Comparative Assessment of Two Objective Forecast Models for Cases of Persistent Extreme Precipitation Events in the Yangtze–Huai River Valley in Summer 2016

BAIQUAN ZHOU
State Key Laboratory of Severe Weather, Chinese Academy of Meteorological Sciences, China Meteorological Administration, and University of Chinese Academy of Sciences, Beijing, China

PANMAO ZHAI
State Key Laboratory of Severe Weather, Chinese Academy of Meteorological Sciences, China Meteorological Administration, Beijing, China

RUOYUN NIU
National Meteorological Center, China Meteorological Administration, Beijing, China

(Manuscript received 21 March 2017, in final form 16 December 2017)

ABSTRACT

Two persistent extreme precipitation events (PEPEs) that caused severe flooding in the Yangtze–Huai River valley in summer 2016 presented a significant challenge to operational forecasters. To provide forecasters with useful references, the capacity of two objective forecast models in predicting these two PEPEs is investigated. The objective models include a numerical weather prediction (NWP) model from the European Centre for Medium-Range Weather Forecasts (ECMWF), and a statistical downscaling model, the Key Influential Systems Based Analog Model (KISAM). Results show that the ECMWF ensemble provides a skillful spectrum of solutions for determining the location of the daily heavy precipitation (>25 mm day$^{-1}$) during the PEPEs, despite its general underestimation of heavy precipitation. For lead times longer than 3 days, KISAM outperforms the ensemble mean and nearly one-half or more of all the ensemble members of ECMWF. Moreover, at longer lead times, KISAM generally performs better in reproducing the meridional location of accumulated rainfall over the two PEPEs compared to the ECMWF ensemble mean and the control run. Further verification of the vertical velocity that affects the production of heavy rainfall in ECMWF and KISAM implies the quality of the depiction of ascending motion during the PEPEs has a dominating influence on the models’ performance in predicting the meridional location of the PEPEs at all lead times. The superiority of KISAM indicates that statistical downscaling techniques are effective in alleviating the deficiency of global NWP models for PEPE forecasts in the medium range of 4–10 days.

1. Introduction

Persistent extreme precipitation (PEP) has attracted a great deal of attention because of its long duration and association with natural disasters. A regional PEP event (PEPE) is capable of inducing severe flooding of rivers and streams, which may threaten people’s lives and property and cause significant disruption to regional economies (Galarneau et al. 2012). The two historical flood events that struck central eastern China in 1998 and Pakistan in 2010 are typical examples that were triggered by clusters of PEPEs (Galarneau et al. 2012; Lau and Kim 2012; Lu 2000). However, accurately predicting PEPEs remains a challenge for forecasters in operational meteorology (Ebert et al. 2003). The problem they face is primarily the insufficient skill of operational global numerical weather prediction (NWP) models to correctly predict the location, intensity, and duration of PEPEs (Ralph et al. 2010; Sukovich et al. 2014). Numerous evaluation studies have been carried out on forecasts of precipitation and the related large-scale atmospheric circulation produced by the NWP models from The Observing System Research and Predictability Experiment Interactive Grand Global Ensemble (TIGGE). With regard to heavy precipitation, poor performance of low-resolution global NWP models with general underestimation has been reported in
many studies (Krishnamurti et al. 2009; Ralph et al. 2010; Willems et al. 2012; Zhi et al. 2013; Su et al. 2014; Sharma et al. 2017). For an individual extreme precipitation event, Kobold and Suselj (2005) found that the European Centre for Medium-Range Weather Forecasts (ECMWF) underestimated the 27–28 June 1997 precipitation in a Slovenian catchment by 60%. Research conducted by Lavers and Villarini (2013) indicated that ensemble mean rainfall patterns from NWP models in TIGGE initialized at a roughly 4-day lead time failed to forecast the persistent nature of a PEPE. Characterized by an extreme and persistent nature, PEPEs pose enormous difficulties to operational forecasters and NWP models. There is thus an urgent need for effective forecast methods for PEPEs, for disaster prevention and mitigation, particularly across regions like the Yangtze–Huai River valley (YHRV), which is one of the most important agricultural zones and densely populated regions in China (Ralph et al. 2005; Lu 2000). Within the context of global warming, with more intense and frequent extreme precipitation events projected (Hartmann et al. 2013; Hirsch and Archfield 2015), the YHRV will be more likely to suffer severe socioeconomic damage due to the flooding and landslides caused by PEPEs.

Helpfully, studies on PEPEs in China are now emerging. Instead of investigating extreme precipitation primarily on a daily basis (e.g., Karl and Knight 1998; Zhai et al. 2005), Chen and Zhai (2013) introduced a new definition for regional PEPEs based on the extremity, temporal persistence, and spatial contiguity of the precipitation. The criteria for this PEPE definition include 1) that daily precipitation amounts of not less than 50 mm should be observed by at least three neighboring stations and 2) that the condition in the first criterion should persist for at least 3 days at every station. The PEPE comes to an end when the first criterion is not satisfied for the following two consecutive days. In their subsequent research, they reported that PEPEs typically result from concurrent combinations of persistent anomalies from the displacement of the South Asian high and jets, the development of blocking patterns, and a westward shift of the western Pacific subtropical high (WPSH) (Chen and Zhai 2014). Therefore, utilizing the revealed circulation patterns to downscale PEPEs is a sensible strategy to alleviate the deficiency of global NWP models for extreme precipitation forecasts, since they exhibit better skill in forecasting large-scale circulation variables and key synoptic systems than extreme precipitation (Park et al. 2008; Pelly and Hoskins 2003; Niu and Zhai 2013; Niu et al. 2015; Zhou et al. 2015). Statistical downscaling models are usually developed and employed for weather and climate prediction (Wilby and Wigley 1997; Zorita and Von Storch 1999; Fan et al. 2008; Wang and Fan 2009; Caillouet et al. 2016; Ben Daoud et al. 2016). Considering the performance of ECMWF is often better than other NWP models (Park et al. 2008; Niu and Zhai 2013; Zhou et al. 2015; Swinbank et al. 2016), Zhou and Zhai (2016) used large-scale circulation variables derived from the ECMWF ensemble mean as predictors to establish a new statistical downscaling model called the Key Influential Systems Based Analog Model (KISAM) for PEPE forecasts in the YHRV. Because of the superior skill of KISAM in reforecasting, we are motivated to investigate the performance of KISAM and ECMWF in real-time operational forecasts of PEPEs.

Two PEPEs that occurred in the YHRV during summer 2016 and met the criteria specified by Chen and Zhai (2013) were selected for our performance assessment. Operational forecast products based on NWP models failed to correctly capture the meridional locations of the rainbands of these two PEPEs, which resulted in an ineffective deployment of flood prevention strategies. To improve the performance of the NWP models for PEPEs, forecast verification is crucial (Sukovich et al. 2014). Therefore, this study aims to verify the performance of ECMWF, including all ensemble members and its ensemble mean, as well as KISAM, for PEPE forecasts. Their strengths and weaknesses will be analyzed to provide forecasters with useful guidelines to employ the NWP models and statistical downscaling models effectively. Furthermore, through comparison of the atmospheric circulation and physical conditions affecting the process of precipitation production in ECMWF and KISAM, we will further address the leading factors behind their advantages and disadvantages in PEPE forecasts.

The data employed and the methodology are described in section 2, followed by an introduction of the characteristics of the two PEPEs and corresponding circulations in section 3. A comparative assessment of the two objective forecasts for the two PEPEs is given in section 4. Section 5 presents an analysis of the superiority of KISAM for longer lead times, with conclusions and further discussion provided in section 6.

2. Data and methodology

a. Data

The observed daily precipitation dataset from about 2200 stations covering China in summer 2016 is provided by the National Meteorological Center of the China Meteorological Administration (CMA). Daily atmospheric reanalysis data from the ERA-Interim dataset, including wind vectors, geopotential height, and relative humidity in the summers of 1981–2010 and 2016, with a horizontal resolution of $1^\circ \times 1^\circ$, are used to depict the
climatology and circulation features during summer 2016, respectively. ERA-Interim is the latest global atmospheric analysis produced by the ECMWF and is based on the 2006 version of the integrated forecasting system (Cy31r2; Dee et al. 2011). Considering KISAM is based on an analog method, daily precipitation data from 756 weather stations over China during 1951–2010 from the National Meteorological Information Center of the CMA are used to obtain the precipitation for the historical PEPE records. Re-analysis data from the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR; Kalnay et al. 1996) serve as the historical pool of large-scale circulation for matching analogs. The daily large-scale variables in the summers of 1951–2010, with a horizontal resolution of 2.5⁰ × 2.5⁰, from the NCEP–NCAR reanalysis, comprise the zonal wind at 200 hPa, geopotential height and vertical velocity at 500 hPa, and zonal and meridional water vapor transport at 700 hPa.

The predicted large-scale circulation variables and total precipitation from the ECMWF ensemble, alongside a control run, are retrieved from initializations at 1200 UTC on each day from 1 June to 6 July in summer 2016, for lead times of up to 15 days with a forecast step of 1 day. The control forecast, a single member of the ECMWF ensemble, is usually an unperturbed analysis generated by a data assimilation procedure. The ensemble-mean forecasts are derived by equally weighting all ensemble members. The horizontal resolution of the total precipitation is 1° × 1°. The large-scale circulation variables obtained from the ECMWF ensemble at a 2° × 2° horizontal resolution also comprise the same variables that are obtained from the NCEP–NCAR reanalysis. The forecasts of vertical velocity at 500 hPa from the ECMWF ensemble mean in the medium range of up to 11 days, with a forecast step of 6 h, initialized on each day from 1 June to 6 July at a 2° × 2° horizontal resolution, are also used.

For ease of comparison, the station precipitation data are interpolated onto an identical grid of 1° × 1° through the Cressman interpolation method, and circulation variables are interpolated onto an identical grid of 2.5° × 2.5° through bilinear interpolation.

b. Brief introduction to KISAM

KISAM is trained and verified using the 25 PEPEs that Chen and Zhai (2013) identified in the YHRV (26°–34°N, 112°–121°E) over 1951–2010. The schematic process of KISAM is displayed in Fig. 1. Based on the circulation patterns from upper to lower levels conducive to the PEPEs identified by Chen and Zhai (2014), KISAM selects four large-scale atmospheric variables from the circulation patterns as predictors, including zonal wind at 200 hPa, geopotential height at 500 hPa, and zonal and meridional water vapor transport at 700 hPa. The forecast data for the predictors used by KISAM are the large-scale circulation variables obtained directly from ECMWF at lead times of 1–15 days. The similarity metric, with the calculation shown in the orange box in Fig. 1, is the weight-assigned cosine of the angle between two patterns. The variables $A$ and $B$ represent two atmospheric fields, and $n$ is the total number of grid points. The weight function $G(i)$ is the
normalized absolute correlation coefficient over the study region between each predictor at each grid point and the area-averaged precipitation over the training period. The weights assigned can help emphasize the role in matching analogs of those key large-scale circulation systems that have larger correlations with the PEPEs. The similarity is measured between forecasts and the records in the historical pool for the selected four predictors individually, and the resulting four similarities $S_{200}$, $S_{500}$, $S_{700\text{qu}}$, and $S_{700\text{qv}}$ are then combined to obtain one integrated similarity score. As the equation in the right-hand box in the third row of Fig. 1 illustrates, the normalized weights $P_{200}$, $P_{500}$, $P_{700\text{qu}}$, and $P_{700\text{qv}}$ are assigned to the four similarities, respectively, to get the combined similarity. Another parameter, $P_{cv}$, introduced in KISAM, is a critical value for judging whether a historical record is sufficiently analogous to the target day. These parameters are determined by an optimization algorithm known as a “cuckoo search” in the training process to optimize KISAM. The optimization is designed to seek one set of parameters ($P_{200}$, $P_{500}$, $P_{700\text{qu}}$, $P_{700\text{qv}}$, and $P_{cv}$) to maximize the threat score of the heavy precipitation forecast for the historical PEPE days and give KISAM the best capability to discriminate PEPE days from non-PEPE days. The parameters obtained after training show that the weights assigned generally show a decrease from lower to upper levels. This suggests that the lower layer has a more crucial influence on the analog scheme. When the obtained integrated similarity between a historical record and the target day is larger than $P_{cv}$, KISAM identifies the historical record as an analog. Circulation and physical conditions such as vertical motion on these analogous days actually control the precipitation that KISAM produces. In the end, if KISAM successfully finds three or more analogs, it will issue a precipitation forecast using the weighted average of the three most similar analogs. The weights are the similarity scores of the analogs. Otherwise, KISAM will not produce a precipitation forecast and assigns missing values for the precipitation field. More specific details of KISAM are described by Zhou and Zhai (2016).

c. Analysis method

Verification of the precipitation forecasts from the ECMWF ensemble members and ensemble mean, as well as from KISAM, is conducted with respect to the observed precipitation collected from about 2200 stations geographically covering China during summer 2016. Verifications of vertical velocity at 500 hPa leading to the production of heavy precipitation in ECMWF and KISAM are performed against the NCEP–NCAR reanalysis in summer 2016. Vertical velocity is not used by KISAM as a predictor and is not involved in the downscaling process. Thus, the vertical velocity examined in KISAM is the average vertical velocity obtained from the NCEP–NCAR reanalysis of the three most analogous historical records. We use the equitable threat score (ETS) to measure the fraction of observed events that are correctly predicted, and it accounts for hits associated with random chance (Wang 2014). The ETS ranges from $-\frac{1}{3}$ to 1, with 1 indicating a perfect score. It is defined as

\[
ETS = \frac{\text{hits} - \text{hits}_{\text{random}}}{\text{hits} + \text{misses} + \text{false alarms} - \text{hits}_{\text{random}}} \quad \text{and} \quad \text{hits}_{\text{random}} = \frac{\text{hits} + \text{misses} + \text{false alarms}}{\text{hits} + \text{misses} + \text{false alarms} + \text{correct negatives}},
\]

where hits, misses, false alarms, and correct negatives are the four combinations of forecasts (yes or no) and observations (yes or no), called the joint distribution. The categorical bias score (BS) is used to measure the forecast performance for the frequency of precipitation exceeding different thresholds. It is calculated as

\[
BS = \frac{\text{hits} + \text{false alarms}}{\text{hits} + \text{misses}}.
\]

The perfect score of BS is 1. Rainfall in eastern China is usually concentrated in a zonal belt and moves progressively toward the north along with the summer monsoon advancing northward (Tao 1980; Ding 1992; Samel et al. 1999). Therefore, the meridional location of the rainband is an important measure when evaluating a model’s PEPE forecast performance. To indicate the direction of deviations of the meridional location, the zonal-mean accumulated precipitation over 112°–121°E during PEPEs is investigated. Moreover, the precipitation root-mean-square error (RMSE) is calculated to quantitatively examine the average magnitude of the forecast errors for the PEPE forecast. It is noted that when KISAM fails to issue a precipitation forecast, the constant precipitation field of missing values will be taken as zero when calculating the RMSE. Besides verification of PEPE forecasts from ECMWF and KISAM, the investigation will focus on the key large-scale circulation and physical processes that are critical for the production of precipitation and subsequently affect PEPE forecast accuracy in these two models. When verifying daily precipitation during a PEPE, we aim to investigate the average performance of a single 15-day forecast initialized on one selected lead day for the whole PEPE process, rather than examine the mean performance
of multiple 15-day forecasts with the same forecast length corresponding to each day of a PEPE. Therefore, forecast lead time for a PEPE in this paper refers to the number of days in advance of the first day of the PEPE. Taking the first PEPE (18–22 June 2016) as an example, a 3-day lead time means the forecast is initialized on 15 June. In this case, the forecasts for 19–22 June are in fact at lead times of 4–7 days. Thus, for a 15-day forecast, a 5-day PEPE can be forecasted as early as 11 days in advance. Following Zhou et al. (2015) and Niu and Zhai (2013), indices for the meridional location of the WPSH and the East Asian subtropical westerly jet (EASWJ) are calculated. The average axis of the EASWJ is defined as the average of the latitude with maximum zonal wind above 30 m s\(^{-1}\) at 200 hPa at each longitude over 110\(^\circ\)–130\(^\circ\)E. The mean position of the WPSH ridge is determined as the mean latitudinal position of the high pressure ridge (\(u = 0\) and \(\partial u/\partial y < 0\), where \(u\) represents zonal wind) surrounded by the 588-dagpm contour at 500 hPa to the north of 10\(^\circ\)N over 110\(^\circ\)–130\(^\circ\)E.

### 3. Characteristics of the two PEPEs and corresponding circulations

With the summer monsoon shifting northward, the rainband in eastern China shifts to the YHRV in June (Tao 1980). During this period, PEPEs are prone to occur, with cold and dry air persistently encountering anomalous warm and moist air in the same region (Ding and Reiter 1982; Chen and Zhai 2014). Two PEPEs in summer 2016 that induced severe flooding in the YHRV are chosen as our verification cases; one occurred during 18–22 June (hereafter event 1), and the other took place over 1–6 July (hereafter event 2). From the distribution of the observed accumulated precipitation during these two PEPEs in Figs. 2a and 2b, it can be seen that, for both events, the rainbands are generally located along the Yangtze River. The stations with accumulated rainfall greater than 100 mm are mainly concentrated between 29\(^\circ\) and 33\(^\circ\)N for event 1, as shown in Fig. 2a. Compared to event 1, the stations with heavy rainfall accumulation in event 2 appear to be more concentrated along the Yangtze River. The two PEPEs are both very strong, with the maximum accumulated precipitation reaching 200–300 mm, and above 500 mm, respectively. According to official statistics, event 2 was the most intense precipitation event in 2016 in China. Severe flooding related to these two PEPEs struck the lower reaches of the Yangtze River, resulting in substantial damage to local agriculture. From the time–latitude cross section of daily precipitation along 112\(^\circ\)–121\(^\circ\)E
(Fig. 2c), it is noted that the rainband shifts from south to north during event 1, while it remains relatively steady during event 2. It can also be clearly seen that the rainfall amount in event 1 is relatively low compared to event 2 (Fig. 2c).

As indicated in Fig. 3, the large-scale atmospheric circulations of the two PEPEs are both characterized by the typical double-blocking high circulation patterns described by Chen and Zhai (2014). At 200 hPa, in Figs. 3a1 and 3a2, a subtropical jet stream is located to the north of the YHRV and near the Sea of Japan. The YHRV is located at the right-hand entrance of the jet stream above the Sea of Japan. Compared to the climatology over 1981–2010, the South Asia high is more intense and shifts eastward. These conditions both provide strong upper-level divergence over the YHRV, which is favorable to the occurrence of a PEPE (Samel et al. 1999; Zhu et al. 2010). At 500 hPa, a pair of blocking highs is located over the Ural Mountains and the Sea of Okhotsk, respectively (Figs. 3b1 and 3b2). The WPSH shifts westward and produces an anomalous anticyclone with southwesterly flow transporting

FIG. 3. Averaged atmospheric circulation patterns at different levels during the two PEPEs showing (left) event 1 and (right) event 2. Among them, results are shown for (a1),(a2) 200, (b1),(b2) 500, and (c1),(c2) 700 hPa. The blue lines in all panels represent the Yangtze River and the Yellow River. At 200 and 500 hPa, black lines represent geopotential height contours, and red lines indicate climatological geopotential height ≥ 1252 dagpm west of 150°E at 200 hPa and ≥ 588 dagpm east of 90°E at 500 hPa, over 1981–2010. Shading at 200 hPa indicates the magnitude of the westerly wind ≥ 30 m s⁻¹, whereas at 500 hPa it represents the anomaly between geopotential height during the PEPEs and the climatology over 1981–2010. At 700 hPa, vectors indicate wind with magnitude ≥ 6 m s⁻¹, and blue shading represents anomalies between specific humidity during the PEPEs and the climatology over 1981–2010 (g g⁻¹).
excessive moisture into the YHRV during the event compared to the climatology (Figs. 3c1 and 3c2). These atmospheric conditions all match the double-blocking high circulation pattern and the associated synoptic features reported by Chen and Zhai (2014), which play a significant role in the development and maintenance of the two PEPEs.

4. Comparative assessment of the two objective forecasts

As a result of the northward progression of the rainfall band in event 1 and the large amount of rainfall in event 2, it is very difficult for forecasters to precisely predict these extreme events. Objective forecasts from NWP or downscaling models have potential use for forecasters. Verification and comparison of forecasts from the ECMWF ensemble members and ensemble mean, as well as KISAM, are conducted for the two PEPEs in a medium forecast range of up to 10 days, to provide useful information for the improvement of the PEPE forecast.

Considering the general underestimation of global NWP models for precipitation in excess of 25 mm day$^{-1}$ (Ralph et al. 2010; Su et al. 2014; Sharma et al. 2017), the performance of ECMWF and KISAM is examined using the ETS for daily precipitation at a threshold intensity of 25 mm day$^{-1}$ with respect to the observations. For forecasts initialized on each day from 7 to 17 June and from 21 to 30 June for events 1 and 2 in the YHRV, the ETS is first calculated for precipitation at the 25 mm day$^{-1}$ threshold on all the days of each PEPE process, and then the ETSs are averaged over the whole PEPE process. Therefore, for a given lead day, multiday forecasts corresponding to the whole PEPE process are used to achieve the average verification score. As indicated by the average ETSs in Fig. 4, with the exception of KISAM at several lead times, KISAM and most ECMWF ensemble members show skill in detecting the location of precipitation at a threshold intensity of 25 mm day$^{-1}$ during these two PEPEs with a lead time of up to around 9 days. The ECMWF ensemble mean shows an obvious advantage over KISAM at lead times less than 4 days but performs poorly afterward, as indicated by ETS values nearly approaching 0. This is partly due to the smoothing effect by the ensemble average. In contrast, KISAM generally exhibits better skill than the ECMWF ensemble mean in the forecast range of 5–10 lead days, especially for event 1. Moreover, the ETSs of all ECMWF ensemble members appear to decrease with increasing lead time. The median ETSs of the ECMWF ensemble are all equal to or greater than zero for all lead times, which indicates that nearly one-half or more of the ECMWF ensemble members have skill in locating heavy rainfall during the process of the two events. It is noted that the ECMWF and KISAM ETSs are relatively low at longer lead times, which indicates that PEPE forecasting is also a challenge for the ECMWF ensemble and KISAM. Thus, there is still a major need to improve the prediction of PEPE in consideration of the related catastrophic consequences.

A comparison of KISAM and the ECMWF ensemble results reveals that the skill of KISAM is lower at lead times of 1–3 days. However, for event 1 (Fig. 4a), the ETSs of KISAM surpass the median of the ECMWF ensemble at lead times of 4–10 days. KISAM is superior to 75% or more of the members of ECMWF in scoring better ETSs at lead times of 5–9 days. It is noteworthy that the KISAM forecast does a better job at capturing the location of rainfall at the 25 mm day$^{-1}$ threshold for event 1 when the lead time increases from 1 to 8 days. Similarly, the ETSs of the control forecast also show an increase at lead times of 8 or more days compared to the middle-range forecast. This can be explained by the fact that verification is conducted against these two individual cases in this paper. However, the generally decreasing skill of the NWP models with increasing lead time reported in the literature is achieved by verification scores averaged over many days or seasons (Park et al. 2008; Lavers and Villarini 2013). For these two cases, the location of the rainband is indeed more accurately predicted at longer lead times. This point is evidenced by the well-predicted meridional location of the rainband at longer lead times by KISAM (see Fig. 6c1, which will be discussed later). For event 2 (Fig. 4b), the skill of KISAM is compatible with the median of the ECMWF ensemble at 4–10 days, with exceptions at lead times of 5 and 7 days. In short, at lead times longer than 3 days, the accuracy of KISAM is at least one-half of the ECMWF ensemble members in locating heavy precipitation during the two PEPEs.

To assess the quantitative accuracy of these forecasts in predicting heavy daily rainfall, the RMSE for the forecasting of precipitation at the 25 mm day$^{-1}$ threshold is investigated (figure not shown). RMSE for precipitation above 25 mm day$^{-1}$ is calculated between the precipitation observations that exceed 25 mm day$^{-1}$ and the corresponding forecasts. The ECMWF ensemble mean has the advantage over most ensemble members, including the control run. For event 1, KISAM shows the best performance at lead times of 2–9 days. The performance of KISAM for event 2 at the 25 mm day$^{-1}$ threshold shows dramatic variation with increasing lead time. However, with the exception of lead days 5 and 7, the performance of KISAM can still match that of the ECMWF ensemble mean forecast for lead times longer
than 3 days. To discriminate the forecast error of the rainfall amount from location error, categorical bias scores are explored for the ECMWF ensemble and KISAM at different lead times and precipitation thresholds (Fig. 5). For event 1, KISAM overpredicts light precipitation ($\geq 1$ mm day$^{-1}$) and heavy precipitation ($\geq 25$ mm day$^{-1}$; Figs. 5a1 and 5b1). In contrast, the ECMWF ensemble performs better, with most members achieving perfect bias scores for light precipitation. However, the 25 mm day$^{-1}$ precipitation is underpredicted by the ECMWF ensemble at lead times of 1–8 days; also, at lead times longer than 8 days, about half of the members overpredict heavy precipitation while the other half show an underestimation. For event 2, underestimation of light and heavy precipitation is seen for both the ECMWF ensemble and KISAM, with the larger error for heavy precipitation. It is also noted that the ECMWF ensemble mean shows significant underprediction of the 25 mm day$^{-1}$ precipitation frequency due to the smoothing effect. Summarizing the results from the bias scores, it seems that the ECMWF ensemble performs well for the frequency of light precipitation ($\geq 1$ mm day$^{-1}$), but shows a general tendency to underpredict (BS $< 1$) heavy precipitation ($\geq 25$ mm day$^{-1}$). For KISAM, consistent overestimation or underestimation is found for light and heavy precipitation.

To determine the direction of forecast deviations for the meridional location of the entire PEPE process, which the ETS and BS do not indicate, the forecasts of the zonal-mean accumulated precipitation during each PEPE process averaged over 112°–121°E are verified for lead times of 1–9 days. For each forecast initialization time, a single forecast of zonal-mean multiday total precipitation throughout the PEPE is used to match with the observed precipitation accumulation. Figure 6 shows the meridional location of the precipitation accumulation at lead times ranging from 9 days to 1 day, placed successively from far left to right. The observed meridional locations of precipitation accumulated during 18–22 June and during 1–6 July for events 1 and 2, respectively, are placed on the far right. The discrepancy
between the forecasts at all lead times and the observation clearly indicates a bias in predicting the latitude of the zonal-mean heavy rainfall center during the PEPE. The forecasts include the ECMWF ensemble mean, the ECMWF control forecast, and KISAM. Since the ensemble mean represents the average locations of precipitation centers depicted by all ensemble members, verification is conducted for the ECMWF ensemble mean instead of all ECMWF ensemble members. The ECMWF control run is selected from the ECMWF ensemble as an independent reference. For event 1, the observed zonal-mean precipitation accumulation above 50 mm is located between 30° and 33°N. Rainfall from ECMWF ensemble mean was slightly underestimated during the entire forecast period and placed more northward than the observation, notably at lead times greater than 4 days (Fig. 6a1). The ECMWF control run better captures the precipitation location and amount on the first two lead days, but shows considerable underestimation (particularly at 3–6 days) and displacement error of heavy rainfall for longer lead times (Fig. 6b1). The rainfall pattern is depicted well by KISAM, albeit the amount is overestimated and the locations of the maxima are shifted 2° southward, especially at lead times of 2–3 days (Fig. 6c1). For event 2, as shown in Fig. 6a2, the precipitation maxima predicted by the ECMWF ensemble mean gradually shifts to the north with increasing lead time, and reaches 34°N compared to the observed 31°N at lead day 7. The same northward deviation is found for the ECMWF control forecast (Fig. 6b2). From Fig. 6c2, it is noted that the deviations of KISAM for forecasts of the meridional location of event 2 at various lead times are not consistent. The forecast precipitation center is located about 2° northward at lead times of 8–9 days, and slightly to the south at lead times of 1–7 days (Zhou and Zhai 2016). A point to note is that all three forecasts predict weaker intensities of the accumulated precipitation for event 2.
Nonetheless, the ECMWF control forecast outperforms the other two forecasts. The weak precipitation accumulation of event 2 produced by KISAM is due to its failure to produce forecasts for the first 2 days of event 2; thus, the precipitation on these 2 days is set as missing values and not counted in the accumulation.

To sum up, KISAM is generally as good as the median of all the ensemble members of ECMWF at lead times longer than 3 days for detecting the location of heavy precipitation. In addition, the investigation based on the RMSE and BS indicates that KISAM also performs as well as most ensemble members of the ECMWF in forecasting the frequency and amount of heavy precipitation. With respect to the meridional location of the two PEPE processes, KISAM forecasts are generally closer to the observation at lead times longer than 3 days, due to the significant northward divergence for forecasts of the ECMWF ensemble mean and the control forecast. KISAM uses the large-scale atmospheric variables from the ECMWF ensemble mean to down-scale precipitation over the PEPE. Therefore, the forecasts generated by KISAM and ECMWF are influenced partly by the same large-scale atmospheric background. The underlying reason contributing to the different
levels of forecast performance between KISAM and ECMWF is expounded upon in the next section.

5. Analysis of the superiority of KISAM for longer lead times

As Chen and Zhai (2014) reported, PEPEs typically result from concurrent anomalies of key large-scale systems such as subtropical jets, blocking highs, and the WPSH. Moreover, the predictors of KISAM contain the large-scale variables obtained directly from ECMWF. For the spatial field of these predictors, the key large-scale systems are all assigned more weight in the statistical downscaling scheme of KISAM, which was introduced in section 2b. Therefore, the skill of the ECMWF ensemble mean in predicting these key large-scale systems has significant influence on the performance of KISAM and the ECMWF ensemble for PEPE forecasts. The analysis is therefore focused on the skill of the ECMWF ensemble mean in predicting these key large-scale systems during the two PEPEs.

As mentioned above, KISAM is based on an analog method and uses the large-scale circulation depicted in the ECMWF ensemble mean forecast to search for analogous circulations from the historical pool. Even at the lead time of 9 days for event 1, the three most similar analogs that KISAM finds from the historical pool (Table 1) are almost all from the PEPE cases under the double-blocking high pattern identified by Chen and Zhai (2014). Since the analogous circulations are characterized by the double-blocking high pattern, the synoptic patterns of the predictors forecasted by the ECMWF ensemble mean must also exhibit a double-blocking high pattern. This reflects indirectly that the large-scale circulation of KISAM taken from the ECMWF ensemble mean well reproduces the double-blocking high patterns during these two events. In addition, the performance of the ECMWF ensemble mean for predicting the subtropical jet and WPSH is examined. Figures 7a1 and 7a2 illustrate that, for both events, the subtropical jet is predicted well by ECMWF in terms of extent, intensity, and location at the lead time of 3 days. At 6 days, the jet axis deviates significantly to the north at 90°–120°E during event 1. The performance of ECMWF for the subtropical jet during event 2 at 6 days is comparable to that at 3 days. With the lead time increasing to 9 days, the intensity and extent of the subtropical jet as predicted by ECMWF are both underestimated. For the WPSH, which is usually depicted by geopotential heights of 586 and 588 dagpm, the prediction of event 1 shows good skill, with the extent and intensity close to the observation at the lead time of 3 days (Fig. 7b1). The predicted northern edge of the WPSH at 3 and 6 days is located farther north than observed. Interestingly, ECMWF describes the northern edge of the WPSH more accurately at the lead time of 9 days. For event 2, the extent of the WPSH is overestimated slightly at the lead times of 3, 6, and 9 days (Fig. 7b2). The northwest quadrant of the WPSH deviates a little to the north at 3, 6, and 9 days for event 2. However, the northward deviation of the northwest quadrant of the WPSH at 9 days appears to be smaller than that at 3 and 6 days. In addition, water vapor transport is predicted well for events 1 and 2, with abundant moisture advected to the YHRV (Figs. 7c1 and 7c2).

The forecast errors of the ECMWF ensemble mean for the mean meridional position of the WPSH and the EASWJ over the region 110°–130°E on each day of the two PEPEs are quantified using indices of the average WPSH ridge and EASWJ axis. Positive and negative errors indicate northward and southward deviations, respectively. From Fig. 8a1, it can be seen that the southward deviation of the forecast at a 6-day lead for the average WPSH ridge increases with valid time during event 1. The northward deviation of the average WPSH ridge forecast at 9 days turns to a southward deviation with time during event 2 (Fig. 8a2). With the exception of these two circumstances, the deviations of the WPSH ridge forecast at the lead times of 1, 3, 6, and 9 days are relatively steady over the two PEPEs. For the average EASWJ axis, the forecasts all locate this axis to the north of that observed at lead times of 1, 3, and 6 days, with the deviation more evident at 6 days (Fig. 8b1). This behavior confirms the result shown in Fig. 7a1. For event 2, forecasts of the average EASWJ axis from the ECMWF ensemble mean show general southward deviations at the lead time of 1 day. The direction of the EASWJ axis forecast deviation shifts from

---

**Table 1. Dates of the three most similar analogs found by KISAM for event 1, at 9-day lead time. Boldface indicates that the day is among the cases under the double-blocking high pattern identified by Chen and Zhai (2014).**

<table>
<thead>
<tr>
<th>Objective day</th>
<th>18 Jun</th>
<th>19 Jun</th>
<th>20 Jun</th>
<th>21 Jun</th>
<th>22 Jun</th>
</tr>
</thead>
</table>
northward to southward during event 2 at 3 and 6 days. Deviations of the EASWJ axis forecast at 9 days oscillate around zero and are generally southward. The mean absolute error (MAE) of the WPSH ridge forecast averaged daily during each PEPE is computed for all lead times. The maximum MAE among all lead times for the average WPSH ridge amounts to 1.1°8 and 1.3°8 for events 1 and 2, respectively. Similarly, the maximum MAEs for the average EASWJ axis reach 4° and 2.7° for events 1 and 2, respectively. As Niu and Zhai (2013) reported, the MAE of the mean position of the WPSH ridge over 110°–122.5°E is as large as about 3.5° at lead times of up to 15 days. Based on our previous work related to the evaluation of the EASWJ prediction in the NWP models (Zhou et al. 2015), the MAE of the EASWJ axis averaged over 100°–125°E can reach up to around 4° at the lead time of 15 days. The MAEs of the two indexes in these two studies are computed over many summer days.

Thus, for the two PEPE cases, the bias of the ECMWF ensemble mean in predicting the meridional location of the WPSH and the EASWJ up to 9 days ahead of the PEPE is smaller than the average error reported in the literature (Niu and Zhai 2013; Zhou et al. 2015). With increasing lead time, the forecast deviations of the intensity and meridional position of the key large-scale systems like the WPSH and the EASWJ show poor association with that of the PEPEs for ECMWF.
example, the performance of the ECMWF ensemble mean in depicting the WPSH and EASWJ is comparable at lead times of 1, 3, and 6 days for event 2. However, for the ECMWF ensemble mean and control run, the forecast of precipitation accumulation for event 2 shows an increasing magnitude in northward deviation with increasing lead time. This demonstrates that the meridional displacement of the PEPE from ECMWF is not completely dominated by the forecast deviations of the key large-scale systems. Therefore, the reasonable forecast errors for the key large-scale systems are unable to fully account for the deficiency of the ECMWF in PEPE forecasts, especially for the meridional location of the rainband. For this reason, further investigation is conducted on the essential physical conditions that affect the production of extreme precipitation forecasts for KISAM and ECMWF.

From a rather macroscopic perspective, long duration, abundant moisture, and strong ascending motion are vital physical conditions for extreme heavy precipitation, on top of the conditions needed for ordinary precipitation. The effect of long duration mainly manifests in the intensity of the precipitation accumulation, as verified and discussed above. In addition, the abundant moisture transported to the YHRV is predicted well by ECMWF at the lead time of 6 days, as the verifications show in Figs. 7c1 and 7c2. Since the conditions of long duration and abundant moisture are depicted well in the forecasts from ECMWF, attention is then turned to vertical velocity. Strong ascending motion, which provides a mechanism for triggering convection and large-scale moisture convergence, is also a significant condition for heavy rainfall, with its importance stressed in many studies (Lau and Kim 2012; Lenderink et al. 2017). In view of the above facts, the description of ascending motion in ECMWF is suspected to be a factor in the deficiency of ECMWF in the PEPE forecast. Therefore, efforts are devoted to the depiction of vertical velocity at 500 hPa, which affects the production of precipitation forecast in ECMWF. For comparison, the vertical velocity at 500 hPa related to the precipitation production process in KISAM is also investigated.

![Fig. 8. Forecast error (°) of the ECMWF ensemble mean forecast for the mean position of the WPSH ridge and EASWJ axis over 110°–130°E at lead times of 1, 3, 6, and 9 days: (a1) event 1 for the average WPSH ridge, (b1) event 1 for the average EASWJ axis, (a2) event 2 for the average WPSH ridge, and (b2) event 2 for the average EASWJ axis. The legend in (a1) is applicable to all panels.](https://example.com/fig8.png)
ECMWF provides a direct product of the vertical velocity forecast in the medium forecast range of up to 10-day lead. However, KISAM does not use the vertical velocity forecast from NWP models, nor does it directly produce forecasts of vertical velocity. Since KISAM is based on an analog method, its precipitation forecast is actually regulated by the corresponding weather conditions on the analogous historical days. Hence, vertical velocity, which has an impact on the production of precipitation forecasts in KISAM, is obtained from the analogous days that KISAM finds at different lead times.

Vertical velocities at 500 hPa that affect the precipitation output of the ECMWF ensemble mean, the ECMWF control forecast, and KISAM are verified. The observed meridional evolution of vertical velocity at 500 hPa during event 1 is illustrated as a time–latitude cross section in Fig. 9d. The vertical velocity also shifts from south to north in parallel with the northward shift of the rainband during this event. Predictions of the vertical velocity from the ECMWF ensemble mean with a forecast step of 6 h basically reproduce the shift in the observations at the lead times of 1 and 3 days (Figs. 9a2 and 9a3). However, the ECMWF ensemble mean overpredicts the intensity of the ascending motion at all initialization dates. Furthermore, the forecasts of the time of occurrence of the strong ascending motion from the ECMWF ensemble mean are not quite consistent with the observations. When the forecast is initialized on 12 June for event 1, the ECMWF ensemble mean fails to capture the northward shift of the 500-hPa vertical velocity (Fig. 9a1). The forecasts of vertical velocity from the ECMWF ensemble mean deviate notably to the north of the observations. Interestingly, the performance of the ECMWF ensemble mean in predicting the location of the vertical velocity corresponds well to its performance in capturing the location of precipitation at different lead times during event 1. The

![Figure 9: Observed and predicted time–latitude cross sections of vertical velocity at 500 hPa during event 1: (a1)–(a3) ECMWF ensemble mean forecasts; (b1)–(b3) ECMWF control forecast initialized at 12, 15, and 17 Jun; (c1)–(c3) vertical velocity at 500 hPa averaged over the three most similar historical analogous days that KISAM finds for forecasts initialized at 9, 12, and 15 Jun; and (d) the observed time–latitude cross section of the vertical velocity during event 1. The initializing days are marked in the bottom-left corners.](image-url)
directions of the meridional deviation for the forecasts of vertical velocity and the accumulated precipitation from the ECMWF ensemble mean are consistent as lead time varies. Forecasts of vertical velocity from the ECMWF control run at all initialization times are stronger and more discontinuous in time. The depiction of the meridional evolution of the vertical velocity is similar to that of the ECMWF ensemble mean (Figs. 9b1–b3). Vertical velocities from KISAM fail to show the northward shift at all lead times (Figs. 9c1–c3). Nevertheless, the vertical velocity from KISAM is located between 26° and 34°N, with close agreement to the observations. This means the evolution of vertical velocity from KISAM coincides to a large extent with the observed meridional evolution. For the forecast initialized on 9 June, due to the northward deviation of vertical velocity in the ECMWF ensemble mean, KISAM shows better performance in predicting the vertical velocities in the process of precipitation production. For event 2, similar results are seen. As depicted in Fig. 10, predictions of vertical velocity from the ECMWF ensemble mean show a clear northward bias at all lead times (Figs. 10a1–a3). The degree of the northward deviation becomes more severe as the lead time increases, which corresponds to the larger northward deviation of precipitation accumulation with increasing lead time. The forecast deviations of the ECMWF control run for vertical velocity show similar trends to the meridional displacement of the whole precipitation process (Figs. 10b1–b3). With the exception of the first 2 days, which KISAM fails to predict in event 2, the vertical velocities related to precipitation production in KISAM show an apparent advantage over those depicted in ECMWF, at longer lead times (Figs. 10c1–c3). Therefore, the quality of the depictions of the vertical velocity during a PEPE for ECMWF and KISAM has a dominant influence on their performance in the forecast of a PEPE at various lead times. The northward deviation of the ascending motion is the leading cause of the deficiency of ECMWF in forecasting the meridional location of the PEPE at longer lead times. In contrast, vertical velocity related to precipitation production in KISAM is achieved from the NCEP–NCAR reanalysis, which appears to be better during PEPEs compared to
the ECMWF forecast. This provides a proper elucidation for the superiority of KISAM for the forecast of PEPEs at longer lead times.

6. Conclusions and discussion

Two PEPEs during summer 2016 that caused severe flooding in the YHRV posed a serious challenge to operational forecasters. In accordance with the regional PEPE definition introduced by Chen and Zhai (2013), these two cases are selected as verification cases for two objective forecast models. One is ECMWF, an outstanding NWP model from TIGGE, and the other is a statistical downscaling model using the large-scale atmospheric circulation output from ECMWF, called KISAM. An assessment and comparison of the two forecast models is conducted for the two PEPEs, and the factors affecting their forecast performance are analyzed. The main conclusions that we draw are as follows.

For the location of daily heavy precipitation with amounts no less than 25 mm day\(^{-1}\) during the PEPE, about one-half or more of the possible solutions in the spectrum of the ECMWF ensemble show good skill at all lead times. However, the ECMWF ensemble shows a reduction in skill as the lead time increases, and it shows a general underestimation of heavy precipitation (≥25 mm day\(^{-1}\)). At lead times longer than 3 days, KISAM outperforms nearly one-half or more of all the ensemble members and the ensemble mean of ECMWF, especially for event 1. Moreover, at longer lead times, KISAM also shows an advantage over the ECMWF ensemble and the control forecasts in terms of the meridional location of the two PEPE processes. It is interesting that KISAM achieves higher ETSs at longer lead times, which is also found for the ECMWF control forecast. This can be attributed to the fact that the evaluation is conducted against only two PEPE cases, while the general tendency of decreasing skill as lead time increases is based on a larger number of samples.

Since the large-scale variables from the ECMWF ensemble mean forecast are used by KISAM as predictors, the performances of KISAM and ECMWF for PEPE forecasts are both related to the forecast performance of these key large-scale systems during PEPE. In view of this, the ability of the ECMWF ensemble mean in predicting key large-scale systems during the PEPE is verified for lead times of up to 9 days. This demonstrates that good correspondence between the deviation in latitude of the key large-scale systems and the meridional displacement of the precipitation accumulation does not exist. According to further verifications of the ascending motion, which affects the production of heavy precipitation in ECMWF and KISAM, the directions of the meridional deviation in the ECMWF ensemble mean forecasts of vertical velocity and accumulated precipitation show good correspondence with varying lead time. In addition, the vertical velocities obtained from the NCEP–NCAR reanalysis related to precipitation production in KISAM show an apparent advantage over those depicted in ECMWF, at longer lead times. Therefore, the quality in the depiction of vertical velocity during PEPE for ECMWF and KISAM has a dominant influence on their performance in the forecasting of the meridional location of the PEPE at various lead times.

For ECMWF, a hydrostatic model, the vertical velocity is determined by horizontal advection and pressure gradients, which results in inaccuracy. Moreover, errors in other complicated parameterization schemes, and initial and boundary conditions, may also contribute to the inaccuracy of the precipitation forecast. In contrast, statistical downscaling techniques like analog methods can easily establish nonlinear relationships between large-scale variables and local variables (Fernández and Sáenz 2003). For KISAM, the vertical velocity that actually regulates the production of the precipitation forecast is the average vertical velocity obtained from the NCEP–NCAR reanalysis of the three most analogous historical records. Therefore, the improved capture of the ascending motion related to precipitation production in KISAM explains its superior performance compared to the ensemble mean and half the members of ECMWF for the PEPE at longer lead times. In addition, based on the very thorough understanding of the circulation patterns conducive to PEPEs as described by Chen and Zhai (2014), KISAM combines the key circulation variables from the typical circulation patterns together as predictors. This may help KISAM search corresponding analogous patterns in the PEPE historical pool more precisely and produce better forecasts of PEPEs.

In summary, the ECMWF ensemble provides a skillful envelope of solutions for PEPE forecasts. However, as with other NWP models, the ECMWF forecast performance declines as lead time increases. In view of this, statistical downscaling techniques can serve as an important aid for forecasters to conduct forecasts of heavy precipitation in a medium forecast range of 4–10 days. Moreover, the forecast verification scores of the ECMWF ensemble and KISAM at longer times are still low. There is still a great need to improve our predictions of PEPEs, whether with statistical downscaling models or perhaps with a high-resolution dynamic model with state-of-the-art data assimilation.

Acknowledgments. This study was jointly supported by the National Key Technology Support Program
REFERENCES


