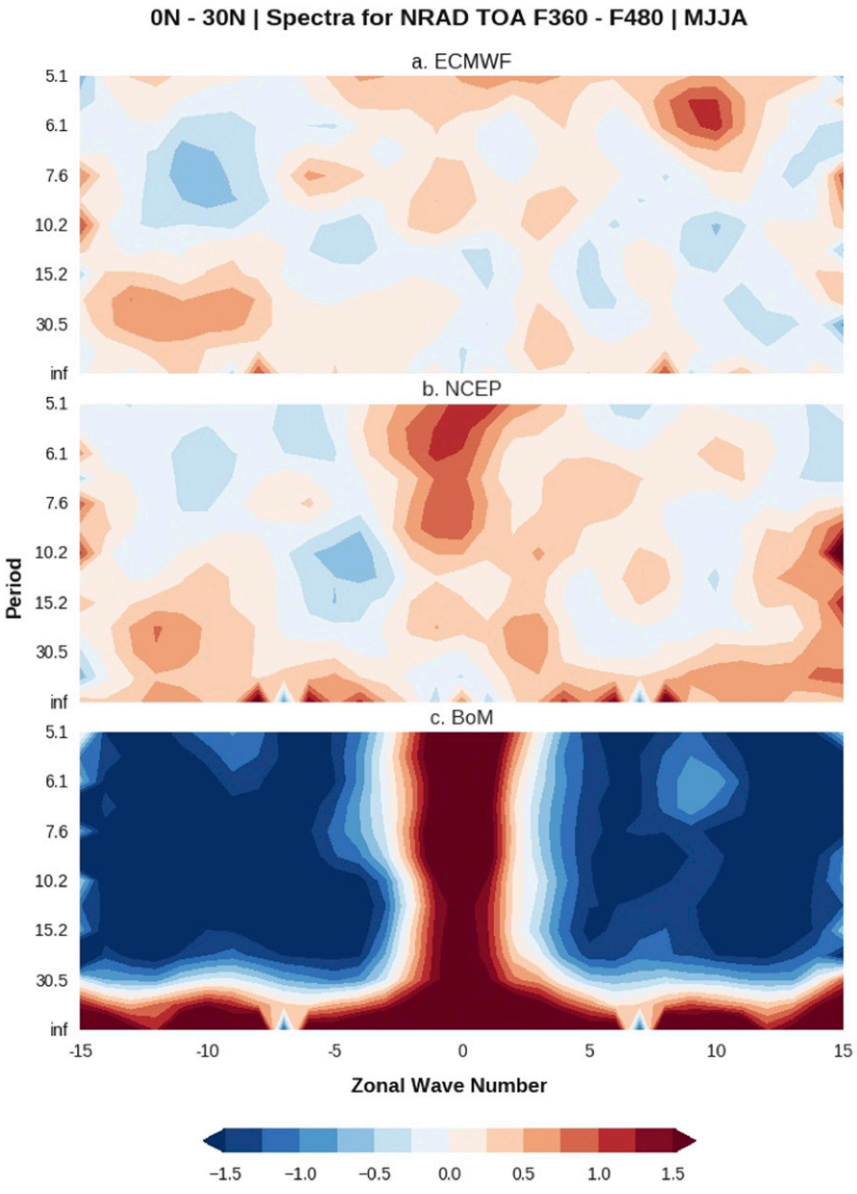

FIG. 18. The F360–F480 difference in wave spectra from reanalysis for NRAD TOA averaged from 0° to 30°N for (a) ECMWF, (b) NCEP, and (c) BoM. Differences tend to be significant at ±0.5.

westward-propagating waves with periods of 5–10 days between wavenumbers 6 and 15 (Fig. 18a). To the extent that tropical waves in this spectral band play a role in agricultural drought transition periods, the decrease in frequency of these waves in the reforecast may be linked to poor model performance during these events.

NCEP's spectral differences occur at different peaks than ECMWF. Overall, NCEP is subject to more power in eastward-propagating waves. NCEP has considerably more power in the spectral band containing wavenumbers 0–5 and periods of from 6 to 30 days (Fig. 18b). Kelvin waves are often characterized by this spectral band (Wheeler and Kiladis 1999), as are extratropical Rossby waves advected eastward on westerly winds. The average latitudinal variance associated with NRAD TOA in this spectral band was plotted to establish that the variance was concentrated equatorward of 15°N, confirming that this increase in power is likely associated with Kelvin waves as opposed to an extratropical signature. NCEP has a decrease in power in westward-propagating wavenumbers 2–5 for periods from 7 to 15 days (Fig. 18a).

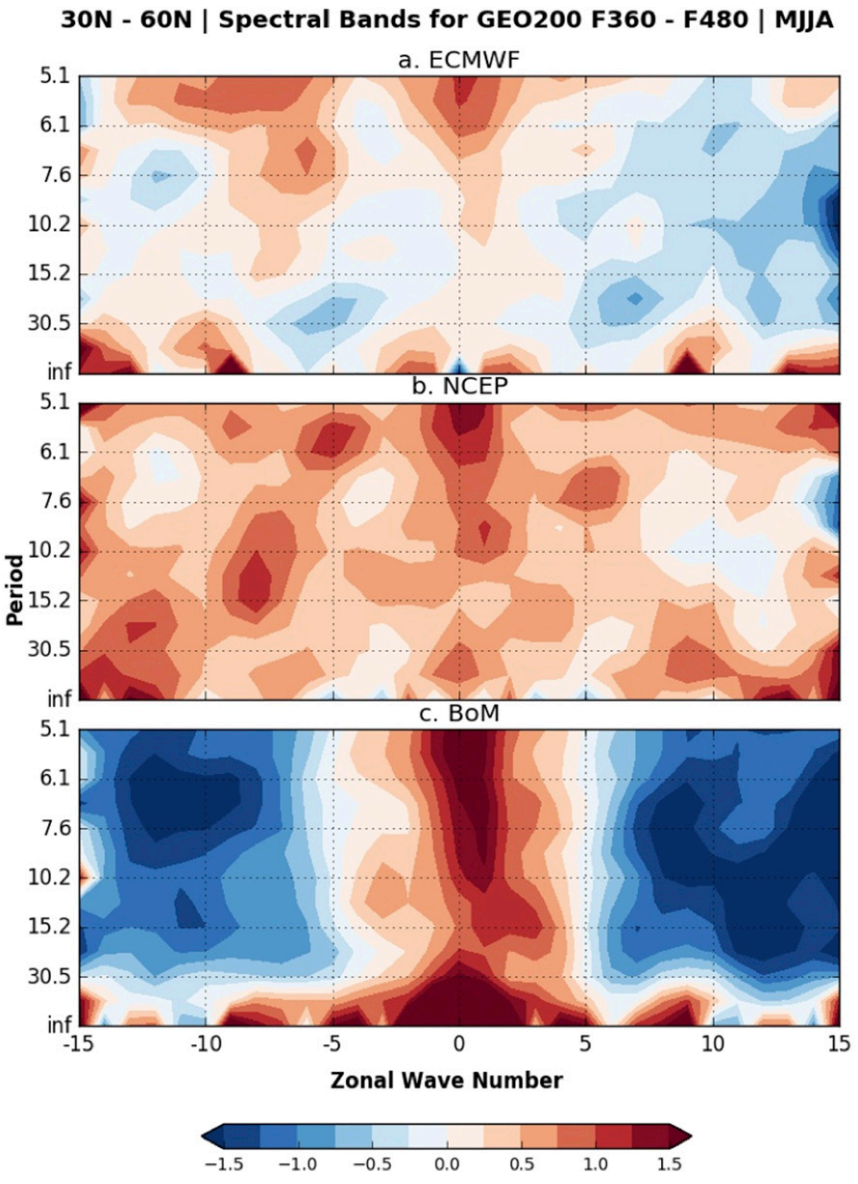

FIG. 19. The F360–F480 difference in wave spectra from reanalysis for Z200 averaged from 30° to 60°N for (a) ECMWF, (b) NCEP, and (c) BoM. Differences tend to be significant at ±0.5.

BoM has the most unique difference from the analyzed OLR of any of the three models. BoM tends to have more power in high wavenumbers and during low-frequency periods, as well as for low wavenumbers and all periods. Outside of those bands, BoM has less power relative to the satellite OLR data.

b. Z200

Shifts in ECMWF Z200 power spectra averaged from 30° to 60°N at F360–F480 reveal reduced eastward-propagating waves from zonal wavenumbers 5–15 and periods of from 6 to 30 days (Fig. 19a). This result suggests that ECMWF does not produce enough synoptic and intraseasonal variability in the extratropical waveguide. The extratropical wave train associated with agricultural drought transition periods is contained in this spectral band, and thus a decrease in variability in this band could be related to a decrease in reforecast performance. NCEP too shows a decrease in power in a similar band, although it is constrained to wavenumbers 10–15 and periods of from 6 to 15 days (Fig. 19b). This reduced power suggests that the model does not sufficiently simulate synoptic variability, which is considerably less important than intraseasonal variability to

these transition events. NCEP has much more power in westward-propagating waves of nearly all frequencies and wavenumbers, which would be interesting to look at in future work (Fig. 19b). Again, BoM tends to have more power in high wavenumbers and during low-frequency periods as well as for low wavenumbers and synoptic periods. Outside of those bands, BoM has less power than in analysis fields.

7. Discussion

The seasonal cycle of bias associated with ECMWF, BoM, and NCEP reveals that biased circulation patterns in Z200 and NRAD TOA are likely tied to each model's ability to predict agricultural drought transition periods. Each model tends to revert to its climatological state over North America in as little as three weeks, as demonstrated by a sharp drop in ACCs for the low-frequency band (Fig. 3b). Since the ACCs presented in this paper are computed on bias-corrected fields, bias correction alone is not sufficient. Results show that errors in the prediction of Z200 over North America during agricultural drought transition periods, on average, are no different than the historical error during the target study period. During MJJA, ECMWF and NCEP show similar skill in predicting Z200 heights over North America, with ACCs falling below 0.2 by lead F360–F480, and BoM results falling below 0.2 at F240–F360.

There is an increase of 0.15 in ACC compared to climatology at lead F360–F480 in NCEP during 20-day drought decay events and an increase of 0.25 during 60-day drought decay events. This increase in ACC is due to improvements in predicting the intraseasonal and synoptic bands. This result demonstrates that to improve the prediction of transition events, the numerical weather prediction community must improve the prediction of a broad basis of weather phenomena. It is likely that improvements in the low-frequency and intraseasonal bands will contribute to achieving better reforecasts, as most of the amplitude of the Z200 and NRAD TOA errors are contained in the intraseasonal band.

Our results demonstrate that while ECMWF has higher ACCs in Z200 over North America through leads F240–F360 in the total and filtered fields compared to NCEP and BoM, ECMWF does not show better skill in leads thereafter during May–August. Overall, Z200 and NRAD TOA have generally lower errors for 60-day agricultural drought onset and decay at lead F360–F480 in ECMWF and NCEP. Thus, during these types of drought onset or decay events, ECMWF and NCEP may be useful operationally in assessing when such agricultural drought transitions might occur.

Acknowledgments. We thank three anonymous reviewers for their thoughtful comments. This work is sponsored by NOAA Grant NA14OAR4310230.

REFERENCES

Adonizio, W., N. Kook, and S. Royales, 2012: Impact of the drought on corn exports: Paying the price. *Beyond the Numbers: Global Economy*, Vol. 1, No. 17, Bureau of Labor and Statistics, 7 pp. [Available online at https://www.bls.gov/opub/btn/volume-1/pdf/impact-of-the-drought-on-corn-exports-paying-the-price.pdf.]

Berrisford, P., and Coauthors, 2011: The ERA-Interim archive, version 2.0. ERA Report Series, No. 1, ECMWF, Reading, United Kingdom, 23 pp. [Available online at http://www.ecmwf.int/sites/default/files/elibrary/2011/8174-era-interim-archive-version-20.pdf.]

Chen, M., W. Shi, P. Xie, V. B. S. Silva, V. E. Kousky, R. W. Higgins, and J. E. Janowiak, 2008: Assessing objective techniques for gauge-based analyses of global daily precipitation. *J. Geophys. Res.*, **113**, D04110, doi:10.1029/2007JD009132.

ECMWF, 2016: Changes in ECMWF model: Evaluation of the IFS. [Available online at http://www.ecmwf.int/en/forecasts/documentation-and-support/changes-ecmwf-model.]

Guo, Z., P. A. Dirmeyer, and T. Delsole, 2011: Land surface impacts on subseasonal and seasonal predictability. *Geophys. Res. Lett.*, **38**, L17808, doi:10.1029/2011GL048611.

Hamill, T. M., J. S. Whitaker, and X. Wei, 2004: Ensemble reforecasting: Improving medium-range forecast skill using retrospective forecasts. *Mon. Wea. Rev.*, **132**, 1434–1447, doi:10.1175/1520-0493(2004)132,1434:ERIMFS.2.0.CO;2.

——, R. Hagedorn, and J. S. Whitaker, 2008: Probabilistic forecast calibration using ECMWF and GFS ensemble reforecasts. Part II: Precipitation. *Mon. Wea. Rev.*, **136**, 2620–2632, doi:10.1175/2007MWR2411.1.

Hoerling, M., J. Eischeid, A. Kumar, R. Leung, A. Mariotti, K. Mo, S. Schubert, and R. Seager, 2014: Causes and predictability of the 2012 Great Plains drought. *Bull. Amer. Meteor. Soc.*, **95**, 269–282, doi:10.1175/BAMS-D-13-00055.1.

Lee, H. T., 2014: Climate Algorithm Theoretical Basis Document (C-ATBD): Outgoing longwave radiation (OLR)—Daily. NOAA Climate Data Record Program Doc. CDRP-ATBD-0526, 46 pp. [Available online at http://www1.ncdc.noaa.gov/pub/data/sds/cdr/CDRs/Outgoing%20Longwave%20Radiation%20-%20Daily/AlgorithmDescription.pdf.]

Li, S., and A. W. Robertson, 2015: Evaluation of submonthly precipitation forecast skill from global ensemble prediction systems. *Mon. Wea. Rev.*, **143**, 2871–2889, doi:10.1175/MWR-D-14-00277.1.

Lorenz, E. N., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, **20**, 130–141, doi:10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2.

Luo, L., and E. F. Wood, 2007: Monitoring and predicting the 2007 U.S. drought. *Geophys. Res. Lett.*, **34**, L05802, doi:10.1029/2006GL028612.

Mo, K. C., and B. Lyon, 2015: Global meteorological drought prediction using the North American Multi-Model Ensemble. *J. Hydrometeor.*, **16**, 1409–1424, doi:10.1175/JHM-D-14-0192.1.

Quan, X. W., M. P. Hoerling, B. Lyon, A. Kumar, M. A. Bell, M. K. Tippett, and H. Wang, 2012: Prospects for dynamical prediction

of meteorological drought. *J. Appl. Meteor. Climatol.*, **51**, 1238–1252, doi:10.1175/JAMC-D-11-0194.1.

Saha, S., and Coauthors, 2014: The NCEP Climate Forecast System version 2. *J. Climate*, **27**, 2185–2208, doi:10.1175/JCLI-D-12-00823.1.

Scheuerer, M., and T. M. Hamill, 2015: Statistical postprocessing of ensemble precipitation forecasts by fitting censored, shifted gamma distributions. *Mon. Wea. Rev.*, **143**, 4578–4596, doi:10.1175/MWR-D-15-0061.1.

Schiraldi, N. J., and P. E. Roundy, 2017: The evolution of agricultural drought transition periods in the U.S. Corn Belt. *Mon. Wea. Rev.*, **145**, 451–472, doi:10.1175/MWR-D-16-0225.1.

Vitart, F., and Coauthors, 2016: The Subseasonal to Seasonal Prediction (S2S) project database. *Bull. Amer. Meteor. Soc.*, **98**, 163–173, doi:10.1175/BAMS-D-16-0017.1.

Wheeler, M., and G. N. Kiladis, 1999: Convectively coupled equatorial waves: Analysis of clouds and temperature in the wavenumber–frequency domain. *J. Atmos. Sci.*, **56**, 374–399, doi:10.1175/1520-0469(1999)056<0374:CCEWAO>2.0.CO;2.

Wilhite, D. A., 2000: Drought as a natural hazard: Concepts and definitions. *Drought: A Global Assessment*, D. A. Wilhite, Ed., Vol. 1, Routledge, 3–18.

Wu, R., and J. L. Kinter, 2009: Analysis of the relationship of U.S. droughts with SST and soil moisture: Distinguishing the time scale of droughts. *J. Climate*, **22**, 4520–4538, doi:10.1175/2009JCLI2841.1.

Yuan, X., and E. F. Wood, 2013: Multimodel seasonal forecasting of global drought onset. *Geophys. Res. Lett.*, **40**, 4900–4905, doi:10.1002/grl.50949.