10.3 PROBABILISTIC FORECASTS OF LOCATION AND TIMING OF CONVECTION DURING THE 2012 NOAA HAZARDOUS WEATHER TESTBED SPRING FORECASTING EXPERIMENT

Stuart D. Miller, Jr.*[3,4,6], John S. Kain[1], Patrick T. Marsh[1,3,4], Adam J. Clark[1,3], Michael C. Coniglio[1], Valliappa Lakshmanan[1,3], James Correia[2,3], David Imy[2], Scott R. Dembek[3], Israel L. Jirak[2], Steven J. Weiss[2], Andrew R. Dean[2], Christopher J. Melick[2,3], Ryan A. Sobash[1,4], Ming Xue[4,5], Fanyou Kong[4,5], and Kevin W. Thomas[5]

[1]NOAA/National Severe Storms Laboratory, Norman, OK

[2]NOAA/NWS/NCEP/Storm Prediction Center, Norman, OK

[3]Cooperative Institute for Mesoscale Meteorological Studies, Norman, OK

[4]University of Oklahoma School of Meteorology, Norman, OK

[5]Center for Analysis and Prediction of Storms, Norman, OK

[6]USAF Institute of Technology Civilian Institution Programs, Wright-Patterson AFB, OH

1. INTRODUCTION

Forecasting of convection initiation (CI) was examined in the 2012 National Oceanic and Atmospheric Administration (NOAA) Hazardous Weather Testbed (HWT) Spring Forecasting Experiment (SFE) (e.g., Kain et al., 2012b). CI is of particular interest since hazards resulting from convection tend to disrupt human activity. This is especially true in general for aviation, or if the convection is severe and threatens life and property. It is therefore logical that the ability to accurately predict when and where CI will occur is critical for issuing the best forecasts possible. In order for this to be done, definitions of convection activity (CA), and CI are needed. Once these are determined, methods can be used to objectively detect CA and CI in gridded model and observational datasets. Using model guidance, human forecasts for CA and CI can be made, and model performance can be compared to human forecast skill.

Defining CA and CI are not trivial tasks since there has not been a universally-established definition of CI to date. Even if one existed, it would potentially be challenging to represent in the context of numerical models, since it is not straightforward what proxies would need to be used. A few possibilities are reflectivity, lightning, and/or vertical velocity, but ultimately a parameter needs to be chosen so that verification with observations can be performed. Another challenge lies in discerning true CI from ongoing convection, and deciding whether to focus on isolated convection, large convective events, or both. Finally, performing verification is difficult because it must be done over a region free of ongoing and transient convection before CI takes place.

During SFE2012, model guidance was provided by the University of Oklahoma Center for Analysis and Prediction of Storms (CAPS) convection-allowing, 4-km ensemble (see http://hwt.nssl.noaa.gov/Spring_2012/OPS_plan_draft.pdf for a complete overview of SFE2012 including detailed information on the CAPS ensemble). The CAPS ensemble configuration varies year to year, but it consisted of two primary sets of members in 2012 – core and PBL (planetary boundary layer). The core set was considered to be the most diverse, as it was comprised of Advanced Research Weather Research and Forecasting (WRF-ARW) members with varying physics and initial and lateral boundary conditions. The PBL members utilized the same initial and lateral boundary conditions,

**Corresponding author address:* Stuart Miller, 120 David L. Boren Blvd, Suite 2200, Norman, OK, 73072
E-mail: stuart.miller@ou.edu

but varied in the PBL scheme that was used. This work will focus on the core members.

Section 2 of this paper describes the techniques used in representing CA and CI in the CAPS ensemble and Section 3 gives an overview of CA and CI forecasting activities during SFE2012. Section 4 explains the verification techniques used on the CA and CI human and ensemble forecasts, Section 5 presents the results of this verification, and Section 6 offers conclusions drawn from those results. Finally, Section 7 lists references used and Section 8 includes all figures.

2. MODEL CA and CI

During SFE2012, CA was defined as the presence of (simulated or observed) reflectivity ≥35 dBZ at the minus 10ºC level (MTR35). For an explanation of this choice, see Kain et al. 2012a.

Model CA was provided at 5-min resolution on the 4-km CAPS grid for each member. In order to produce ensemble probability of CA within 20 km of a point, each grid point that was within 20 km of a CA point at any sample time within a specified time window (e.g. 1 h or 4 h) was assigned a value of 1. All other points were assigned a value of zero. Next, various levels of Gaussian smoothing (e.g. σ=5 grid points, σ=10 grid points, and σ=20 grid points) were applied to the binary fields to effectively generate probabilities as in Sobash et al. (2011). Finally, the mean of the CA probability fields from all members at each grid point was computed to produce a single ensemble probability field.

In addition to CA probabilities, CI guidance was generated from the raw 4-km, 5-min CA field. CA objects were identified and tracked through each 36-h ensemble forecast using the same tracking algorithm employed in Clark et al. (2012). The earliest instance of each unique CA object was labeled as the CI time and grid-point location. CI probabilities were also computed, in the same way as CA probabilities.

3. 2012 SPRING FORECASTING EXPERIMENT

SFE2012 was conducted between 7 May and 8 June 2012 in the NOAA HWT in Norman, OK. The CI desk was comprised of approximately 3-4 invited forecasters, researchers, and academics in addition to SFE facilitators. CI forecasting activities for the entirety of SFE2012 were led by an experienced Storm Prediction Center (SPC) forecaster. There were two primary components to CI forecasting during SFE2012 – probability of CA within 20 km of a point and first CI timing and location, both within a specified region. Each morning, a domain was chosen that was relatively free of convection, but in which convection was expected to initiate and spread throughout the day. This was known as the convection forecast domain, or CFD. Then, a smaller domain within the CFD that was completely free of convection was specified to determine the timing and location of the CI forecast. This second domain was known as the CI forecast domain, or CIFD.

Forecasts for probability of convection within 20 km of a point were made inside of the CFD (Fig. 1), valid for the following consecutive time periods: 1600-2000 UTC and 2000-0000 UTC the same day, and 0000-0400 UTC the next day. These probabilities were issued in the same bins as the SPC's Thunderstorm Outlooks (0-9%, 10-39%, 40-69%, and ≥70%). Probabilistic CA guidance from the CAPS ensemble and calibrated thunder (Bright et al. 2005) from the Short Range Ensemble Forecast (SREF) system (Du et al. 2004) were used, along with operational model guidance and routine observations, to make these forecasts.

Timing (within a 1-h period) and location of CI was forecast using a suite of products from the CAPS ensemble core members. This suite included a cumulative distribution function (CDF) of timing of all CA points (Fig. 2) and a probability distribution function (PDF) of first CI point timing (Fig. 3) from each member. These CDFs and PDFs were valid within the CIFD. When a forecast time for first CI inside the CIFD was agreed upon (e.g., if convection was expected between 1900 and 2000 UTC, the forecast CI time was set at 1930 UTC), a forecast PDF was drawn to convey timing uncertainty. The peak of the PDF was centered at the forecast CI time, and its width, area, and shape represented the overall uncertainty in the forecast for that day.

4. VERIFICATION TECHNIQUES

In order to ensure the fairest possible evaluation of forecast skill for CI timing (minimizing the impact of "spurious" nearby convection), a CI verification domain (CIVD) was chosen each morning, for the forecast valid the day before. This was typically a smaller domain inside of the CIFD that remained free of observed and simulated CA prior to the CI event that was forecasted, but included the location of human-forecasted CI. The CIVD excluded ongoing and/or remnant overnight convection as well as any transient convection irrelevant to the afternoon event of interest. Judicious selection of the CIVD minimized ambiguity in assessments of model and human forecasts in terms of timing error and reliability.

The observed reflectivity at -10ºC in its native temporal resolution on the National Mosaic & Multi-Sensor QPE (NMQ) (Zhang et al. 2011) grid was used, with the -10ºC height determined using the Rapid Update Cycle (RUC) model. Next, this reflectivity field was interpolated to the CAPS grid via a budget scheme where grid points with missing values were handled appropriately. Verification was then conducted in a manner consistent with a definition of CA within 20 km of a grid point.

The reliability diagram was chosen as the verification metric for the spatial probabilities of CA as well as the CI temporal probabilities for both ensemble and human forecasts. Additionally, histograms were used for qualitative assessment of the ensemble timing error of first CI.

5. RESULTS

Reliability for model and human probability of CA is displayed in Fig. 4. This is based on the daily 4-h probability forecasts within the CFDs using the same probability bins as the forecasts. The human forecasts appear to outperform the ensemble guidance slightly, but given the relatively small sample size there likely is not a significant difference between them.

The histogram of the timing error for forecasts of first CI from all of the CAPS ensemble core members is depicted in Fig. 5. This is valid for ±4 h of observed CI, during the period of SFE2012 inside of the CIVD. The distribution is relatively wide; however, a majority of the members do initiate convection within ±2 h of observed CI. There did seem to be an apparent decrease in performance using this metric when compared to similar results from SFE2011 (Fig. 6). The distribution of Fig. 6 is a bit narrower, with a majority of members initiating convection within ±1.5 h of observed CI, and the tails are relatively short compared to the SFE2012 plot (Fig. 5). There were more points on the SFE2011 (Fig. 6) histogram simply because there were more core members in the CAPS ensemble that year.

The most promising result from SFE2012 was reliability of human vs. ensemble forecasts of first CI timing (Fig. 7). This was valid over the same days as the ensemble CI timing error histograms inside of the CIVD (Figs. 5 and 6), with 20% probability bins. Essentially, this can be thought of as reliability based on all ensemble and human forecast PDFs of first CI timing (within ±4 h of observed CI)[1]. This particular metric assumes that CI actually occurred during the day within the CIVD, and there was only one day during SFE2012 on which this did not happen. There was minimal difference between human and ensemble reliability through 40% (both being nearly reliable), but humans added considerable skill to forecasts of CI timing in the 40-60% bin. Neither the ensemble nor human forecasts included an hour where the forecast probability of convection was >60%.

6. CONCLUSION

Though substantial strides have been made in modeling and forecasting CI and CA in the HWT during the past two SFEs, there is still significant work to be done. Subjective assessments suggest that convection-allowing models (CAMs) exhibit appreciable skill in predicting CA and CI, and are capable of providing valuable guidance to forecasters. However, problems still exist in determining the best model proxies to use, probability computation methods, and verification techniques, among others. In essence, subjective assessments suggest that CAMs predict the initiation of significant convective events quite well, but methods to objectively validate this assessment and provide useful CI guidance to forecasters are still lacking.

In this experimental setting, there were limitations on human skill at forecasting probability of CA over a given area, as can be seen in Fig. 4. However, it has been found that operational human forecasts of probability of convection (SPC's Thunderstorm Outlooks) are reliable (A.R. Dean, SPC, 2012, personal communication). Even though an experienced SPC forecaster led the CI desk during SFE2012, that individual allowed for substantial input from participants who tended to be much less experienced at operational forecasting. Additionally, forecasters themselves tend to become "calibrated" to new guidance after seeing verification of forecasts they based on that guidance (D. Imy, SPC, 2012, personal

[1] Computation of the reliability for the forecast PDFs follows the same concept as that for spatial probabilities. Since CI timing was forecast to the nearest hour and reliability was valid for ±4 h of observed CI , there were 9 points evaluated for each forecast – the hour of observed CI and ±4 h of it. For example, consider the cases where probability of CI in an hour was forecast to be 10%. A reliable forecast in this particular probability bin would mean that during SFE2012, observed CI occurred in 10% of the hours where 10%probability of CI was forecast.

communication). The experimental CAPS CA and CI guidance was relied upon in addition to operational and/or calibrated guidance during SFE2012, and it seems this had a detrimental effect on the reliability of human CA forecasts.

It is very promising, though, that human forecasters did add considerable skill to model forecasts of CI timing (Fig. 7). This paper should serve as a starting point for further work in CI forecasting, which needs to address defining CI in a model framework, improving forecast/verification techniques for simulated CI products, and improving model estimates of reflectivity.

7. REFERENCES

Bright, D.R., M.S. Wandishin, R.E. Jewell, and S.J. Weiss, 2005: A physically based parameter for lightning prediction and its calibration in ensemble forecasts. Preprints, Conf. on Meteor. Applications of Lightning Data, San Diego CA. [2339K PDF]

Clark, A.J., J.S. Kain, P.T. Marsh, J. Correia, M. Xue, F. Kong, 2012: Forecasting tornado pathlengths using a three-dimensional object identification algorithm applied to convection-allowing forecasts. Wea. Forecasting, **27**, 1090–1113.

Du, J., J. McQueen, G. DiMego, T. Black, H. Juang, E. Rogers, B. Ferrier, B. Zhou, Z. Toth and M.S. Tracton, 2004: The NOAA/NWS/NCEP shortrange ensemble forecast (SREF) system: evaluation of an initial condition vs. multi-model physics ensemble approach. Preprints, 16th Conference on Numerical Weather Prediction, Seattle, Washington, Amer. Meteor. Soc., CD-ROM (21.3)

Kain, J.S., M.C. Coniglio, J. Correia, A.J. Clark, P.T. Marsh, C.L. Ziegler, V. Lakshmanan, S.D. Miller, S.R. Dembek, S.J. Weiss, F. Kong, M. Xue, R.A. Sobash, A.R. Dean, I.L. Jirak, and C.J. Melick, 2012a: A feasibility study for probabilistic convection initiation forecasts based on explicit numerical guidance. Conditionally accepted for publication in the Bull. Amer. Meteor. Soc.

_____, I.L. Jirak, S.J. Weiss, A.J. Clark, M.C. Coniglio, J. Correia, A.R. Dean, P.T. Marsh, C.J. Melick, S.D. Miller, R.A. Sobash, M. Xue, F. Kong, K.W. Thomas, V. Lakshmanan, D.A. Imy, and S.R. Dembek, 2012b: An overview of the 2012 NOAA Hazardous Weather Testbed Spring Forecasting Experiment. Preprints, 26th Conference on Severe Local Storms, Nashville, TN, Amer. Meteor. Soc., https://ams.confex.com/ams/26SLS/webprogram/Paper211107.html (10.1)

Sobash, R A., J. S. Kain, D. R. Bright, A. R. Dean, M. C. Coniglio, and S. J. Weiss, 2011: Probabilistic forecast guidance for severe thunderstorms based on the identification of extreme phenomena in convection-allowing model forecasts. Wea. Forecasting, **26**, 714-728. doi: http://dx.doi.org/10.1175/WAF-D-10-05046.1

Zhang, J., and Coauthors, 2011: National mosaic and multi-sensor QPE (NMQ) system: description, results, and future plans. Bull. Amer. Meteor. Soc., 92, 1321–1338.

8. FIGURES

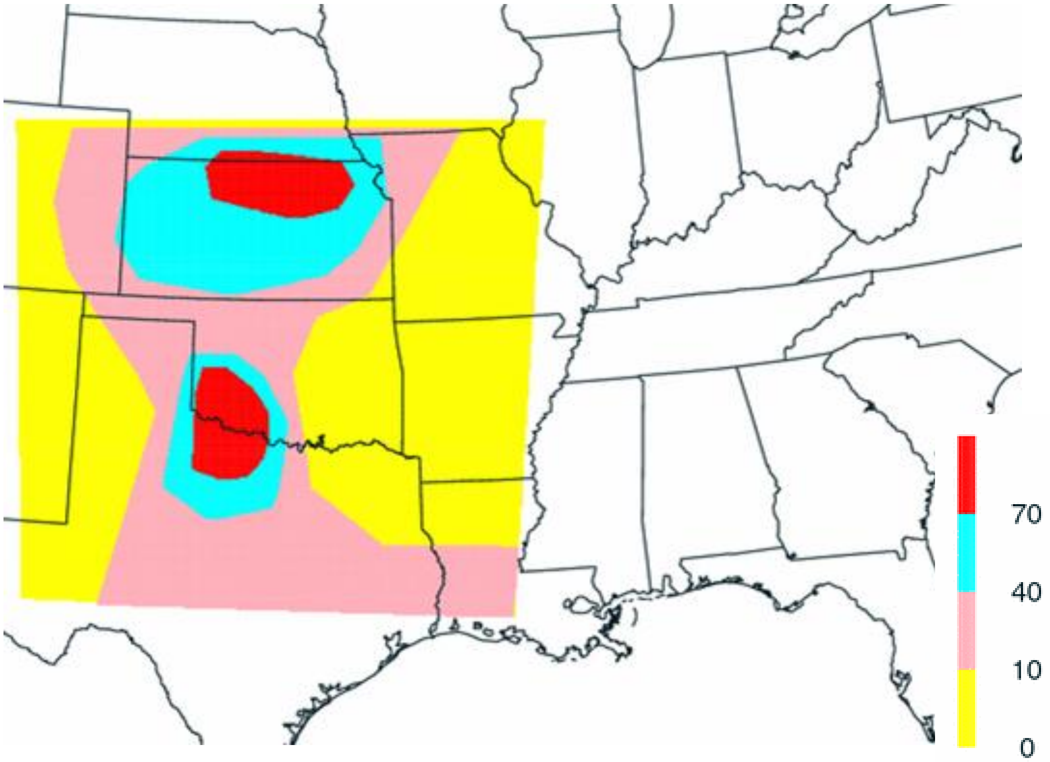

FIG. 1. Example forecast for probability of CA within 20 km of a point issued during SFE2012. The unit of the color bar is percent.

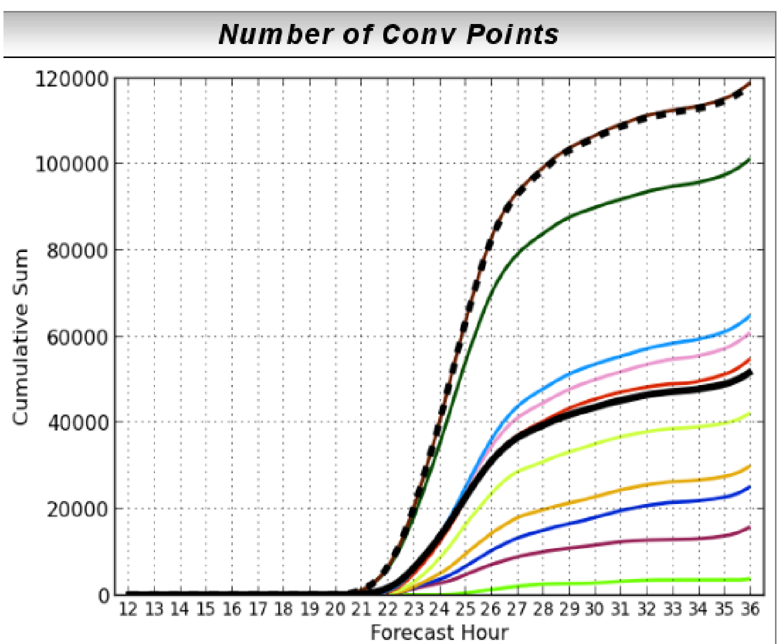

FIG. 2. Example of a CDF of number of CA points inside the CIFD for a day during SFE2012 for the CAPS ensemble core members. Each color represents a different member, the solid black curve is the ensemble mean, and the dashed line representing the observed CA points.

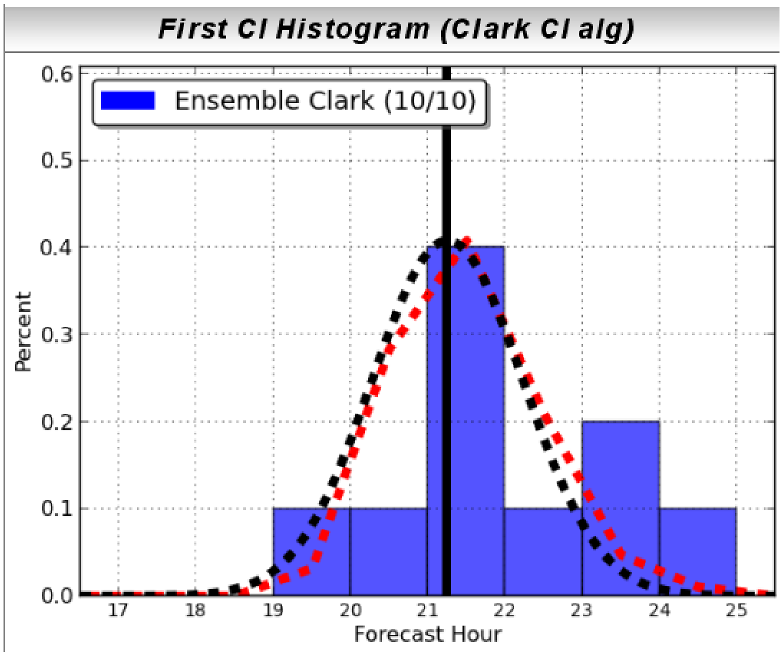

FIG. 3. Example of a PDF of first CI point inside the CIFD for a day during SFE2012 for the CAPS ensemble core members. The red dashed line represents the human forecast PDF of first CI time, the vertical black solid line is observed CI time, and the black dashed line is a Gaussian PDF centered at the observed CI time with the same peak and integrated area as the human PDF.

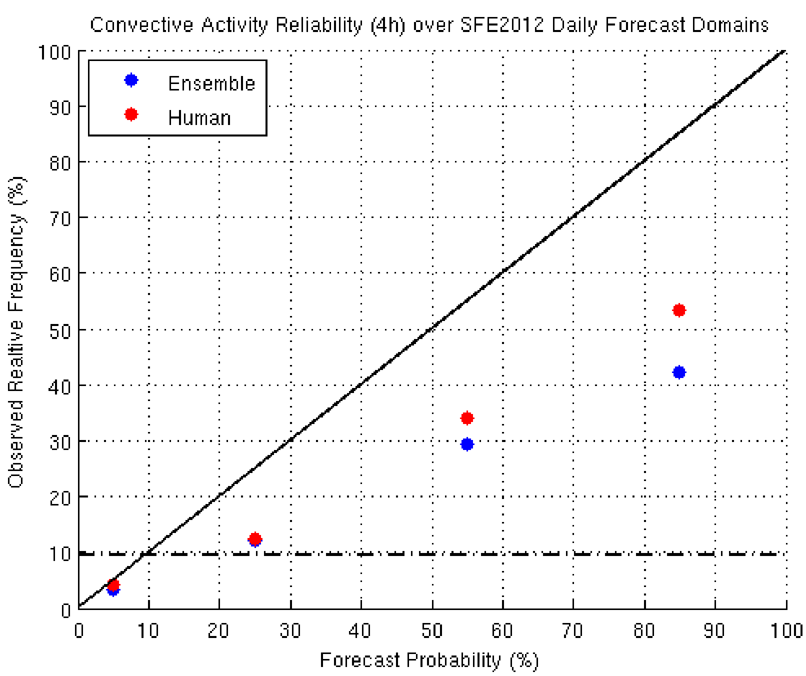

FIG. 4. Reliability of probability of CA within 20 km of a point for all 4-h human outlooks issued during SFE2012 and corresponding ensemble guidance (smoothed at σ=20 grid points) inside the CFD. The diagonal one-one solid line represents perfect reliability, and the horizontal dashed-stippled line represents climatology. The probabilities bins are the same as the probability thresholds used in the SPC's Thunderstorm Outlooks.

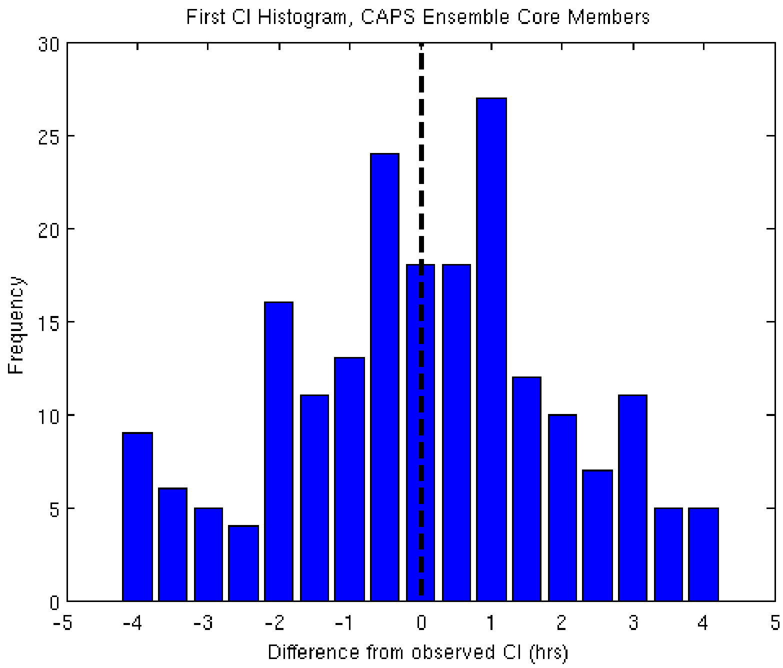

FIG. 5. Histogram of timing error of CI ±4 h of observed CI for all core members of the CAPS ensemble during SFE2012.

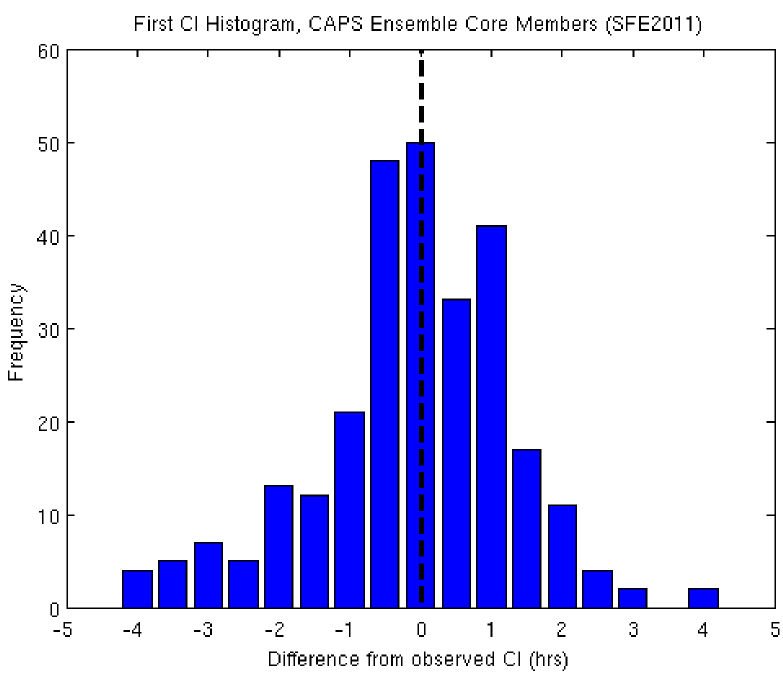

FIG. 6. Same as Fig. 5, except valid for SFE2011.

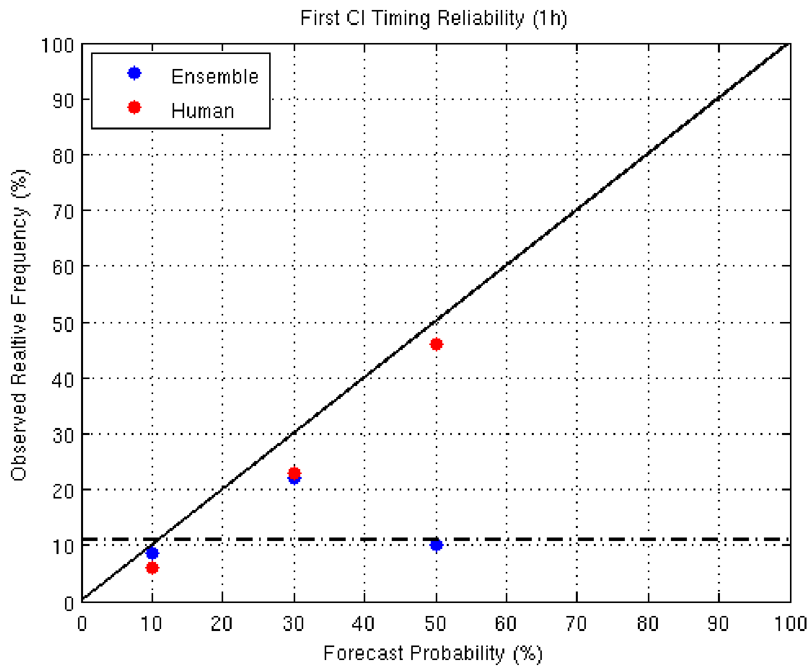

FIG. 7. Reliability of CI timing to the nearest hour for human and ensemble forecast CI timing PDFs inside the CIFD during SFE2012. The diagonal one-one solid line represents perfect reliability, and the horizontal dashed-stippled line represents climatology. The probabilities are binned at 20% intervals.