An Analog Technique to Improve Storm Wind Speed Prediction Using a Dual NWP Model Approach

JAEMO YANG AND MARINA ASTITHA

Department of Civil and Environmental Engineering, University of Connecticut, Storrs, Connecticut

LUCA DELLE MONACHE AND STEFANO ALESSANDRINI

National Center for Atmospheric Research, Boulder, Colorado

(Manuscript received 10 July 2017, in final form 16 September 2018)

ABSTRACT

This study presents a new implementation of the analog ensemble method (AnEn) to improve the prediction of wind speed for 146 storms that have impacted the northeast United States in the period 2005–16. The AnEn approach builds an ensemble by using a set of past observations that correspond to the best analogs of numerical weather prediction (NWP). Unlike previous studies, dual-predictor combinations are used to generate AnEn members, which include wind speed, wind direction, and 2-m temperature, simulated by two state-of-the-science atmospheric models [the Weather Research and Forecasting (WRF) Model and the Regional Atmospheric Modeling System–Integrated Community Limited Area Modeling System (RAMS–ICLAMS)]. Bias correction is also applied to each analog to gain additional benefits in predicting wind speed. Both AnEn and the bias-corrected analog ensemble (BCAnEn) are tested with a weighting strategy, which optimizes the predictor combination with root-mean-square error (RMSE) minimization. A leave-one-out cross validation is implemented, that is, each storm is predicted using the remaining 145 as the training dataset, with modeled and observed values over 80 stations in the northeast United States. The results show improvements of 9%–42% and 1%–29% with respect to original WRF and ICLAMS simulations, as measured by the RMSE of individual storms. Moreover, for two high-impact tropical storms (Irene and Sandy), BCAnEn significantly reduces the error of raw prediction (average RMSE reduction of 22% for Irene and 26% for Sandy). The AnEn and BCAnEn techniques demonstrate their potential to combine different NWP models to improve storm wind speed prediction, compared to the use of a single NWP.

1. Introduction

Analog-based methods using information of past observations, reanalysis, and numerical weather prediction (NWP) have been explored for various forecast fields. These methods have been applied to general circulation models (Radinović 1975; van den Dool 1989, 1994, 2007; Gao et al. 2006; Ren and Chou 2007), forecasting of summer monsoon subseasonal variability (Xavier and Goswami 2007), Southern Oscillation

index (Drosowsky 1994), mesoscale forecasts (Carter and Keislar 2000), temperature (Bergen and Harnack 1982; Livezey and Barnston 1988; Toth 1989), wind speed (Klausner et al. 2009), and precipitation prediction (Hamill and Whitaker 2006; Panziera et al. 2011; Toth 1989). Recently, Delle Monache et al. (2011) proposed a method to capture similarity between current and past forecasts for analog-space Kalman filter (ANKF) and weighted analogs (AN). Their analog-based technique is used for deterministic and probabilistic predictions of a range of parameters, including 1) atmospheric variables (e.g., wind speed, temperature, and relative humidity; Delle Monache et al. 2011, 2013; Mahoney et al. 2012; Nagarajan et al. 2015; Eckel and Delle Monache 2016); 2) improvement for surface particulate matter (PM_{2.5}) forecasts (Djalalova et al. 2015; Delle Monache 2017); 3) wind power forecasts

Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/MWR-D-17-0198.s1>.

Corresponding author: Marina Astitha, marina.astitha@uconn.edu

DOI: 10.1175/MWR-D-17-0198.1

© 2018 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).

(Alessandrini et al. 2015a; Junk et al. 2015a; Davò et al. 2016); 4) solar power prediction (Alessandrini et al. 2015b; Davò et al. 2016); 5) assessment of the economic impact of deterministic and probabilistic wind power forecast (Alessandrini et al. 2014); 6) reconstruction of historical wind speed data for wind resource estimates (Vanvyve et al. 2015; Zhang et al. 2015) and precipitation (Keller et al. 2017); 7) calibration of ensemble forecasts using ensemble model output statistics (EMOS; Junk et al. 2015b); 8) gridded probabilistic forecasts (Sperati et al. 2017); and 9) prediction of tropical cyclone intensity (Alessandrini et al. 2018).

The analog ensemble algorithm (AnEn; Delle Monache et al. 2013) searches for the best-matching past forecasts (analogues) to the current forecast at specific locations and forecast lead times. Single or multiple physical predictors from deterministic model predictions are used in the analog metric to find the best-matching analogues. The AnEn performance can be improved by selecting appropriate predictor combinations. For example, Delle Monache et al. (2013) found that a set of wind speed and direction, 2-m temperature, and surface pressure is reasonable for wind speed prediction, but better results in terms of error and correlation can be obtained if pressure is not included in the selected predictors. Junk et al. (2015a) explored predictor-weighting techniques to assign unequal weights to the predictors, resulting in improved results, as also shown by Alessandrini et al. (2015b). Although the weight term of the analog metric has been included in the initial study (Delle Monache et al. 2011), most of the studies (Delle Monache et al. 2011, 2013; Mahoney et al. 2012; Alessandrini et al. 2014; Nagarajan et al. 2015; Djalalova et al. 2015; Alessandrini et al. 2015a; Vanvyve et al. 2015; Zhang et al. 2015; Eckel and Delle Monache 2016) have not attempted to find optimal weights for the predictors (i.e., all weights are set to 1). Also, in previous studies, only a single deterministic forecast or a mean of ensemble predictions has been used to generate analogues. Table 1 outlines some of the AnEn applications that researchers have implemented so far and the one used in this work.

In this study, two NWP models and observations are combined to postprocess predicted wind speed during storms, such as those that had significant impacts on infrastructure and the environment. For the first time, AnEn uses multiple predictors from two atmospheric modeling systems [the Weather Research and Forecasting (WRF) Model and the Regional Atmospheric Modeling System–Integrated Community Limited Area Modeling System (RAMS–ICLAMS)] with a predictor-weighting technique (Junk et al. 2015a; Alessandrini et al. 2015b). The approach in this work

aims to improve the prediction of wind speed associated with a storm. The latter is the main driving force behind tree damages in the northeast United States, and it results in interruptions in the electric power grid that can last from hours to days, depending on storm severity. The correlation between storm severity and power outages has been investigated in recent studies and has motivated new ways to improve weather prediction accuracy (Yang et al. 2017; Wanik et al. 2015, 2017; He et al. 2017). Additionally, the performance of AnEn for high-impact tropical storms such as Irene (2011) and Sandy (2012) is investigated through the application of a bias-corrected analog ensemble (BCAnEn) to further improve the prediction. The paper is structured as follows. Section 2 describes the model configuration and observations; section 3 provides details about the methodology for AnEn and BCAnEn; section 4 contains a discussion on the results; and major findings are summarized in section 5.

2. Atmospheric modeling systems and observations

The storm set is composed of 146 storms over the years 2005–16 and includes thunderstorms, extratropical storms, and two major tropical storms (Irene and Sandy). The storms are selected based on the impacts caused to the environment and electric power network in the northeast United States [from 20 to >15 000 power outages defined as locations that require manual intervention to restore power (Wanik et al. 2015); data provided from Eversource Energy]. In the storm database, 18 occurred during spring, 43 during summer, 30 in the fall, and 55 during winter. Hourly wind speed NWP forecasts for the selected storms are simulated using two mesoscale meteorological modeling systems: the WRF Model (WRF–ARW version 3.4.1; referred to as WRF; Skamarock et al. 2008) and the RAMS–ICLAMS (referred to as ICLAMS; Cotton et al. 2003; Solomos et al. 2011). For each storm, WRF and ICLAMS are initialized 12–24 h before the peak of the storm defined by the strongest wind speed (for thunderstorms, we use 12 h). The selection of the simulation start time also reflects the timing of the first power outage report running in the Eversource network (Wanik et al. 2015). The duration of the simulation encapsulates the event by placing the peak of the storm approximately in the middle of the timeline.

Both models are configured with three nests with horizontal grid spacing of 18 (outer domain), 6 (inner-intermediate domain), and 2 km (inner domain). The innermost domain is the focus area in this study (Figs. 1a,b).

TABLE 1. Previous studies on AnEn (2011–17). Acronyms are as follows: Global Environmental Multiscale model (GEM); Community Multiscale Air Quality model (CMAQ); Kalman filter in analog space (KFAS, or ANKF); Kalman-filtered AnEn mean (KFAN); European Centre for Medium-Range Weather Forecasts (ECMWF) Ensemble Prediction System (EPS); ECMWF high-resolution (HRES); North American Mesoscale Forecast System (NAM); Rapid Update Cycle (RUC); mean, maximum, and minimum values and standard deviation (MMMS); hybrid ensemble (HyEn); Consortium for Small-Scale Modeling (COSMO) reanalysis at 6-km horizontal grid spacing (COSMO REA6); and Hurricane WRF (HWRF).

Author	Deterministic/ probabilistic	Raw model	Analyzed variable	Analog techniques	Weight optimization, bias correction
Delle Monache et al. (2011)	Deterministic	WRF	10-m wind speed	AN, ANKF	—
Mahoney et al. (2012)	Deterministic	RFTDDA	80-m wind speed	ANKF, ANKF+QR	—
Delle Monache et al. (2013)	Probabilistic	GEM	10-m wind speed, 2-m temperature	AnEn	—
Alessandrini et al. (2014)	Deterministic/ probabilistic	RAMS	Wind power, wind power income	AnEn	—
Alessandrini et al. (2015a)	Deterministic/ probabilistic	RAMS	Wind power	AnEn	—
Alessandrini et al. (2015b)	Deterministic/ probabilistic	RAMS	Solar power	AnEn	Weight optimization
Djalalova et al. (2015)	Deterministic	CMAQ	PM _{2.5}	KFAS, AN, KFAN	—
Junk et al. (2015a)	Deterministic/ probabilistic	ECMWF	Wind power	AnEn	Weight optimization
Junk et al. (2015a)	Probabilistic	ECMWF EPS	100-m wind speed	AnEn, analog-based EMOS (AN-EMOS)	Weight optimization
Nagarajan et al. (2015)	Deterministic	GFS, NAM, RUC	10-m wind speed, 2-m temperature, 2-m RH	AN, ANKF	—
Vanyve et al. (2015)	Deterministic/ probabilistic	MERRA	80-m wind speed (historical data reconstruction)	AnEn	—
Zhang et al. (2015)	Deterministic/ probabilistic	MERRA	100-, 80-, and 50-m wind speed (historical data reconstruction)	SSEn	—
Davò et al. (2016)	Deterministic/ probabilistic	RAMS (wind power), GEFS (solar irradiance)	Wind power, solar irradiance	AnEn, PCA+AnEn, MMMS+AnEn	Weight optimization
Eckel and Delle Monache (2016)	Probabilistic	GEM	10-m wind speed, 2-m temperature	AnEn, HyEn	—
Sperati et al. (2017)	Deterministic/ probabilistic	HRES, ECMWF EPS	10-m wind speed	AnEn	Weight optimization
Delle Monache (2017)	Deterministic/ probabilistic	CMAQ	O ₃ , PM _{2.5}	AnEn	Weight optimization
Keller et al. (2017)	Deterministic/ probabilistic	COSMO REA6	Precipitation (historical data reconstruction)	AnEn	Weight optimization
Alessandrini et al. (2018)	Deterministic/ probabilistic	HWRF	Tropical cyclone intensity	AnEn	Weight optimization, bias correction
This study	Deterministic	WRF, ICLAMS	10-m wind speed	AnEn, BCAnEn	Weight optimization, bias correction

The National Centers for Environmental Prediction (NCEP) Global Forecast System ($1^\circ \times 1^\circ$, 6-hourly intervals) analyses (NCEP/NWS/NOAA/DOC 2007) and the GFS Final Analysis ($1^\circ \times 1^\circ$, 6-hourly intervals) data (NCEP/NWS/NOAA/DOC 2000) are used to initialize WRF and ICLAMS, respectively. For each storm, the two models produce hourly outputs for a maximum lead time of 61 h.

WRF utilizes the Thompson scheme for cloud microphysics (Thompson et al. 2008); Grell 3D scheme for convective parameterization (Grell and Dévényi 2002); Goddard for shortwave radiation (Chou and Suarez 1994); Rapid Radiative Transfer Model (RRTM) for longwave radiation (Mlawer et al. 1997); Noah for land surface scheme (Tewari et al. 2004); and the Yonsei scheme for the planetary boundary layer (PBL; Hong et al. 2006).

FIG. 1. Model domains (the inner rectangle box indicates the fine model domain) from (a) WRF and (b) RAMS-ICLAMS; (c) NCEP/NWS/NOAA stations over the northeast United States (circles) and elevation (shaded area).

ICLAMS uses a two-moment bulk microphysics scheme (Walko et al. 1995; Meyers et al. 1997) with an explicit cloud droplet activation (Nenes and Seinfeld 2003; Fountoukis and Nenes 2005); Kain–Fritsch cumulus parameterization for convective parameterization; RRTM for shortwave/longwave radiation (Mlawer et al. 1997); Land Ecosystem–Atmosphere Feedback version 3 (LEAF-3; Walko et al. 2000) as surface–atmosphere

interaction scheme; and Mellor–Yamada scheme for PBL (Mellor and Yamada 1982). Detailed information regarding the WRF and ICLAMS configuration is summarized in Table 2. All WRF and ICLAMS predicted storms are based on retrospective simulations.

Observations comprise hourly 10-m AGL wind speed from 80 meteorological terminal aviation routine weather report (METAR) stations in the northeast United States

TABLE 2. WRF and ICLAMS configuration.

	WRF	ICLAMS
Grid structure (three grids)	Grid spacing (dx): 18–6–2 km Vertical: 28 levels	Grid spacing (dx): 18–6–2 km Vertical: 50 levels
Nesting	Two-way nesting	Two-way nesting
Initial conditions	NCEP GFS ($1^\circ \times 1^\circ$, 6 h)	NCEP FNL ($1^\circ \times 1^\circ$, 6 h)
Cumulus scheme	Grell 3D scheme (Grell and Dévényi 2002)	Kain–Fritsch cumulus parameterization
Cloud microphysics	Thompson et al. (2008) scheme	Two-moment bulk scheme (Walko et al. 1995; Meyers et al. 1997); explicit cloud droplet activation scheme (Nenes and Seinfeld 2003; Fountoukis and Nenes 2005) with prescribed aerosols.
PBL	Yonsei scheme (Hong et al. 2006)	Mellor and Yamada (1982)
Boundary conditions	SST (NCEP GFS); topography (USGS GTOPO30, $30''$); land cover (USGS, $30''$); soil texture (FAO, $5'$; North America STATSGO, $30''$)	SST daily; NDVI (USGS, $30''$); topography (NASA SRTM90 v4.1, $3''$); land cover (USGS OGE, $30''$); soil texture (FAO, $2'$)
Radiation	Goddard for shortwave radiation (Chou and Suarez 1994); RRTM for longwave radiation (Mlawer et al. 1997)	RRTM for shortwave/longwave radiation (Mlawer et al. 1997)
Land surface	WRF Noah (Tewari et al. 2004)	LEAF-3 (Walko et al. 2000)

(Fig. 1c). Model outputs are interpolated to the METAR station locations with bilinear interpolation using four neighboring grid points. In addition, METAR observations for wind direction and temperature are used to evaluate model performance.

3. Methodology

a. Analog ensemble

The AnEn is formed with observations that correspond to past forecasts, which better match the current forecast and are referred to as *analog forecasts*. The following are the main steps of the algorithm: 1) at each station location and lead time, the best analog forecasts are selected based on the analog metric defined below, which quantifies the degree of analogy between the current forecast, for which AnEn is being generated, and forecasts available in the training dataset (the training dataset comprises available storms in the database and the selection of analogs depends on the lead time); and 2) observations corresponding to the best analog forecasts form the members of AnEn (Fig. 2).

The analog metric is defined as follows (Delle Monache et al. 2011, 2013):

$$\|F_t, A_{t'}\| = \sum_{i=1}^N \frac{w_i}{\sigma_{f_i}} \sqrt{\sum_{j=-\tilde{t}}^{\tilde{t}} (F_{i,t+j} - A_{i,t'+j})^2}, \quad (1)$$

where F_t is the current forecast at future time t ; $A_{t'}$ is an analog forecast with the same forecast lead time but valid at a past time t' ; N and w_i are the number of

atmospheric predictors and their weights; σ_{f_i} is the standard deviation of the time series of past forecasts of a given predictor at the same location (to normalize the contribution to the metric of predictors with different units); $\tilde{t}$ is half of the time window over which the analog metric is computed; and $F_{i,t+j}$ and $A_{i,t'+j}$ are the values of the current forecast and the analog of the atmospheric predictors in the time window.

Three model variables are used as parameters to select the best single-model analog sets from past forecasts: wind speed (WSPD), wind direction (WDIR), and temperature (TEMP). All parameters combined (i.e., three from WRF and three from ICLAMS) are used to select the combined-model analog set. An analog predictor-weight optimization is implemented for each station given the weight constraint $\sum_{i=1}^{\# \text{ of predictors}} w_i = 1$ and $w_i \in \{0, 0.1, 0.2, \dots, 1\}$ (phase 2 in Fig. 2). Note that we use a brute-force predictor-weighting strategy (Junk et al. 2015b) based on RMSE minimization. A training period of 145 storms is used for this phase, testing the 66 possible weight combinations for three-predictor AnEn (AnEn_{WRF} and AnEn_{ICLAMS}) and 3003 possible weight combinations for six-predictor AnEn (AnEn_{DUAL}) individually (the starting values for weight optimization is 1.0 for one predictor and zero for the rest of the predictors). Even though the computational cost of weight optimization using six predictors from two NWP models is affordable, the hybrid analog ensemble (HyEn; Eckel and Delle Monache 2016) would have to be considered in the case that AnEn is implemented with a larger number of predictors and NWP models. The HyEn is constructed by searching m analogs for each NWP member, so it is faster for weight optimization when

FIG. 2. A schematic diagram of the three phases to implement the AnEn technique.

compared to the AnEn using the predictors of multiple NWP models. HyEn is tested by selecting the best five analog ensemble members separately from each model (HyEn using BCAnEn_{WRF} and BCAnEn_{ICLAMS} is referred to as HyEn_{BC}) and constructing the ensemble of the 10 members. The optimized weights, using WRF and ICLAMS predictors separately, are higher for WSPD and WDIR across the 80 stations (Fig. 3). The median weights of WSPD and WDIR over the 80 stations are 0.6 and 0.3, respectively, for both AnEn_{WRF} and AnEn_{ICLAMS}. For AnEn_{DUAL}, the two WSPD predictors from WRF and ICLAMS have higher weights (median 0.3 for WRF WSPD and 0.4 for ICLAMS WSPD), compared to the other predictors. TEMP does not contribute significantly to find the optimal analog forecasts for AnEn in the cases analyzed here.

b. Bias-corrected analog ensemble

A bias-correction scheme is applied to improve the AnEn performance for wind speed (S. Alessandrini et al. 2017, meeting presentation). The AnEn introduces a conditional negative bias when predicting high wind events

of the forecast probability density function (PDF). This underestimation increases as the predicted event is rarer. This error is found to be dependent on the difference between the mean of the past analog forecasts and the current forecast. The approach used in this work is the simplest among those proposed by S. Alessandrini et al. (2017, meeting presentation). An adjustment factor is added to each AnEn member, which are past observations corresponding to past analog forecasts. This method is applied under the assumption that a linear relationship between predicted and observed wind speed holds when predictions are in the right tail of the PDF. This has been verified by fitting a linear regression with predicted wind speed as the response variable and the observed wind speed as the explanatory variable.

The optimal weights change after the application of bias correction (Fig. 3). Specifically, the weights of WDIR increase for all AnEn implementations: the median values are 0.4 for BCAnEn_{WRF}, 0.5 for BCAnEn_{ICLAMS}, and 0.3 for both WRF and ICLAMS. We speculate that the latter can be explained by the fact that high

FIG. 3. AnEn and BCanEn predictor weights for 80 stations (bar: median, box: interquartile range, whiskers: range, and error bars: minimum and maximum).

wind conditions, which are better predicted with the bias correction, are associated with specific wind directions.

c. Sensitivity of the analog ensemble size and length of the training set

The AnEn forecast is constructed by observations that match the past forecast analogs. AnEn forecasts are created for each individual model. That from WRF predictors is referred to as $AnEn_{WRF}$; that from ICLAMS predictors is referred to as $AnEn_{ICLAMs}$; and that for both models combined is referred to as $AnEn_{DUAL}$ (Table 3 presents the variables included in each AnEn model). The AnEn is verified using

leave-one-out cross validation (LOOCV), where a single storm is held out and the remaining storms compose the training set. The training dataset used in this study is restricted to past storms based on the need to improve wind speed prediction when a storm is imminent and does not include daily weather forecasts. We cannot speculate on the analog forecast performance when applied to daily weather forecasts, but in that case, the training dataset would have to include a wider range of meteorological conditions and not, like in this case, only past storms selected for their intensity.

A sensitivity analysis is conducted to define the AnEn size (phase 1 in Fig. 2). Multiple AnEn forecasts are

TABLE 3. Predictor combinations for AnEn and BCanEn.

AnEn model	Predictor						Optimization	Bias correction
	WRF			ICLAMs				
	WSPD	WDIR	TEMP	WSPD	WDIR	TEMP		
AnEn _{WRF}	✓	✓	✓	—	—	—	✓	—
AnEn _{ICLAMs}	—	—	—	✓	✓	✓	✓	—
AnEn _{DUAL}	✓	✓	✓	✓	✓	✓	✓	—
BCAnEn _{WRF}	✓	✓	✓	—	—	—	✓	✓
BCAnEn _{ICLAMs}	—	—	—	✓	✓	✓	✓	✓
BCAnEn _{DUAL}	✓	✓	✓	✓	✓	✓	✓	✓

created with 1, 3, 5, 7, 10, 15, 20, 25, and 30 members using Eq. (1) and weights w_i set to one. The ensemble size is chosen to minimize root-mean-square error (RMSE) between the ensemble mean and observations over the testing period (146 storms). In each AnEn implementation (i.e., AnEn_{WRF}, AnEn_{ICLAMS}, and AnEn_{DUAL}), the RMSE values progressively improve with increasing ensemble member size, and the trend reaches a plateau after 10 members to an almost constant value (not shown here). Therefore, 10 analog members are used as a reasonable ensemble size in this work.

To identify an optimal number of training storms, AnEn_{DUAL} and BCanEn_{DUAL} are implemented using an increasing number of randomly selected storms (20, 40, 60, 80, 120, 140, and 145). The comparison of RMSE for the raw models (WRF and ICLAMS), AnEn_{DUAL}, and BCanEn_{DUAL} computed with all available pairs of observations and predictions for each season shows that AnEn_{DUAL} starts to outperform ICLAMS in RMSE reduction when 60–80 storms are included in the training dataset (Figs. S1–S3 in the online supplemental material). BCanEn_{DUAL} exhibits significant RMSE reduction with only 20 training storms when compared to ICLAMS. Although all AnEn models produce the best results in terms of the global RMSE (Fig. S1) when using 145 storms for all seasons, the trend reaches a plateau after using 140 training storms to an almost constant value. Thus, the number of training storms to improve wind speed prediction should be at least 140 for all models to be comparable. We opted to use the maximum available training storms to get the best error reduction possible.

d. Data processing and evaluation

The first 6 h of the simulations are treated as the model spinup time and are discarded from the analysis. Zero and missing values for hourly wind speed observations are not included as modeled–observed pairs to generate AnEn (zero observed wind speed is predominantly wind speed below the instrument's threshold; by discarding those values, we avoid including a nonrealistic model bias in the analysis). Since we focus on an improvement of raw deterministic forecasts from WRF and ICLAMS, only deterministic verification scores are used in this study. WRF and ICLAMS serve as baselines to compare with the ensemble mean of AnEn and BCanEn models in a deterministic framework; thus, any probabilistic scores are not considered. Statistical metrics are calculated globally (use of all available model–observation pairs), temporally (at each lead time), and spatially (for each station separately). Global error metrics are also estimated for

wind direction and temperature [see Table S1; wind direction is treated using circular statistics as described by Jammalamadaka and Sengupta (2001)]. The common statistical metrics used are RMSE, mean bias (BIAS), and coefficient of determination R^2 . Note that the ensemble mean is used from AnEn and BCanEn throughout the manuscript.

4. Results and discussion

a. Seasonal analysis of global and event-based error

WRF and ICLAMS winds exhibit large scatter that is caused by over-/underestimation of wind speed (density scatterplots binned by 2 ms^{-1} shown in Figs. 4a,b). AnEn partially improves the raw model forecasts, in that it corrects the WRF and ICLAMS overestimations of wind speeds higher than 20 ms^{-1} and increases the density of correctly predicted wind speeds across the 1:1 regression line. However, the AnEn results show that it penalizes accurate forecasts of high wind speeds, as underestimation of winds larger than 20 ms^{-1} situated along the diagonal is introduced (Figs. 4c–e). BCanEn alleviates this underestimation problem of high wind speed with respect to AnEn (Figs. 4f–h; BCanEn_{WRF}, BCanEn_{ICLAMS}, BCanEn_{DUAL}).

Errors in the prediction in terms of RMSE and BIAS are estimated across seasons and wind categories to further assess the improvements gained by AnEn and BCanEn (Figs. 5, 6). The four wind categories include light breeze (group 1), moderate breeze (group 2), strong breeze (group 3), and gale/storm (group 4), following the World Meteorological Organization (WMO) classification (Table 4). RMSE values of models in groups 1 and 2 are lower than those in groups 3 and 4 in all seasons because the error metrics are proportional to the wind speed magnitude. This is also the main reason that for lower wind speed, AnEn and BCanEn do not show noticeable RMSE reduction relative to the raw models since the error is already small (e.g., group 1 in Fig. 5). For all seasons and groups, the best AnEn performance is obtained when applied to both models, as in AnEn_{DUAL} and BCanEn_{DUAL}, with the exception of group 1 (light breeze; $0.0 < \text{observed wind speed} \leq 3.1 \text{ ms}^{-1}$), where all models behave similarly with small errors and biases. AnEn reduces the RMSE for group 2 (moderate breeze: $3.1 < \text{observed} \leq 8.2 \text{ ms}^{-1}$) and group 3 (strong breeze: $8.2 < \text{observed} < 13.9 \text{ ms}^{-1}$) with the biases almost zero for all models. The results clearly demonstrate that BCanEn outperforms AnEn with improvements of RMSE and BIAS values for strong breeze/gale/storm (groups 3 and 4: $8.2 \text{ ms}^{-1} < \text{observed wind speed}$). Among the three BCanEn

FIG. 4. Density scatterplots of observed and modeled 10-m wind speed binned by 2 m s^{-1} intervals for (a) WRF, (b) ICLAMS, (c)–(e) AnEn, and (f)–(h) BCanEn models (146 storms).

implementations, the most skillful is BCanEn_{DUAL}, with reduced RMSEs for all seasons when compared to raw models (group 3: 20%–38% of WRF, 22%–25% of ICLAMS; group 4: 13%–27% of WRF, 15%–22% of ICLAMS).

Since high wind speed storms are the main focus of this study, the models' performance is also assessed for wind speed during the observed storm peak at each station (i.e., the maximum value of wind speed for each storm and location; Figs. 7, 8). BCanEn_{DUAL} is consistently performing best across

all seasons, with summer being the only exception, where performance is similar to AnEn_{WRF} and BCanEn_{WRF}. In the summer, the models exhibit lower forecast errors in terms of storm peaks when compared to errors in the other seasons because summer wind speeds are lower. About 98% of low wind speeds in groups 1 and 2 (Table 4) are observed during summer.

The AnEn skill depends on the raw model's ability to estimate the observations (Nagarajan et al. 2015). In this regard, AnEn and BCanEn take advantage of

RMSE

FIG. 5. Seasonal RMSE calculated by observed 10-m wind speeds classified in four groups as described in Table 4. The number of model–observation pairs is indicated for each group.

the WRF and ICLAMS independent estimates in searching for analogs. Furthermore, from the previous discussion, it is evident that BCAnEn using dual predictors is the best-performing method for the prediction of high wind speeds associated with extreme storms (Figs. 5–8). Thus, in the analysis that follows, the forecast skill is compared among WRF, ICLAMS, and BCAnEn_{DUAL}.

Errors calculated individually for each storm and season reveal the main differences between raw model outputs and the BCAnEn method (Fig. 9). WRF exhibits higher RMSE values than ICLAMS for most of the storms, with the exception of summer. Regardless of the different performances of the two raw models, BCAnEn_{DUAL} reduces their RMSE and BIAS over all seasons. BCAnEn_{DUAL} achieves RMSE improvements for all storms in the range of 9%–42% and 1%–29% when compared to WRF and ICLAMS, respectively. In addition, BCAnEn_{DUAL} reduces the RMSE values for Irene to 1.73 m s^{-1} (28 August 2011; 24% and 20% reduction for WRF and ICLAMS, respectively) and Sandy to 1.93 m s^{-1} (29 October 2012; 27% and 24% reduction for WRF and ICLAMS, respectively). On the contrary, AnEn_{DUAL} does not reduce the error for the two tropical storms (not shown in Fig. 9).

WRF has a negative bias for 88% of the storms, ranging from -1.55 to -0.01 m s^{-1} , with biased predictions being especially predominant in winter (Fig. 9). ICLAMS performs better than WRF in terms of bias (-0.52 m s^{-1} for WRF and -0.19 m s^{-1} for ICLAMS), but a noticeable positive BIAS for ICLAMS greater than 0.9 m s^{-1} is evident for several storms (e.g., 25 October 2005 and 1 and 17 December 2015; not shown in Fig. 9). BIAS is almost entirely removed for most of the storms with BCAnEn_{DUAL}, having an average BIAS value of 0.03 m s^{-1} . BCAnEn_{DUAL} also brings the median BIAS among all seasons closer to zero, compared to the raw models (spring: -0.06 ; summer: -0.01 ; fall: 0.16 ; and winter: 0.01 m s^{-1} ; Fig. 9). The results for the event-based analysis discussed in this section indicate that BCAnEn_{DUAL} is indeed able to significantly reduce the error of raw forecast for individual storms. The behavior of BCAnEn_{DUAL} in terms of errors for the two extreme cases of Tropical Storms Irene and Sandy is analyzed in detail in section 4c.

b. Seasonal analysis of temporal and spatial error

In this section, the temporal and spatial variation of the errors are discussed at each forecast lead time and station and analyzed separately for each season. Figure 10 shows the temporal variation of RMSE, BIAS,

BIAS

FIG. 6. As in Fig. 5, but for seasonal BIAS.

and R^2 for WRF (blue), ICLAMS (red), BCAnEn_{DUAL} (purple), and HyEn_{BC} (orange) (each error metric is calculated using all available model–observation pairs for the specified lead time). Temporal RMSE values feature similar patterns among the three predictions with peaks between the 20th and 30th lead times. This is because RMSE peaks correspond to maximum wind speed, where the magnitude of the variable and, thus, its error, is higher. The start and end times of the simulation are selected as such that they encapsulate the entire storm duration, having the peak of the storm approximately in the middle of the forecast timeline. Compared to WRF and ICLAMS, BCAnEn_{DUAL} and HyEn_{BC} show error reductions across all forecast lead times in the range of 9%–30% (spring), 8%–25% (summer), 6%–32% (fall), and 14%–36% (winter). Additionally, seasonal RMSE

peaks are consistently reduced by both AnEn applications, with percentage improvements in the range of 13%–32%, compared to raw models across the four seasons. Nonoverlapping bootstrapped intervals (95%) provide confidence that results are significantly different for all RMSE values for BCAnEn_{DUAL} and HyEn_{BC} (not shown).

BCAnEn_{DUAL} provides a nearly constant-in-time BIAS around zero, whereas WRF and ICLAMS BIAS values fluctuate over forecast lead times. WRF has negative BIAS for most forecast lead times in all seasons. During winter, WRF wind speeds show a stronger negative BIAS, compared to the other seasons, but nevertheless BCAnEn_{DUAL} removes the BIAS of the raw forecasts. The correlation of determination R^2 also indicates the improvements succeeded with BCAnEn_{DUAL}, compared to individual

TABLE 4. Wind speed classification. (Source: NOAA, Beaufort wind scale; <http://www.spc.noaa.gov/faq/tornado/beaufort.html>.)

Group	Condition (m s^{-1})	Classification	WMO classification
1	$0.0 < \text{observed} \leq 3.1$	Light breeze	Calm/light air/light breeze
2	$3.1 < \text{observed} \leq 8.2$	Moderate breeze	Gentle breeze/moderate breeze
3	$8.2 < \text{observed} \leq 13.9$	Strong breeze	Fresh breeze/strong breeze
4	$13.9 < \text{observed}$	Gale/storm	Near gale/gale/strong gale/storm

RMSE

FIG. 7. Seasonal RMSE calculated with 10-m wind speeds during the observed storm peaks (maximum wind speed).

raw models WRF and ICLAMS (Fig. 10). It is noteworthy that BCAnEn_{DUAL} and HyEn_{BC} exhibit similar skill in terms of R^2 , RMSE, and BIAS for all seasons.

AnEn forecast variability is compared to observed and raw model variability using the standard deviation of observed and predicted wind speed at a given forecast lead time (shown in Fig. S5). AnEn_{DUAL} forecast variability is lower than the observed variability across forecast lead times, but BCAnEn_{DUAL} gets closer to the observations and similar to ICLAMS. To see how that translates in time series of wind speed, we randomly selected four stations and plotted observed and modeled wind speeds for Hurricane Sandy (Fig. S6). This gives an indication of the temporal variability during a storm and how it is depicted by all models. It is clear that AnEn_{DUAL} underestimates wind speed maxima during these extreme events, but BCAnEn_{DUAL} brings them closer to the observed ones. Last, maximum values of observed and predicted wind speeds for each station and lead time are shown in Fig. S7. Consistent with the results shown in the manuscript (Figs. 7, 8), all models show underestimations of maximum wind speed when compared to the observations. However, AnEn_{DUAL} and BCAnEn_{DUAL} exhibit similar variability with the raw models and closer to the observed variability.

The spatial distribution of RMSE and BIAS analyzed for 80 stations in the northeast United States for

the four seasons is shown by geographical maps of the area of interest (Figs. 11, 12). Each colored circle represents the value of the statistical metric calculated with observations for each station and season using all available data pairs. Consistent with the results reported earlier, BCAnEn_{DUAL} is effective at reducing the RMSE, compared to raw models (third row in Fig. 11). BCAnEn_{DUAL} achieves average RMSE reduction (i.e., spatially averaged RMSE) in the range of 18%–30% and 15%–18% for WRF and ICLAMS, respectively, across the four seasons. In the summer period, the models exhibit lower RMSE values, explained by the fact that over that period, the metric is computed with a greater percentage of low wind speeds than the other seasons, and this is also shown in the event-based and temporal RMSE analyses (Figs. 7–10).

BCAnEn_{DUAL} removes the BIAS of raw models for most of the stations, indicated by the whiter colors in Fig. 12. WRF consistently exhibits negative bias values, similar to temporal BIAS results for 47 (spring), 34 (summer), 38 (fall), and 53 (winter) stations ($\text{BIAS} < -0.5 \text{ m s}^{-1}$). For BCAnEn_{DUAL}, 21%–79% stations have a BIAS in the range of -0.1 to 0.1 m s^{-1} over the four seasons (WRF: 1%–10% stations; ICLAMS: 5%–11% stations). Overall, BCAnEn_{DUAL} is shown to be an effective method to reduce RMSE and BIAS spatially and temporally.

FIG. 8. As in Fig. 7, but for seasonal BIAS.

c. Analysis of model performances for Tropical Storms Irene and Sandy

The bias-correction scheme applied to AnEn is successful in significantly reducing the error of the raw models, according to the analysis presented in the previous sections. Here, we analyze the AnEn performance for two tropical storms (Irene and Sandy) to investigate how much improvement can be obtained with AnEn_{DUAL}, BCanEn_{DUAL}, and HyEn_{BC}. Tropical Storms Irene and Sandy have resulted in significant impacts for the northeast United States (power network interruptions, critical infrastructure damages, and coastal area flooding), which is the main reason for focusing on improving wind speed prediction for these storms using AnEn.

BCanEn_{DUAL} is an effective method for reducing the peak RMSE values, compared to AnEn_{DUAL} (38% and 39% RMSE reduction for Irene and Sandy, respectively). The peak RMSE generally coincides with the storm peak (9–13-h forecast lead time for Irene and 20–28 h for Sandy), and the error reduction is statistically significant for 7-, 11-, and 13-h forecast lead times for Irene, and 18 and 20 h for Sandy (shown in Figs. 13a,b). The RMSE peaks produced by the 10 members of AnEn_{DUAL} are in the range of 3.7–5.0 and 5.3–6.8 m s⁻¹ for Irene and Sandy, which are higher

than the BCanEn_{DUAL} members (Irene: 2.7–3.6 m s⁻¹; Sandy: 3.9–4.7 m s⁻¹; not shown in Figs. 13a,b). The AnEn_{DUAL} members are biased (shown in Fig. S4), which elucidates why the ensemble mean is not improving the prediction for Irene and Sandy. This suggests that the bias-correction method is necessary to address AnEn_{DUAL}'s notable underestimation of rarely observed wind speed, as indicated by the temporal BIAS of the ensemble mean for the two tropical storms (Figs. 13c,d).

The biggest impact of BCanEn_{DUAL} is the significant RMSE reduction at the storm peak. The overall BIAS for Irene is not significantly different from the two raw models (Fig. 13c), while for Sandy, it is significantly different for the first 10 h (Fig. 13d). BCanEn_{DUAL} performs better than AnEn_{DUAL} not only for the temporal, but also for the spatial errors. Out of the 80 available stations, AnEn_{DUAL} decreases the RMSE for 33 and 38 stations for Irene and 31 and 36 stations for Sandy, compared to WRF and ICLAMS, whereas BCanEn_{DUAL} achieves improvements at double the number of stations (Irene: 64 for WRF and 69 for ICLAMS; Sandy: 69 for WRF and 65 for ICLAMS; Fig. 14).

BCanEn_{DUAL} shows positive BIAS values across some stations but also achieves greater bias reduction when

RMSE

BIAS

FIG. 9. RMSE and BIAS of WRF, ICLAMS, and BCAnEn_{DUAL} for storms in each season. The modeled box-and-whisker plots include 18 (spring), 43 (summer), 30 (fall), and 55 (winter) storms. RMSE and BIAS calculated with all data available for individual storms.

compared to WRF, ICLAMS, and AnEn_{DUAL} for Irene and Sandy. BCAnEn_{DUAL} has 49 (Irene) and 47 (Sandy) stations in the BIAS range of -0.5 to 0.5 m s^{-1} , compared to 32 (Irene) and 19 (Sandy) for WRF, 25 (Irene) and

14 (Sandy) for ICLAMS, and 36 (Irene) and 26 (Sandy) for AnEn_{DUAL}. This demonstrates that BCAnEn_{DUAL} effectively improved the bias for most stations when compared to the other models.

FIG. 10. Temporal variation of R^2 , RMSE, and BIAS for WRF (blue), ICLAMS (red), BCAnEn_{DUAL} (purple), and HyEn_{BC} (orange) for each season. The HyEn_{BC} is constructed by the best five analog ensemble members selected separately from BCAnEn_{WRF} and BCAnEn_{ICLAMS}. Statistical metrics use all available observed–modeled pairs at a given lead time.

Given the BCAnEn_{DUAL}'s computational cost for weight optimization using six predictors, there is interest in looking at the performance of HyEn, selecting five analogs from each NWP member, and constructing the 10 ensemble members to assess how different HyEn and BCAnEn_{DUAL} are for the two tropical

storms. For Irene and Sandy, two HyEn models constructed by the best five analog ensemble members for each model (from AnEn_{WRF} and AnEn_{ICLAMS} is referred to as HyEn; from BCAnEn_{WRF} and BCAnEn_{ICLAMS} is referred to as HyEn_{BC}) are evaluated against AnEn_{DUAL} and BCAnEn_{DUAL}.

RMSE

FIG. 11. Spatial distribution of seasonal RMSE for WRF, ICLAMS, and BCAnEn_{DUAL}. The topography is denoted with the gray-shaded area. Statistical metrics use all available data for each station.

Although HyEn_{BC} employs the analog ensemble members from two different BCAnEn models, HyEn (green) and AnEn_{DUAL} (orange), and HyEn_{BC} (black) and BCAnEn_{DUAL} (purple) models are almost identical in terms of their performance for RMSE and BIAS across all forecast hours (Fig. 13). The improvements of RMSE and BIAS are not achieved by HyEn in predicting Irene and Sandy because the members of AnEn_{WRF} and AnEn_{ICLAMS} are biased (see Fig. S4). However, HyEn_{BC} and BCAnEn_{DUAL} improve wind speed prediction of Irene and Sandy with similar RMSE and BIAS over forecast lead times (Fig. 13). For the spatial distribution of the error metrics analyzed for 80 stations, both AnEn_{DUAL} and HyEn models and both BCAnEn_{DUAL} and HyEn_{BC} models show similar spatial patterns with regard to RMSE and BIAS for each station (Fig. 14). The decrease in RMSE and BIAS for HyEn_{BC} model is significant and useful to improve the performance of WRF and ICLAMS. This

indicates that a hybrid AnEn method is effective at producing improved prediction of wind speed for two tropical storms when employing the members from two BCAnEn models.

5. Summary

This study presents a new implementation of the analog ensemble (AnEn) and a bias-corrected analog ensemble (BCAnEn) for improving wind speed prediction of extreme storms that have impacted the northeast United States during a 12-yr period (2005–16). The novelty of the presented approach is the use of two NWP models (WRF and ICLAMS) simultaneously to provide multiple predictors to the AnEn, as well as the implementation to extreme storms where high wind speeds have significant impacts on the infrastructure and the environment. Optimal analogs for AnEn and BCAnEn are searched using multiple

BIAS

FIG. 12. As in Fig. 11, but for spatial distribution of seasonal BIAS.

predictors from WRF, ICLAMS, and a combination of both models (DUAL), given optimized weights that are found by minimizing the root-mean-square error (RMSE).

A leave-one-out cross validation (LOOCV) is implemented to build the training dataset and to apply and evaluate AnEn and BCanEn for each storm. Both AnEn and BCanEn are most efficient for wind speed prediction improvements when using dual predictors from the two NWP models (six atmospheric variables/predictors; AnEn_{DUAL} and BCanEn_{DUAL}). The seasonal statistical metrics for four wind speed classification groups indicate that the inclusion of all six, rather than three, predictors from WRF or ICLAMS is beneficial for error reduction. For strong wind associated with rare events, AnEn does not reduce the largest errors, which instead is accomplished with BCanEn. For the gale/storm group, BCanEn reduces the seasonal RMSE of WRF and ICLAMS by 13%–27% and 15%–22%, respectively. BCanEn reduces the RMSE

also for individual storms by an amount ranging between 9% and 42% for WRF and 1% and 29% for ICLAMS.

Results from the spatial and temporal error analysis indicate that BCanEn consistently produces significantly lower RMSE than WRF and ICLAMS at all forecast lead times. Similar results are achieved by BCanEn also for individual stations, indicated by the spatially averaged RMSE reduction of raw models in the range of 18%–30% (WRF) and 15%–18% (ICLAMS) across all seasons.

Performances of AnEn and BCanEn (using six predictors) for Irene and Sandy differ substantially. The temporal and spatial analysis of RMSE and BIAS for BCanEn shows that the bias-correction scheme applied to AnEn offers benefits for the two tropical storms, which are characterized by winds significantly higher than the other storms included in this study. Consequently, the improvements obtained by BCanEn for Irene and Sandy are shown with average

FIG. 13. Temporal variation of (a),(b) RMSE and (c),(d) BIAS for WRF (blue), ICLAMS (red), AnEn (orange), BCAnEn_{DUAL} (purple), HyEn (green), and HyEn_{BC} (black) for Irene and Sandy. Statistical metrics use all available observed–modeled pairs at a given lead time.

RMSE reductions of around 22% for Irene and 26% for Sandy with respect to the raw models, and the most significant benefit is that the error is reduced for the peak wind speed associated with each storm. The hybrid AnEn approach (HyEn_{BC}) is also proven to have similar skill with BCAnEn, with the advantage of a lower computational burden since the optimization is done on fewer predictors and separately for each NWP.

Overall, AnEn is proven a successful technique to improve wind speed predictions for extreme storms in the northeast United States. The results in this study highlight that using dual-model predictors and a bias-correction scheme results in high-quality AnEn-based deterministic forecasts. A future implementation of this technique involves error correction of wind speed during real-time forecast of storms, given the advantage of using past storm forecasts and matching

FIG. 14. Spatial distribution of RMSE and BIAS of WRF, ICLAMS, AnEn_{DUAL} and BCanEn_{DUAL}, HyEn, and HyEn_{BC} for Irene and Sandy. Topography is denoted with the gray-shaded area. Statistical metrics use all available data for each station.

observations to build the analogs in the domain of interest.

Acknowledgments. The work was supported by Eversource Energy through a research grant awarded by the Eversource Energy Center at the University of Connecticut and by the Graduate Visitor program of the Advanced Study Program of National Center for Atmospheric Research.

REFERENCES

- Alessandrini, S., F. Davò, S. Sperati, M. Benini, and L. Delle Monache, 2014: Comparison of the economic impact of different wind power forecast systems for producers. *Adv. Sci. Res.*, **11**, 49–53, <https://doi.org/10.5194/asr-11-49-2014>.
- , L. Delle Monache, S. Sperati, and J. N. Nissen, 2015a: A novel application of an analog ensemble for short-term wind power forecasting. *Renewable Energy*, **76**, 768–781, <https://doi.org/10.1016/j.renene.2014.11.061>.
- , —, —, and G. Cervone, 2015b: An analog ensemble for short-term probabilistic solar power forecast. *Appl. Energy*, **157**, 95–110, <https://doi.org/10.1016/j.apenergy.2015.08.011>.
- , —, C. Rozoff, and W. Lewis, 2018: Probabilistic prediction of the tropical cyclone intensity with an analog ensemble. *Mon. Wea. Rev.*, **146**, 1723–1744, <https://doi.org/10.1175/MWR-D-17-0314.1>.
- Bergen, R. E., and R. P. Harnack, 1982: Long-range temperature prediction using a simple analog approach. *Mon. Wea. Rev.*, **110**, 1083–1099, [https://doi.org/10.1175/1520-0493\(1982\)110<1083:LRTPUA>2.0.CO;2](https://doi.org/10.1175/1520-0493(1982)110<1083:LRTPUA>2.0.CO;2).
- Carter, R. G., and R. E. Keislar, 2000: Emergency response transport forecasting using historical wind field pattern matching. *J. Appl. Meteor.*, **39**, 446–462, [https://doi.org/10.1175/1520-0450\(2000\)039<0446:ERTFUH>2.0.CO;2](https://doi.org/10.1175/1520-0450(2000)039<0446:ERTFUH>2.0.CO;2).
- Chou, M. D., and M. J. Suarez, 1994: An efficient thermal infrared radiation parameterization for use in general circulation models. NASA Tech. Memo. 104606, Vol. 3, 85 pp.
- Cotton, W. R., and Coauthors, 2003: RAMS 2001: Current status and future directions. *Meteor. Atmos. Phys.*, **82**, 5–29, <https://doi.org/10.1007/s00703-001-0584-9>.
- Davò, F., S. Alessandrini, S. Sperati, L. Delle Monache, D. Airolidi, and M. T. Vespucci, 2016: Post-processing techniques and principal component analysis for regional wind power and solar irradiance forecasting. *Sol. Energy*, **134**, 327–338, <https://doi.org/10.1016/j.solener.2016.04.049>.
- Delle Monache, L., 2017: Postprocessing of air quality predictions. *2017 Fall Meeting*, San Francisco, CA, Amer. Geophys. Union, Abstract A33L-01.
- , T. Nipen, Y. Liu, G. Roux, and R. Stull, 2011: Kalman filter and analog schemes to postprocess numerical weather predictions. *Mon. Wea. Rev.*, **139**, 3554–3570, <https://doi.org/10.1175/2011MWR3653.1>.
- , F. A. Eckel, D. L. Rife, B. Nagarajan, and K. Searight, 2013: Probabilistic weather prediction with an analog ensemble. *Mon. Wea. Rev.*, **141**, 3498–3516, <https://doi.org/10.1175/MWR-D-12-00281.1>.
- Djalalova, I., L. Delle Monache, and J. Wilczak, 2015: PM_{2.5} analog forecast and Kalman filter post-processing for the Community Multiscale Air Quality (CMAQ) model. *Atmos. Environ.*, **119**, 431–442, <https://doi.org/10.1016/j.atmosenv.2015.05.057>.
- Drosowsky, W., 1994: Analog (nonlinear) forecasts of the Southern Oscillation index time series. *Wea. Forecasting*, **9**, 78–84, [https://doi.org/10.1175/1520-0434\(1994\)009<0078:AFOTSO>2.0.CO;2](https://doi.org/10.1175/1520-0434(1994)009<0078:AFOTSO>2.0.CO;2).
- Eckel, F. A., and L. Delle Monache, 2016: A hybrid NWP–analog ensemble. *Mon. Wea. Rev.*, **144**, 897–911, <https://doi.org/10.1175/MWR-D-15-0096.1>.
- Fountoukis, C., and A. Nenes, 2005: Continued development of a cloud droplet formation parameterization for global climate models. *J. Geophys. Res.*, **110**, D11212, <https://doi.org/10.1029/2004JD005591>.
- Gao, L., H. Ren, J. Li, and J. Chou, 2006: Analogue correction method of errors and its application to numerical weather prediction. *Chin. Phys.*, **15**, 882, <https://doi.org/10.1088/1009-1963/15/4/038>.
- Grell, G. A., and D. Dévényi, 2002: A generalized approach to parameterizing convection combining ensemble and data assimilation techniques. *Geophys. Res. Lett.*, **29**, 1693, <https://doi.org/10.1029/2002GL015311>.
- Hamill, T. M., and J. S. Whitaker, 2006: Probabilistic quantitative precipitation forecasts based on reforecast analogs: Theory and application. *Mon. Wea. Rev.*, **134**, 3209–3229, <https://doi.org/10.1175/MWR3237.1>.
- He, J., D. W. Wanik, B. M. Hartman, E. N. Anagnostou, M. Astitha, and M. Frediani, 2017: Nonparametric tree-based predictive modeling of storm outages on an electric distribution network. *Risk Anal.*, **37**, 441–458, <https://doi.org/10.1111/risa.12652>.
- Hong, S.-Y., Y. Noh, and J. Dudhia, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev.*, **134**, 2318–2341, <https://doi.org/10.1175/MWR3199.1>.
- Jammalamadaka, S. R., and A. Sengupta, 2001: *Topics in Circular Statistics*. Series on Multivariate Analysis, Vol. 5, World Scientific, 322 pp.
- Junk, C., L. Delle Monache, S. Alessandrini, G. Cervone, and L. von Bremen, 2015a: Predictor-weighting strategies for probabilistic wind power forecasting with an analog ensemble. *Meteor. Z.*, **24**, 361–379, <https://doi.org/10.1127/metz/2015/0659>.
- , —, and —, 2015b: Analog-based ensemble model output statistics. *Mon. Wea. Rev.*, **143**, 2909–2917, <https://doi.org/10.1175/MWR-D-15-0095.1>.
- Keller, J., L. Delle Monache, and S. Alessandrini, 2017: Statistical downscaling of a high-resolution precipitation reanalysis using the analog ensemble method. *J. Appl. Meteor. Climatol.*, **56**, 2081–2095, <https://doi.org/10.1175/JAMC-D-16-0380.1>.
- Klausner, Z., H. Kaplan, and E. Fattal, 2009: The similar days method for predicting near surface wind vectors. *Meteor. Appl.*, **16**, 569–579, <https://doi.org/10.1002/met.158>.
- Livezey, R. E., and A. G. Barnston, 1988: An operational multifield analog/antianalog prediction system for United States seasonal temperatures. Part I: System design and winter experiments. *J. Geophys. Res.*, **93**, 10 953–10 974, <https://doi.org/10.1029/JD093iD09p10953>.
- Mahoney, W. P., and Coauthors, 2012: A wind power forecasting system to optimize grid integration. *IEEE Trans. Sustainable Energy*, **3**, 670–682, <https://doi.org/10.1109/TSTE.2012.2201758>.
- Mellor, G. L., and T. Yamada, 1982: Development of a turbulence closure model for geophysical fluid problems. *Rev. Geophys.*, **20**, 851–875, <https://doi.org/10.1029/RG020i004p00851>.
- Meyers, M. P., R. L. Walko, J. Y. Harrington, and W. R. Cotton, 1997: New RAMS cloud microphysics parameterization. Part II: The two-moment scheme. *Atmos. Res.*, **45**, 3–39, [https://doi.org/10.1016/S0169-8095\(97\)00018-5](https://doi.org/10.1016/S0169-8095(97)00018-5).
- Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the

- longwave. *J. Geophys. Res.*, **102**, 16 663–16 682, <https://doi.org/10.1029/97JD00237>.
- Nagarajan, B., L. Delle Monache, J. P. Hacker, D. L. Rife, K. Searight, J. C. Knivel, and T. N. Nipen, 2015: An evaluation of analog-based postprocessing methods across several variables and forecast models. *Wea. Forecasting*, **30**, 1623–1643, <https://doi.org/10.1175/WAF-D-14-00081.1>.
- NCEP/NWS/NOAA/DOC, 2000: NCEP FNL Operational Model Global Tropospheric Analyses, continuing from July 1999. NCAR Research Data Archive Computational and Information Systems Lab, accessed September 2016, <https://doi.org/10.5065/D6M043C6>.
- , 2007: NCEP Global Forecast System (GFS) analyses and forecasts. NCAR Research Data Archive Computational and Information Systems Lab, accessed September 2016, <https://doi.org/10.5065/D65Q4TSG>.
- Nenes, A., and J. H. Seinfeld, 2003: Parameterization of cloud droplet formation in global climate models. *J. Geophys. Res.*, **108**, 4415, <https://doi.org/10.1029/2002JD002911>.
- Panziera, L., U. Germann, M. Gabella, and P. V. Mandapaka, 2011: NORA—Nowcasting of orographic rainfall by means of analogues. *Quart. J. Roy. Meteor. Soc.*, **137**, 2106–2123, <https://doi.org/10.1002/qj.878>.
- Radinović, D., 1975: An analogue method for weather forecasting using the 500/1000 mb relative topography. *Mon. Wea. Rev.*, **103**, 639–649, [https://doi.org/10.1175/1520-0493\(1975\)103<0639:AAMFWF>2.0.CO;2](https://doi.org/10.1175/1520-0493(1975)103<0639:AAMFWF>2.0.CO;2).
- Ren, H., and J. Chou, 2007: Strategy and methodology of dynamical analogue prediction. *Sci. Chin.*, **50D**, 1589–1599, <https://doi.org/10.1007/s11430-007-0109-6>.
- Skamarock, W. C., and Coauthors, 2008: A description of the Advanced Research WRF version 3. NCAR Tech. Note NCAR/TN-475+STR, 113 pp., http://www.mmm.ucar.edu/wrf/users/docs/arw_v3.pdf.
- Solomos, S., G. Kallos, J. Kushta, M. Astitha, C. Tremback, A. Nenes, and Z. Levin, 2011: An integrated modeling study on the effects of mineral dust and sea salt particles on clouds and precipitation. *Atmos. Chem. Phys.*, **11**, 873–892, <https://doi.org/10.5194/acp-11-873-2011>.
- Sperati, S., S. Alessandrini, and L. Delle Monache, 2017: Gridded probabilistic weather forecasts with an analog ensemble. *Quart. J. Roy. Meteor. Soc.*, **143**, 2874–2885, <https://doi.org/10.1002/qj.3137>.
- Tewari, M., and Coauthors, 2004: Implementation and verification of the unified Noah land-surface model in the WRF Model. *20th Conf. on Weather Analysis and Forecasting/16th Conf. on Numerical Weather Prediction*, Seattle, WA, Amer. Meteor. Soc., 14.2a, https://ams.confex.com/ams/84Annual/techprogram/paper_69061.htm.
- Thompson, G., P. R. Field, R. M. Rasmussen, and W. D. Hall, 2008: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. *Mon. Wea. Rev.*, **136**, 5095–5115, <https://doi.org/10.1175/2008MWR2387.1>.
- Toth, Z., 1989: Long-range weather forecasting using an analog approach. *J. Climate*, **2**, 594–607, [https://doi.org/10.1175/1520-0442\(1989\)002<0594:LRWFUA>2.0.CO;2](https://doi.org/10.1175/1520-0442(1989)002<0594:LRWFUA>2.0.CO;2).
- van den Dool, H. M., 1989: A new look at weather forecasting through analogues. *Mon. Wea. Rev.*, **117**, 2230–2247, [https://doi.org/10.1175/1520-0493\(1989\)117<2230:ANLAWF>2.0.CO;2](https://doi.org/10.1175/1520-0493(1989)117<2230:ANLAWF>2.0.CO;2).
- , 1994: Searching for analogues, how long must we wait? *Tellus*, **46A**, 314–324, <https://doi.org/10.3402/tellusa.v46i3.15481>.
- , 2007: *Empirical Methods in Short-Term Climate Prediction*. Oxford University Press, 240 pp.
- Vanvyve, E., L. Delle Monache, A. J. Monaghan, and J. O. Pinto, 2015: Wind resource estimates with an analog ensemble approach. *Renewable Energy*, **74**, 761–773, <https://doi.org/10.1016/j.renene.2014.08.060>.
- Walko, R. L., W. Cotton, M. Meyers, and J. Harrington, 1995: New RAMS cloud microphysics parameterization. Part I: The single-moment scheme. *Atmos. Res.*, **38**, 29–62, [https://doi.org/10.1016/0169-8095\(94\)00087-T](https://doi.org/10.1016/0169-8095(94)00087-T).
- , and Coauthors, 2000: Coupled atmosphere–biophysics–hydrology models for environmental modeling. *J. Appl. Meteor.*, **39**, 931–944, [https://doi.org/10.1175/1520-0450\(2000\)039<0931:CABHMF>2.0.CO;2](https://doi.org/10.1175/1520-0450(2000)039<0931:CABHMF>2.0.CO;2).
- Wanik, D. W., E. N. Anagnostou, B. M. Hartman, M. E. Frediani, and M. Astitha, 2015: Storm outage modeling for an electric distribution network in northeastern USA. *Nat. Hazards*, **79**, 1359–1384, <https://doi.org/10.1007/s11069-015-1908-2>.
- , J. R. Parent, E. N. Anagnostou, and B. M. Hartman, 2017: Using vegetation management and LiDAR-derived tree height data to improve outage predictions for electric utilities. *Electr. Power Syst. Res.*, **146**, 236–245, <https://doi.org/10.1016/j.epsr.2017.01.039>.
- Xavier, P. K., and B. N. Goswami, 2007: An analog method for real-time forecasting of summer monsoon subseasonal variability. *Mon. Wea. Rev.*, **135**, 4149–4160, <https://doi.org/10.1175/2007MWR1854.1>.
- Yang, J., M. Astitha, E. Anagnostou, and B. Hartman, 2017: Using a Bayesian regression approach on dual-model windstorm simulations to improve wind speed prediction. *J. Appl. Meteor. Climatol.*, **56**, 1155–1174, <https://doi.org/10.1175/JAMC-D-16-0206.1>.
- Zhang, J., C. Draxl, T. Hopson, L. Delle Monache, E. Vanvyve, and B. M. Hodge, 2015: Comparison of numerical weather prediction based deterministic and probabilistic wind resource assessment methods. *Appl. Energy*, **156**, 528–541, <https://doi.org/10.1016/j.apenergy.2015.07.059>.