African Easterly Wave Forecast Verification and Its Relation to Convective Errors within the ECMWF Ensemble Prediction System

TRAVIS J. ELLESS AND RYAN D. TORN

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

(Manuscript received 30 August 2017, in final form 6 December 2017)

ABSTRACT

African easterly waves (AEWs) are the primary synoptic-scale weather feature found in sub-Saharan Africa during boreal summer, yet there have been few studies documenting the performance of operational ensemble prediction systems (EPSs) for these phenomena. Here, AEW forecasts in the 51-member ECMWF EPS are validated against an average of four operational analyses during two periods of enhanced AEW activity (July–September 2007–09 and 2011–13). During 2007–09, AEW position forecasts were mainly underdispersive and characterized by a slow bias, while intensity forecasts were characterized by an over-intensification bias, yet the ensemble-mean errors generally matched the forecast uncertainty. Although 2011–13 position forecasts were still underdispersive with a slow bias, the ensemble-mean error is smaller than for 2007–09. In addition, the 2011–13 intensity forecasts were overdispersive and had a negligible intensity bias. Forecasts from 2007 to 2009 were characterized by higher precipitation in the AEW trough center and high correlations between divergence errors and intensity errors, suggesting the intensity bias is associated with errors in convection. By contrast, forecasts from 2011 to 2013 have smaller precipitation biases than those from 2007 to 2009 and exhibit a weaker correlation between divergence errors and intensity errors, suggesting a weaker connection between AEW forecast errors and convective errors.

1. Introduction

African easterly waves (AEWs) are the main synoptic-scale weather phenomena occurring over sub-Saharan Africa during boreal summer and provide a significant fraction of the seasonal rainfall in the region (e.g., Diedhiou et al. 1999; Gu et al. 2004). Recently, studies have explored the impact of precipitation on the overall evolution of AEWs. Norquist et al. (1977) were among the first to hypothesize that differences in convection can result in different AEW growth rates. However, because of data limitations, Hsieh and Cook (2007) were the first to try to validate this hypothesis. They concluded that AEWs grow through both barotropic and baroclinic processes, but they also noted that convective activity leads to larger baroclinic energy conversions. Enhanced baroclinic energy conversion is not the only pathway for convective features to impact AEW growth; latent heat release is associated with the redistribution of potential vorticity (PV), which in turn can impact AEW intensity. Berry and Thorncroft (2005) hypothesized that PV generated from convection to the west of the AEW trough center can enhance the westward propagation across the African continent, while diabatically generated PV over the Guinea highlands in West Africa can merge with the PV maximum in the AEW trough, resulting in the creation of a stronger wave. Moreover, Berry and Thorncroft (2012) found diabatic processes lead to the generation of PV within the AEW trough leading to the intensification of the wave. Additionally, they noted that simulations using different convective parameterizations produced variability in the location of the convection and hence diabatically generated PV, which in turn can grow upscale to impact the AEW strength variability. As a consequence, the location of convection relative to the AEW, and hence these positive PV anomalies, can impact the propagation and/or amplification of AEWs.

Variability associated with convection can also affect the predictability of AEWs. Torn (2010) evaluated AEW position and amplitude uncertainty for two initialization times for a strong AEW during September 2006 using ensemble forecasts from the Advanced Research version of the Weather Research and Forecasting...
(WRF) Model. Whereas short-term (12 h) AEW intensity forecasts were most sensitive to the overall wave structure, longer-lead-time forecasts (24–48 h) were more sensitive to downstream thermodynamic variables, suggesting errors in convective processes could play a key role in AEW predictability at longer lead times. This result is consistent with other studies that show convective-scale errors can rapidly grow upscale and begin to influence large-scale features (e.g., Lorenz 1969; Zhang et al. 2002, 2003). However, since Torn (2010) considered only two initialization times, additional research is needed to confirm the hypothesis that convective processes impact the predictability of AEWs.

One way to assess the predictability of AEWs is through ensemble forecasts; however, this type of study requires properly calibrated forecasts of these features. For a well-designed ensemble prediction system (EPS), the ensemble standard deviation should be consistent with the ensemble-mean error (Murphy 1988). While work has been carried out to assess EPS performance for Lagrangian features, including extratropical cyclone tracks (Froude et al. 2007), tropical cyclone (TC) genesis (Halperin et al. 2013; Komaromi and Majumdar 2014), TC tracks (Buckingham et al. 2010; Hamill et al. 2011), and tropical waves, such as the Madden–Julian oscillation (Ling et al. 2014), it appears there are no previous studies that have documented an EPS’s skill for AEWs.

The aim of this study is to verify AEW track and intensity forecasts within the European Centre for Medium-Range Weather Forecasts (ECMWF) EPS and diagnose the impact of convection on AEW forecast errors during two different 3-yr periods. Section 2 of this paper discusses the data and methodology used to calculate the AEW verification. Section 3 presents the verification of AEW position and intensity forecasts. The relationship between forecast errors and convection is discussed in section 4, while a summary of the results is given in section 5.

2. Methodology

a. Tracking AEWs

ECMWF ensemble forecasts of AEWs initialized at 0000 UTC are validated using forecast grids available through The Observing System Research and Predictability Experiment (THORPEX) Interactive Grand Global Ensemble (TIGGE; Bougeault et al. 2010) during July–September (JAS) of 2007–09 and 2011–13. The choice to evaluate ECMWF ensemble forecasts is due to the previous skill shown by the ECMWF EPS in predicting other tropical phenomena, such as tropical cyclone position (e.g., Hamill et al. 2011). This performance is at least partially due to the application of stochastic physics, which yields greater forecast spread in the tropics (e.g., Buizza et al. 1999). On average, the 2007–09 and 2011–13 periods were also associated with an above average value of 2–6-day-filtered 700-hPa eddy kinetic energy\(^1\) (Fig. 1a) and, thus, provide a large number of cases to study. Furthermore, the ECMWF EPS experienced several major upgrades during 2010 (Richardson et al. 2010, 2011), including, but not limited to, increasing the horizontal resolution from T399 (approximately 50-km grid spacing) to T639 (approximately 32-km grid spacing); implementing an ensemble data assimilation technique to generate initial ensemble spread; adding prognostic variables to the microphysics scheme; and modifying the cumulus scheme trigger, to make it harder for convection to initiate over land; therefore, it is important to consider these periods separately. In addition, AEWs are identified and tracked within the control member analysis (i.e., the 0-h unperturbed ensemble member forecast) from the ECMWF, National Centers for Environmental Prediction (NCEP), China Meteorological Administration (CMA), and Met Office (UKMO); these analyses were chosen because they were present in the TIGGE dataset during the entire study period. As will be discussed in section 2b, tracking AEWs in multiple analyses allows this study to assess the uncertainty in AEW position and intensity estimates and does not bias our verification results toward a single analysis value.

The AEW identification and tracking algorithm utilized here is similar to what was used in Brammer and Thorncroft (2015) and is documented below. This algorithm is an extension of the Berry et al. (2007) methods, which employs 700-hPa curvature vorticity to track AEW troughs, because the 700-hPa total vorticity field over Africa is dominated by the shear vorticity associated with the African easterly jet. For simplicity, all data are degraded from their native resolution onto a 1° × 1° grid, which has been shown to be adequate for resolving gross features of AEWs (e.g., Janiga and Thorncroft 2013). Despite the degraded resolution, curvature vorticity on a 1°-resolution grid can be a noisy field with several maxima located within a trough; therefore, an area-averaged curvature vorticity is computed for each grid point by summing the curvature vorticity within 600 km. This area-average curvature vorticity denotes the macroscale rotation associated with the AEW, leading to a smoother vorticity field, which makes it easier to identify and track the AEW center. The first step of the algorithm computes

\(^1\) This metric has been used to quantify the seasonal AEW activity in past studies (e.g., Leroux et al. 2010).
curvature vorticity averaged in 10°-latitude segments between 5° and 20°N at each longitude (Fig. 1b). For each maxima in excess of $0.2 \times 10^{-5} \text{ s}^{-1}$, the feature is considered a continuation of a previously identified AEW if it is within 1000 (500) km of a previously identified analysis (forecast) AEW. A new AEW is declared if a maximum has area-averaged curvature vorticity in excess of $0.4 \times 10^{-5} \text{ s}^{-1}$, is over the African continent (17°W–40°E; Fig. 1b) and is not associated with a previously identified system. These criteria were determined through extensive sensitivity testing; more stringent initial AEW thresholds often miss the early part of many AEW life cycles, while lower thresholds identified numerous spurious features that are not associated with AEWs. The combination of the lower intensity and distance criteria allows for the continuing identification for weakening AEWs while suppressing the identification of non-AEW features. For all features passing the threshold test, the longitude of the vorticity maximum is used as a first-guess location to define the AEW position. The final position is determined via a mass-centering technique applied to the area-averaged curvature vorticity, similar to what has been used for TCs (e.g., Nguyen et al. 2014). For the remainder of this paper, AEW intensity refers to the value of area-averaged curvature vorticity at the final AEW location.

### b. Error calculations

AEW forecasts are validated using methods similar to other feature-following verification studies (e.g., Elsberry and Carr 2000; Buckingham et al. 2010; Hamill et al. 2011). Many of these studies were able to make use
of an independent best-track record of a system’s position and intensity; however, such a database does not exist for AEWs. As a consequence, forecasts are validated against an ensemble of deterministic model analysis values. While it is common to verify an ensemble against its own analysis value, this choice often leads to lower forecast errors compared to validating against other analyses (e.g., Froude et al. 2007). In situ observation coverage in sub-Saharan Africa is also relatively sparse relative to other continental areas; therefore, the analysis can be weighted heavily toward the model’s first-guess field, which could mask biases when validating against the same model’s analysis. Given this, the verification dataset consists of AEW position and intensity values obtained by averaging over the ECMWF, NCEP, CMA, and UKMO analyses. Although it is likely that these forecast centers are characterized by different levels of analysis quality, it is impossible to quantify whether one center’s analysis provides a more accurate estimate of the atmosphere over Africa; therefore, the validation dataset treats all analysis estimates equally. While the multicenter analysis approach allows for appropriate error calculation, it also provides uncertainty in the validation value, allowing for a more appropriate measure of forecast uncertainty as well. Finally, for a unique AEW position to enter this best-track dataset, the AEW must be tracked over Africa; therefore, the validation dataset treats all cases where an AEW is identified after 0 h). These calculations are only computed for AEWs identified at 0 h in the forecast (i.e., verification does not include cases where an AEW is identified after 0 h). Moreover, forecast errors are computed only for times when the AEW is tracked in at least 26 ensemble members (>50% of the ensemble); this choice produces a comparable sample size at each lead time between the 2007–09 and 2011–13 forecast periods (Table 1) and ensures that the statistics focus on robust AEWs. While this choice is a more stringent threshold than may be found in other feature-following studies that use 10%–38% of the ensemble members (e.g., Froude et al. 2007; Buckingham et al. 2010), the case counts and statistics calculated show little sensitivity when a less stringent threshold, such as 13 members (>25% of the ensemble; not shown), is used.

### TABLE 1. Number of AEW forecasts with valid analysis times for each lead time.

<table>
<thead>
<tr>
<th>Lead time (h)</th>
<th>0</th>
<th>12</th>
<th>24</th>
<th>36</th>
<th>48</th>
<th>60</th>
<th>72</th>
<th>84</th>
<th>96</th>
<th>108</th>
<th>120</th>
</tr>
</thead>
<tbody>
<tr>
<td>2007–09 case count</td>
<td>258</td>
<td>259</td>
<td>221</td>
<td>217</td>
<td>191</td>
<td>186</td>
<td>169</td>
<td>161</td>
<td>145</td>
<td>138</td>
<td>123</td>
</tr>
<tr>
<td>2011–13 case count</td>
<td>257</td>
<td>247</td>
<td>217</td>
<td>203</td>
<td>190</td>
<td>178</td>
<td>162</td>
<td>152</td>
<td>137</td>
<td>128</td>
<td>120</td>
</tr>
</tbody>
</table>

The forecast bias-corrected root-mean-square error (RMSE) and bias for a given lead time $t$ are computed via

$$
\text{RMSE}(t) = \sqrt{\frac{1}{M(t)} \sum_{m=1}^{M(t)} [\bar{x}_m(t) - b(t) - \bar{\sigma}_m(t)]^2} \quad \text{and (1)}
$$

$$
b(t) = \frac{1}{M(t)} \sum_{m=1}^{M(t)} \bar{x}_m(t) - \bar{\sigma}_m(t), \quad \text{and (2)}
$$

where $\bar{x}_m(t)$ is the ensemble-mean forecast value for the $m$th AEW forecast at lead time $t$, $\bar{\sigma}_m(t)$ is the corresponding mean analysis value, $b(t)$ is the bias at lead time $t$ [Eq. (2)], and $M(t)$ is the total number of verifying forecasts at lead time $t$. For a well-calibrated EPS, the expected value of the deviations from the ensemble mean (i.e., ensemble variance) should equal the mean-square error in the ensemble mean (Murphy 1988). When computing the ensemble variance, one also needs to account for uncertainty in the verification quantity itself, which here is the analysis uncertainty. As a consequence, the forecast uncertainty (SD) for a given lead time $t$ is computed via

$$
\text{SD}(t) = \sqrt{\frac{1}{M(t)} \sum_{m=1}^{M(t)} [\sigma_{om}^2(t) + \sigma_{en}^2(t)]} \quad \text{and (3)}
$$

$$
\sigma^2 = \frac{1}{N-1} \sum_{n=1}^{N} (y_n - \bar{y})^2, 
$$

where $\sigma_{om}^2(t)$ is the forecast (analysis) variance [Eq. (4)] for the $m$th AEW forecast for lead time $t$, $N$ is the number of forecast (analysis) ensemble members, $y_n$ is an individual forecast (analysis) ensemble member value, and $\bar{y}$ is the ensemble mean value. For simplicity, these calculations are only computed for AEWs identified at 0 h in the forecast (i.e., verification does not include cases where an AEW is identified after 0 h). Moreover, forecast errors are computed only for times when the AEW is tracked in at least 26 ensemble members (>50% of the ensemble); this choice produces a comparable sample size at each lead time between the 2007–09 and 2011–13 forecast periods (Table 1) and ensures that the statistics focus on robust AEWs. While this choice is a more stringent threshold than may be found in other feature-following studies that use 10%–38% of the ensemble members (e.g., Froude et al. 2007; Buckingham et al. 2010), the case counts and statistics calculated show little sensitivity when a less stringent threshold, such as 13 members (>25% of the ensemble; not shown), is used.

### 3. AEW verification

#### a. AEW position errors

AEW position errors exhibit consistent rate of error growth for these two periods (Fig. 2). The 2007–09 forecast period exhibits a fairly constant growth rate [approximately 70 km (24 h)$^{-1}$ during the first 48 h, with a smaller rate [40 km (24 h)$^{-1}$] beyond 72 h. By contrast, the 2011–13 period has a large error growth rate [200 km (24 h)$^{-1}$] for the first 12 h, which decreases considerably to approximately 50 km (24 h)$^{-1}$ beyond 12 h, yet is consistently smaller than that for 2007–09. For both periods, the SD consistently increases at a rate of 20 km (24 h)$^{-1}$, with the 2011–13 period showing a
slightly larger SD value at all lead times. The statistically
significant differences between the RMSE and SD at all
lead times for 2007–09 indicates that the ECMWF EPS
is underdispersive for this metric. By contrast, the dif-
fferences between RMSE and SD during the 2011–13
period are significant after 60 h, suggesting that short-
term position forecasts are better calibrated in the more
recent model version.

Separating the position errors into their zonal and
meridional components provides further insight into the
nature of the errors. The RMSE and SD for the merid-
ional direction exhibit different characteristics than the
zonal direction (Fig. 3a). From 0 to 60 h, the meridional
position forecast SD is larger than the RMSE, which is
opposite of what is seen for the total position forecasts.
The meridional position bias for both periods remains
relatively small and is rarely statistically different from
zero (Fig. 3b). Furthermore, the similarity between the
2007–09 and 2011–13 periods suggests the 2010 model
changes do not have much impact on meridional
position errors.

For the zonal direction, the 2011–13 RMSE is slightly
smaller and the SD is slightly larger than the 2007–09
period (Fig. 3c). The relationship between the RMSE
and SD shows that the 2011–13 forecasts are fairly well
calibrated in zonal position forecasts from 12 to 72 h,
while the 2007–09 period is underdispersive beyond 24 h.
Moreover, both periods are characterized by a positive
(eastward) zonal position bias at all lead times (Fig. 3d),
which increases as a function of lead time, meaning that
AEWs move too slowly within the model. Whereas the
bias in 2007–09 increases at a rate of roughly 80–100 km day$^{-1}$,
the bias in the 2011–13 forecasts remains relatively constant after 72 h (between 140 and 200 km).

Section 4 explores the relationship between biases in
AEW longitude and convection to diagnose if model changes during 2010 could be responsible for the
observed AEW longitude bias reduction.

b. AEW intensity errors

Similar to AEW position, AEW intensity forecasts
have distinct errors and biases for each period (Fig. 4a).
The 2007–09 period is characterized by the RMSE and
SD increasing at nearly the same rate at all lead times,
suggesting the EPS had well-calibrated AEW intensity
forecasts. By contrast, the 2011–13 period has a 40% reduction in RMSE relative to the 2007–09 period at all
lead times, while the implementation of the ensemble
data assimilation during the 2010 upgrades (Buizza et al.
2008; Richardson et al. 2010) likely aids in generating a
comparable SD between the two periods. The re-
duction in error coupled with comparable spread
yields an overdispersive ensemble for AEW intensity
for this period.

Forecast biases for these two periods highlight dis-
tinctive differences. During the first 24 h, intensity biases
are similar between both periods; however, differences
between the two periods become more apparent after 24 h (Fig. 4b). The 2007–09 period is characterized by an
increasing positive bias (i.e., overintensification), while
the 2011–13 intensity bias decreases to near zero and is
statistically insignificant.

While these periods are associated with different bias
trends, a consistent diurnal signal, which is not caused by
verifying against the multicenter analysis (not shown), is
present in the forecast bias for both periods. Given this
study validates forecasts initialized at 0000 UTC, the
negative intensity bias relative to the forecast trend oc-
curs for lead times that verify at 1200 UTC (i.e., 12, 36,
and 60 h), while lead times verifying at 0000 UTC (i.e.,
24, 48, and 72 h) have a positive intensity bias relative to
the trend. This observed diurnal cycle follows the di-
urnal cycle of mesoscale convective systems (MCSs)
over West Africa, where a minimum (maximum) in
AEW-related MCS activity is typically found during the
early morning (late afternoon/evening) hours (e.g.,
Payne and McGarry 1977; Reed et al. 1977; Jackson
et al. 2009). As a consequence, it is possible that AEW
convection influences AEW intensity biases; therefore,
further investigation is warranted to assess whether the
reduction in intensity bias from 2007–09 to 2011–13 is
related to differences in the model’s representation of convection.

4. Relationship with convection

The previous section highlighted distinct differences in AEW forecast intensity and zonal position errors occurring between forecasts from the 2007–09 and 2011–13 periods. A model upgrade in-between the two periods changed the cumulus scheme trigger, making it harder for convection to initiate over the African continent. Furthermore, the upgrade also implemented a finer resolution and a new microphysics scheme, which could also impact the convective evolution within the forecast. Therefore, it is prudent to assess the modeled AEW convection and its relationship to AEW position and intensity errors to ascertain if these model changes are the likely source of AEW forecast error improvement. Unfortunately, the TIGGE dataset contains limited fields and vertical resolution; therefore, a number of proxy metrics related to convection are employed. Moreover, it is not possible to run the two model configurations on the same cases; therefore, the following section provides circumstantial evidence for the connection between AEW forecast errors and convection.

a. Convective environment

The relationship between convection and AEW forecast biases is first diagnosed by verifying large-scale fields related to convection against analysis fields.
Previous TC genesis studies have measured the large-scale tropical convective environment through moisture availability, primarily through relative humidity (RH), and upper-level divergence (e.g., Komaromi and Majumdar 2014, 2015). While AEWs are not TCs, they are influenced by convection; therefore, these metrics could translate and are used here. Sufficient moisture is required to sustain long-lived convection and the upper-level divergence is a measurement of deep convection outflow regardless of location, including for AEWs (e.g., Reed et al. 1977). Following Komaromi and Majumdar (2014, 2015), the near-wave convective environment is assessed by radially averaging RH and upper-level divergence fields within 500 km from the wave center. This radius highlights the near-wave convective environment and was used by Brammer and Thorncroft (2015) to study AEW synoptic-scale patterns. Furthermore, Berry and Thorncroft (2012) noted convection associated with the AEWs occurs within a similar distance. It is worth noting that the results show little sensitivity to small changes in the averaging radius (400–600 km; not shown).

Previous studies of AEW and TC genesis have quantified upper-level divergence at 200 hPa (e.g., Reed et al. 1977; Komaromi and Majumdar 2014, 2015), while other studies investigated divergence between 300 and 200 hPa (e.g., Brammer and Thorncroft 2015). Given this ambiguity, divergence errors are calculated at 200 and 300 hPa by comparing ensemble-mean divergence against the multicenter analysis mean value. Both periods are characterized by positive divergence biases within this layer (Figs. 5a,b), meaning the forecast has too much divergence. For the 2007–09 period, the 200-hPa divergence bias has a noticeable diurnal cycle with the maximum (minimum) in bias corresponding to the same times as one might expect to have a maximum (minimum) in MCS activity while the bias is smaller at 300 hPa. This result suggests that 200-hPa divergence may be more representative of errors in convection than 300-hPa divergence. By contrast, a diurnal signal is not observed in the 2011–13 divergence bias at any level, and the magnitude of the bias for 2011–13 is smaller than the 2007–09 period at all lead times. The missing diurnal cycle and reduction in biases suggest convective activity is reduced for the 2011–13 forecasts relative to the 2007–09 forecasts.

The availability of moisture can impact convection; therefore, the forecasts of this field are also verified at multiple levels. Both 925- and 850-hPa RH exhibit a positive bias at 0 h that remains through the majority of the forecast lead times for both periods (Figs. 5c,d). While a positive low-level bias might be expected to make it easier for convection to initiate in the model, the midlevel RH plays a larger role in determining the depth and maintenance of convection. For the 2007–09 period, levels above 700 hPa have a dry bias that transitions to a moist bias after the first diurnal cycle and subsequently becomes more positive over time (Fig. 5c). This increasing bias could be related to excessive convection moistening the column as maximum (minimum) in biases follow the same diurnal signal observed in MCS activity (e.g., Jackson et al.
2009); however, this is speculative. Whereas the 0-h 700–500-hPa RH bias for 2011–13 is also negative, the RH bias remains negative for all lead times (Fig. 5d). Although there is a diurnal signal within the RH bias, there is not an upward trend with lead time, suggesting that the model does not become systematically more moist with time as during the 2007–09 period. The negative RH bias also suggests reduced convection during the 2011–13 period, especially around 1200 UTC verifying times.

One way to try to relate AEW biases and convective environment errors is by looking at the joint distribution of these quantities over all forecasts. The joint distribution matches AEW forecast errors with corresponding environmental errors on a case-by-case basis. Komaromi and Majumdar (2015) used this technique to highlight a relationship between circulation errors and relative humidity errors in TC genesis cases, showing a weak circulation bias in the forecast was associated with a 700-hPa dry bias. Whereas Komaromi and
Majumdar (2015) compared the instantaneous error in convective environment fields at individual forecast hours, this might not be the most appropriate choice. Errors in the convective scale can grow upscale over time (e.g., Lorenz 1969; Zhang et al. 2003); therefore, 48-h AEW position and intensity errors are likely influenced by environmental errors at earlier lead times (i.e., 24–48 h). The 24–48-h convective environment period occurs within the second diurnal cycle, which could have greater impact on convective errors at this time compared to initialization errors (Fig. 5). The 48-h AEW position and intensity biases reflect the forecast errors at the end of the previous diurnal cycle; the results discussed below remain similar if 72-h position and intensity biases were used instead of the 48-h biases (not shown). If there is no relationship between AEW metrics and the convective proxies, one would expect a two-dimensional Gaussian distribution with roughly equal numbers of cases in each quadrant. A positive relationship between the AEW and convective environment biases would show in the joint distributions by having the majority of cases located in the top-right and bottom-left corners of the distribution. Cases falling in these quadrants would have higher (lower) position or intensity biases occurring with higher (lower) environmental metric biases, suggesting more (less) convection would be favored.

The 2007–09 joint distributions for divergence and AEW zonal position bias (Fig. 6a) and AEW intensity bias (Fig. 6c) indicate a positive relationship as 66.49% and 63.35% of cases, respectively, and are located within the top-right quadrant. Both distributions also feature a fairly large correlation between position errors and divergence errors (0.233) and intensity errors and divergence errors (0.531; Table 2). This strong correlation is seen in the positive tilt of the distribution, as the largest positive position and intensity biases are associated with the largest divergence bias, suggesting that large AEW position and intensity biases are usually
TABLE 2. Correlation coefficients between the variables utilized in the joint distributions (Figs. 6 and 7).

<table>
<thead>
<tr>
<th>Distribution</th>
<th>2007–09</th>
<th>2011–13</th>
</tr>
</thead>
<tbody>
<tr>
<td>Longitude–divergence</td>
<td>0.233</td>
<td>0.002</td>
</tr>
<tr>
<td>Intensity–divergence</td>
<td>0.531</td>
<td>0.117</td>
</tr>
<tr>
<td>Longitude–RH</td>
<td>0.338</td>
<td>0.116</td>
</tr>
<tr>
<td>Intensity–RH</td>
<td>0.125</td>
<td>0.120</td>
</tr>
</tbody>
</table>

Table 2. Correlation coefficients between the variables utilized in the joint distributions (Figs. 6 and 7). Similarly, the 2011–13 AEW intensity error joint distribution (Fig. 6d) has a much weaker correlation than in 2007–09 (0.170, Table 2), where 35.26% of cases are in the top-right quadrant and 35.79% are in the bottom-right quadrant, suggesting that positive divergence bias does not necessarily result in a positive intensity bias. Given that upper-level divergence in part quantifies the overall strength of the convection, it appears that convective-related biases may have a larger impact on AEW position and intensity errors for the 2007–09 forecasts compared to the 2011–13 forecasts.

Similar to the AEW divergence error joint distributions, the AEW 700-hPa RH error joint distributions for 2007–09 have the largest numbers of cases in the top-right quadrant for AEW zonal position error (Fig. 7a) and intensity error (Fig. 7c) (53.76% and 46.77%, respectively), with a stronger correlation found with position errors (0.338) compared to intensity errors (0.125; Table 2). In particular, positive position and intensity biases tend to occur when the atmosphere is more moist than the corresponding analysis, which could be due to excessive convection occurring in the AEW trough center. Prior studies suggest that the diabatic generation of PV downstream of the AEW helps propagate the AEW westward (e.g., Berry and Thorncroft 2005; Tomassini et al. 2017). By contrast, anomalous convection in the trough center generates PV within the trough center, which slows the westward movement of the AEW (positive longitude bias) and intensifies the already existing AEW trough (positive intensity bias). As a consequence, the dry bias observed in 2011–13 should reduce the amount of anomalous convection occurring within the AEW trough center and theoretically result in faster and weaker AEWs than the 2007–09 period as the diabatic PV may form farther downstream (closer to the observations), allowing for a more accurate representation of AEW propagation by downstream relocation. However, this dry bias does not seem to fully alleviate the slow bias, as demonstrated by the joint 2011–13 zonal position bias–700-hPa RH bias distribution (Fig. 7b). The distribution prominently shows the dry 700-hPa RH bias (66.84% of cases are on the distribution’s left side), while forecasts still primarily have slower moving AEWs (48.42% of cases occurring in the top-left quadrant). This observed difference in the joint error distribution also highlights a weaker relationship between position and RH errors than during 2007–09 (a correlation of 0.1157; Table 2). The 2011–13 dry bias also does not present a distinct weakening in AEW intensity forecasts; forecasts have almost an equal chance of having stronger (31.58% of cases) or weaker (35.58% of cases) AEWs when a dry bias is present (Fig. 7d). The relationship between intensity and RH errors, however, remains relatively consistent with 2007–09 (Table 2). Consequently, moisture errors, and presumably convective errors, appear to have a much larger influence on forecast errors during 2007–09.

b. Composite precipitation differences

AEW forecasts are likely sensitive to the amount of convection and its location relative to the AEW center. If the convection is located to the west of the trough center, then the PV generated will aid in AEW propagation rather than the intensity (e.g., Berry and Thorncroft 2005). By contrast if the convection is close to the center, it can amplify the wave through vortex stretching. As a consequence, it is important to validate precipitation forecasts with respect to AEWs.

Forecast precipitation biases are computed by verifying the model precipitation rate against the Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis (TMPA; TRMM product 3B42; Huffman et al. 2007) values within an AEW-centric framework. To provide a fair comparison between the model and observations, the TRMM precipitation estimate, which has been used by previous studies to assess convective activity over Africa (e.g., Janiga and Thorncroft 2013), is degraded from its 0.25° × 0.25° resolution to a 1° × 1° grid. For each forecast lead time, all ensemble members are shifted to an AEW-centric reference frame (located at 9°N and 0° in these figures), and the ensemble-mean streamfunction anomaly and forecast precipitation rate are calculated at each grid point. The resulting ensemble-mean precipitation forecasts are then averaged over all initialization times for a given lead time. The TRMM values are composited using the same methodology based on the analysis AEW position; this choice removes any position error that may exist between the forecast and analysis positions, which in turn can result in precipitation errors. To relate precipitation errors with AEW forecast errors, this
methodology focuses on understanding how the precipitation rate and coverage in the model compares to the observations over many cases.

Forecasts from 2007–09 exhibit significant biases in precipitation location and intensity. The 12-h forecast composite shows precipitation is centered on the AEW trough center, with a maximum rate value of 6.30 mm (6 h)$^{-1}$ located 1° east of the trough center (Fig. 8a). By contrast, the verifying TRMM precipitation rates are maximized 4° to the west of the AEW center (Fig. 8b), similar to observational studies of AEW convection (e.g., Reed et al. 1977; Berry and Thorncroft 2005), with a maximum precipitation rate of 4.48 mm (6 h)$^{-1}$.

Taking the difference between the model forecast and TRMM analysis precipitation rates indicates that the model produces larger amounts of precipitation within and upstream of the AEW trough, with a maximum difference of 4.13 mm (6 h)$^{-1}$ found in the trough center (Fig. 8c). By contrast, the downstream precipitation differences are fairly small 3° to the west of the AEW. A slightly different pattern is observed at 24 h (Fig. 8d) with negative precipitation differences (i.e., more precipitation occurring in the observations than in the forecast) observed downstream and to the north of the trough center while similar positive precipitation differences occur within the AEW trough. By 36 h (Fig. 8e), the precipitation distribution resembles the 12-h pattern, but with a much larger magnitude; a maximum difference of 6.94 mm (6 h)$^{-1}$ is found 1° east of the center. By 48 h (Fig. 8f), the precipitation differences again resemble the 24-h forecast. This alternating pattern of differences with time continues throughout the remainder of the forecast as lead times verifying at 1200 UTC have a pattern similar to Fig. 8e and lead times verifying at 0000 UTC resemble Fig. 8f (not shown).

Forecast biases from 2011–13 cases are generally smaller relative to 2007–09. The 12-h forecast precipitation (Fig. 9a) has a maximum precipitation rate of 6.28 mm (6 h)$^{-1}$ located at the AEW center while the verifying analysis (Fig. 9b) has a precipitation maximum of 3.85 mm (6 h)$^{-1}$ focused 2° downstream of the trough center. It is worth pointing out that subtle differences...
exist in the observed precipitation distribution for the 2007–09 and 2011–13 periods. Compared to the 2007–09 cases, the spatial extent of heavy precipitation is smaller for 2011–13, especially on the eastern side of the wave, where precipitation rates are 0.5 mm (6 h)\(^{-1}\) lower. This slight shift westward and decrease in forecast

**FIG. 8.** Wave-centered composite of the 700-hPa streamline anomaly (contours every \(2 \times 10^6\) s\(^{-1}\); dashed contours represent negative values) and precipitation rate [mm (6 h)\(^{-1}\); color fill] during 2007–09 for (a) 12-h forecasts and (b) verifying TRMM analyses. Crosses denote the center of the AEW trough. Shown are the forecast streamline anomaly (contour) and precipitation differences between composite forecast and TRMM analysis [mm (6 h)\(^{-1}\); color fill] for (c) 12-, (d) 24-, (e) 36-, and (f) 48-h forecasts. The composite center is at 9°N and 0°. Geographic background is for illustrative purposes only.
precipitation magnitude yields a lower magnitude difference between the forecast and observations (Fig. 9c); the maximum difference is only 3.06 mm (6 h)$^{-1}$ at the AEW trough center for 2011–13. The precipitation differences for the remainder of the forecast (Figs. 9d–f) follow a pattern similar to that of the 2007–09 period, with positive differences found within the AEW trough at all lead times with negative differences found both down- and upstream of the AEW. Unlike the 2007–09 period, the magnitudes of the positive precipitation differences are drastically smaller (maximum differences are reduced >30% for both lead times), which might be expected to result in reduced latent heat release.

**FIG. 9.** As in Fig. 8, but for the 2011–13 period.
Given the positive precipitation errors that occurred within the center of the trough at all lead times, there is likely a direct feedback between precipitation biases and the AEW intensity and position biases. Excessive convection within the trough would be expected to lead to positive PV anomalies in that region, resulting in an amplification of the AEW. Since the 2007–09 period saw substantial positive precipitation biases within the wave at all lead times, the anomalous PV that is generated likely leads to an overintensification of the AEW, which is evident in the constant positive intensity bias for all lead times (cf. Fig. 4b), while lower-amplitude precipitation biases in the 2011–13 period likely lead to less anomalous PV and a reduced intensity bias. Furthermore, given the propagation of the AEW is in part driven by the reorganization of the AEW center toward the downstream PV maxima (Berry and Thorncroft 2005), the reduction of precipitation (and subsequent PV generation) downstream of the AEW can result in slower AEW motion, which may at least partially account for the observed slow bias during both periods (cf. Fig. 3b).

\[ c. \text{Drivers of precipitation changes} \]

Given that forecasts of the same cases with the two versions of the ECMWF EPS are not available, it is not possible to conclusively say why these two forecast sets were characterized by different forecast errors. The difference in performance could be explained by two factors: the cases from 2007 to 2009 were characterized by distinctive biases due to the nature of the AEWs during this period or changes to the forecast model itself led to the improved performance.

One way to determine the role of the periods themselves is to compare the mean characteristics of AEWs from the two periods. This method is meant to assess whether differences in the forecast AEW structure (position and intensity) or the near-wave convective environment (500-km area-averaged 700-hPa relative humidity and 200-hPa divergence) could be responsible for the observed differences in forecast performance. Previous work has shown that the location of the wave can result in precipitation differences since precipitation types can vary across Africa (e.g., Janiga and Thorncroft 2013). For all lead times, the mean AEW longitude is similar for the two periods (Fig. 10a). In addition, the initial AEW intensity is also similar between the two periods (Fig. 10b). Although the 2007–09 period has AEWs that are statistically stronger beyond 48 h, it is unlikely that this is the source of the differences in bias since the 2007–09 and 2011–13 intensity biases have a different character prior to 48 h. Minimal differences between both periods also exist in the convective environment. The midlevel RH shows small, insignificant differences between the two periods through the forecast (Fig. 10c), while the initial upper-level divergence is also similar (Fig. 10d). The 2007–09 period does not have a significantly larger upper-level divergence until after 24 h (after the precipitation differences are apparent), suggesting the differences in the convective environment are also not the source of the convective differences.

Given that the wave structure does not show a coherent explanation for the differences in forecast precipitation, it seems more likely the upgrade to ECMWF version Cy36r4 in 2010 is the source of the difference. Increased horizontal resolution, adding prognostic variables to the microphysical scheme, and modifying the cumulus trigger could all have contributed to the decrease in forecast errors. Increasing the horizontal resolution allows better representation of the AEW structure in the forecast while also decreasing the spatial extent of cumulus convection at an individual grid point. Furthermore, the 2010 upgrade changed the original two-variable prognostic microphysics scheme to a five-variable scheme, which allowed rainwater to become a prognostic variable (Forbes et al. 2011). Forbes et al. (2011) indicated that since rain is now a prognostic variable, falling rainwater can be advected by the horizontal wind, which led to improvements in overall precipitation skill. In the cumulus scheme, the convective trigger was made a function of being over land or ocean. Before the 2010 upgrades, convection was triggered if the cloud water content exceeded a threshold of 0.3 g kg\(^{-1}\) (ECMWF 2010), while the 2010 upgrades raised this threshold to 0.5 g kg\(^{-1}\) over land, while keeping the 0.3 g kg\(^{-1}\) threshold over water (ECMWF 2011). The stricter convective trigger over land would make it more difficult to trigger convection, hence decreasing the convective coverage for the 2011–13 period.

5. **Summary and conclusions**

The goal of this study is to assess ECMWF EPS forecasts of AEW position and intensity forecast errors during two distinct periods of enhanced AEW activity: 2007–09 and 2011–13. Verifying forecasts against a multicenter analysis indicate that the ECMWF EPS position forecasts are underdispersive with an overall slow bias observed for both periods. Ensemble AEW intensity forecasts are calibrated during 2007–09, while a reduction in forecast errors led to the forecasts being overdispersive for 2011–13. Furthermore, AEW intensity forecasts for
2007–09 have an overintensification bias while the 2011–13 period has a near-zero intensity bias. The diurnal signal observed in the intensity bias, combined with the fact the ECMWF EPS changes altered key components that relate to convection in 2010, suggests convective activity could be a contributing factor in the position and intensity bias differences between the two periods.

A comparison of forecast biases against convective environment errors suggests that convective-related biases could be responsible for the AEW track and intensity bias during 2007–09. Forecasts from 2007 to 2009 consistently had higher precipitation rates within the AEW trough at all lead times. Generation of midlevel PV from this anomalous convection can strengthen the AEW through a positive feedback process, which leads to a growing intensity bias with time. The reduction of the in-trough precipitation bias during 2011–13 likely reduces erroneous PV generation and, hence, reduces the rate of AEW growth within the forecast period. Given that both periods presented similar intensity and zonal position biases through the first 12 h, it is likely that differences in convective activity starting in the 12–24-h time frame could be responsible for generating the observed bias difference at later lead times.

Many factors other than the ECWMF EPS changes may also contribute in determining the amount of convective activity over the African continent. The influence of large-scale patterns, most likely equatorial waves, on 25–60- and 10–25-day time scales have been shown to affect the variability of convective activity.

**Fig. 10.** Mean forecast (a) longitude, (b) intensity, (c) 500-km area-averaged 700-hPa RH, and (d) 500-km area-averaged 200-hPa divergence as a function of lead time for 2007–09 (blue lines) and 2011–13 (red lines). Dots indicate times when the two periods are statistically different at the 95% confidence interval using bootstrap resampling.
(Janicot and Sultan 2001; Sultan et al. 2003), while the background convective states across different regions of Africa favor different modes and strengths of convective activity (Janiga and Thorncroft 2013). As a consequence, it is possible that forecasts occurring within these varying conditions could produce significantly different results. Since this present study focused on assessing domain-wide errors, further work is needed to focus on individual cases to determine the role these factors may play in producing large errors. Finally, while this study suggests a broad relationship between AEW forecast errors and convective activity, PV generation from convection may not be the only mechanism that can lead to AEW intensification. Energy conversions from barotropic and baroclinic processes can act as a source of energy for AEW growth; therefore, further work is also needed to assess how uncertainty in these processes can impact AEW forecast errors.

Acknowledgments. We thank our colleagues at the University at Albany, Yuan-Ming Chen and Chris Thorncroft, for providing valuable feedback throughout the completion of this work. We would also like to thank the three anonymous reviewers for their constructive feedback that helped to improve this manuscript. This work was supported through the National Science Foundation (NSF Award 1321568).

REFERENCES


