Quality Assessment of the National Convective Weather Forecast Product

Quality Assessment Group
Barbara G. Brown[1] and Jennifer L. Mahoney[2]

21 April 2000

1. Introduction

The National Convective Weather Forecast Product (NCWFP) is an automated system to predict the locations of convective storm areas out to two hours (Mueller et al. 1999). This system was developed by the Convective Weather Product Development Team (CWPDT). This document provides an assessment of the forecasting capability of the NCWFP, including quantitative verification of one- and two-hour forecasts provided by this product. The verification statistics are compared to statistics for other relevant forecast products, including the convective SIGMETs (C-SIGMETs), which are produced operationally by the Aviation Weather Center (AWC), and an earlier version of the NCWFP (NCWFP-98). This report was prepared in response to the Quality Assessment Plan for the NCWFP, as a step in the TEGO (path-to-operations) process for the FAA test-bed.

The NCWFP was evaluated in real time during the summer of 1999 (1 June through 31 August) by the Real-Time Verification System (RTVS) of the NOAA Forecast Systems Laboratory (Mahoney et al. 1997), as part of a convective forecast intercomparison exercise. The other products of interest here (e.g., C-SIGMETs) also were included in this evaluation. Verification results for the NCWFP and the other products are summarized in this report. Other results from this intercomparison are presented in a separate report (Mahoney et al. 2000).

The evaluation reported here is limited to the main convective season (i.e., the summer months, June through August) since this is the primary period when convection is a concern and convective forecasts can be verified unambiguously. Several categories and stratifications are considered, including forecast length (one and two hours), observation type, and observation filtering. In addition, day-to-day variability in the statistics, as well as statistics for some individual days, are considered.

The report is organized as follows. The forecasts that were included in the intercomparison are described in Section 2, along with a description of the data that were available for verification of the forecasts; mechanics of the analyses are considered in Section 3, and verification methods and statistics are described in Section 4. Section 5

[1] Resarch Applications Program, National Center for Atmospheric Research, Boulder CO
[2] Forecast Systems Laboratory, National Oceanic and Atmospheric Administration, Boulder CO

contains the results of the analyses, and Section 6 includes some conclusions. Some additional results are presented in Appendix A.

2. Data

2.1 Algorithms/forecasts

The NCWFP and the other forecast products considered in this evaluation are briefly described here.

1999 National Convective Weather Forecast Product (NCWFP): The 1999 NCWFP was developed by the CWPDT (Mueller et al. 1999) with funds provided by the FAA's Aviation Weather Research Program. Every 5 minutes the NCWFP generates 1- and 2-h convective extrapolation forecasts that are valid at a specific time. However, only forecasts produced on the hour are evaluated here. Enhancements to the NCWFP that were made in the transition from the 1998 version to the 1999 version include modifications to the extrapolation algorithm, inclusion of a stratiform/convection partition, smoothing of the forecast polygons, and production of 2-h forecasts (Convective Weather PDT, 1999).

1998 National Convective Weather Forecast Product (NCWFP98): The 1998 NCWFP, also developed by the CWPDT, is an earlier version of the NCWFP. During the summer of 1999, the 1998 NCWFP was run at the AWC, utilizing some slightly different data sources (e.g., for radar data) than were used for the 1999 version of the NCWFP, which was run at NCAR.

Convective SIGMET (C-SIGMET): This product, generated by AWC forecasters, is an operational text forecast of convective activity. The forecast is produced hourly and is valid for up to 2 h (NWS 1991). That is, the C-SIGMET is not intended to be valid at a particular time, but rather, is valid across a period. The text is decoded into latitude and longitude vertices. The C-SIGMETs are evaluated in three ways to allow comparisons with the NCWFP. These three treatments of the C-SIGMETs include: (a) as a forecast of length 1 h, valid at the *end* of the period; (b) as a forecast of length 2 h, valid at the *end* of the period; and (c) as a forecast of length 2h, valid *throughout* the entire 2-h period.

Convective SIGMET Outlook (C-SIGMET Outlook): The convective outlook is an operational text forecast of convective activity, generated by AWC meteorologists, issued hourly, and valid from 2-6 h after the issuance time of the C-SIGMET outlook (NWS 1991). The text is decoded into latitude and longitude vertices. The forecast area encompasses moving and changing weather over the 4-h period. For this evaluation, the outlooks are considered in two ways: (a) as a forecast of length 6 h, valid at the *end* of the period and (b) as a forecast of length 4 h, valid *throughout* the 2-to-6-h period after issuance.

2.2 Observations

The observations used as verification data include lightning reports, radar data, and output of the National Convective Weather Detection Product (NCWDP).

Lightning data were obtained from the National Lightning Data Network (NLDN; Orville 1991). These data include information regarding the locations (latitude and longitude) and times of specific lightning strikes. The lightning observations were used alone and in combination with radar data to infer areas of active convection for verification of the forecasts.

Radar reflectivity (dBZ) fields, available on a 4-km grid, were used as a second type of observed convective field. A threshold of 40 dBZ was used to define areas of convection.

Finally, the NCWDP combines a 2-dimensional mosaic of radar reflectivity with radar-derived cloud top data and a grid of lightning detections from the NLDN. The cloud top data primarily are used to remove anomalous propagation and ground clutter, and the lightning data help to keep the NCWDP current, since lightning data have a lower latency than radar data. The NCWDP fields were available on a 4-km grid, with convective storms delineated by a threshold of 40 dBZ, or more than 3 lightning strikes in 10 minutes.

3. Mechanics

3.1 Forecast/observation matching procedures

The following procedures were used to match the forecasts to the observation grids. These methods were developed in collaboration with the CWPDT, prior to the start of the real-time intercomparison.

Before the forecasts were matched to the observations, a 20-km grid was laid over the observation field. Each box on the overlay grid was assigned a *Yes* or *No* value depending on whether a positive observation fell within the 20-km box. Specific criteria used to define a positive observation (depending on which type of observation was being used) for each 20-km box included: (a) 4 lightning strikes anywhere in the 20-km box; (b) one 4-km box with radar reflectivity greater than 40 dBZ anywhere in the 20-km box; and (c) one 4-km NCWDP observation with a reflectivity greater than 40 dBZ anywhere in the 20-km box. The same procedures were applied to the forecasts, with a 20-km box labeled as *Yes* when any part of the forecast polygon intersected that box. If a forecast polygon did not intersect the 20-km box, then a *No* forecast was assigned to the box.

For some analyses, a filter was applied to the NCWDP observations in an attempt to screen out isolated short-lived convection. In this case, a 20-km box was assigned a *Yes*

observation when 12 or more (approximately half) of the enclosed 4-km NCWDP boxes met the 40 dBZ criteria. Otherwise, a *No* observation was assigned to the 20-km box.

Once the matching process was completed, each box on the 20-km observation grid was matched to each 20-km box on the forecast grid. This technique produced the forecast/observation pairs used to generate the verification statistics. For example, a *Yes* forecast box overlapping a *Yes* observation box produced a *Yes*-*Yes* pair. Similarly, a *Yes* forecast and *No* observation produced a *Yes*-*No* pair, and so on.

It is important to note that the matching approach utilized here was developed to allow comparisons among a variety of types of forecasts. In particular, it was designed to allow comparisons between the NCWFP and the C-SIGMETs and C-SIGMET Outlooks, which have a different form from the NCWFP, and generally cover broader, contiguous areas. This method differs from the method used by the CWPDT in developing the NCWFP. For their verification of the NCWFP, the CWPDT directly compared the 4-km forecast and observation boxes, with a 2-grid relaxation of both the forecasts and observations. In contrast, the method used here (particularly without filtering) can extend a single 4-km forecast observation to cover a much larger area. The 12-grid-box filter somewhat reduces the impact of this effect.

Forecasts were evaluated over the entire national domain, extending from the Atlantic Ocean on the East to the Pacific Ocean on the West, and from the U.S.-Canadian border on the North to the U.S.- Mexican border on the South. The entire period (24 h per day) between 1 June and 31 August 1999 was included in the evaluation.

3.2 Time window

Observations that fell within a 10-minute time window prior to the forecast valid time were mapped to the 20-km grid and used for verification. To ensure consistency among the results, all forecast products, excluding the 4-h C-SIGMET Outlook and the 2-h C-SIGMET were subjected to this criterion. Additional criteria were applied to the 4-h C-SIGMET Outlook and the 2-h C-SIGMET; for these forecasts, all observations within the 4-h and 2-h periods, respectively, were mapped to the grid.

4. Verification methods

4.1 Contingency tables

The *Yes/No* forecast/observation pairs were used to create counts, to fill in a 2x2 contingency table like the one shown in Table 1. That is, for a given forecast, all of the 20-km boxes with a *Yes* forecast and a *Yes* observation were counted to obtain YY; all of the 20-km boxes with a *Yes* forecast and a *No* observation were counted to obtain YN; and so on. Individual forecast contingency tables were accumulated to obtain tables representing particular days, months, or other periods (including the entire forecast period).

Table 1. Basic contingency table for evaluation of dichotomous (e.g., Yes/No) forecasts. Elements in the cells are the counts of forecast-observation pairs.

	Observation		
Forecast	***Yes***	***No***	***Total***
Yes	YY	YN	YY+YN
No	NY	NN	NY+NN
Total	YY+NY	YN+NN	YY+YN+NY+NN

4.2 Statistical measures

Table 2 lists the verification statistics that were included in the NCWFP evaluation, with PODy, PODn, and FAR representing the basic verification statistics. General descriptions of these statistics include the following:

- PODy and PODn are estimates of the proportions of *Yes* and *No* observations, respectively, that were correctly forecast (e.g., Brown et al. 1997).
- FAR is the proportion of *Yes* forecasts that were incorrect.
- Bias is the ratio of the number of *Yes* forecasts to the number of *Yes* observations, and is a measure of over- or under-forecasting.
- The Critical Success Index (CSI), also known as the Threat Score, is the proportion of hits that were either forecast or observed.
- The True Skill Statistic (TSS) (e.g., Doswell et al. 1990) is a measure of the ability of the forecast to discriminate between *Yes* and *No* observations; TSS also is known as the Hanssen-Kuipers discrimination statistic (Wilks 1995).
- The Heidke Skill Score (HSS) is the percent correct, corrected for the number expected to be correct by chance.
- The Gilbert Skill Score (GSS) (Schaefer 1990), also known as the Equitable Threat Score, is the CSI corrected for the number of hits expected by chance.
- The % Area is the percent of the total possible area that had a *Yes* forecast (Brown et al. 1997).

5. Results

Throughout the 1999 evaluation of the NCWFP, statistical results (tables and displays) were generated by the RTVS and presented on the web-based interface in near-real-time. A vast wealth of results is still available on the web site[3]. Because so many comparisons were included in the evaluation, only a selection of them can be presented in this report. Thus, this section includes an overview of the most relevant statistical results. Additional results related to the Collaborative Convective Forecast Product (CCFP) and

[3] http://www-ad.fsl.noaa.gov/afra/rtvs/convective/main_convective.html

Table 2. Verification statistics used in this study.

Statistic	*Definition*	*Description*
PODy	YY/(YY+NY)	Probability of Detection of "Yes" observations
PODn	NN/(YN+NN)	Probability of Detection of "No" observations
FAR	YN/(YY+YN)	False Alarm Ratio
CSI	YY/(YY+NY+YN)	Critical Success Index
Bias	(YY+YN)/(YY+NY)	Forecast Bias
TSS	PODy + PODn – 1	True Skill Statistic
HSS	[(YY+NN)-C1]/(N-C1), where N=YY+YN+NY+NN C1=[(YY+YN)(YY+NY) + (NY+NN)(YN+NN)] / N	Heidke Skill Score
GSS	(YY-C2)/[(YY-C2)+YN+NY], where C2=(YY+YN)(YY+NY)/N	Gilbert Skill Score
% Area	(Forecast Area) / (Total Area) x 100	% of the area of the continental U.S. where convection is forecast to occur

C-SIGMET Outlooks can be found in Mahoney et al. (2000). A future report will focus on an in-depth evaluation of the quality of the C-SIGMETs and C-SIGMET Outlooks.

Overall results for the NCWFP and other relevant products are presented in the following section. Some hourly and daily results for the NCWFP and the C-SIGMETs are presented in Section 5.2, followed by results for a few specific cases in Section 5.3.

5.1 Overall results

Overall results for the various types of forecasts and verification observations are shown in Table 3. The statistics in this table were created by combining all of the counts for all of the days and hours for which data were available between 1 June and 31 August 1999. In Table 3, the 1999 version of the NCWFP is simply denoted as "NCWFP," and the 1998 version is denoted "NCWFP98."

The results in Table 3 are stratified according to the type of observations used to evaluate the forecasts (i.e., NCDP, Lightning, Radar). Comparisons of the results based on radar observations to those based on the NCDP indicate that the verification statistics are very similar for these two types of verification data. In contrast, when lightning observations are used alone to verify the forecasts, the FAR and Bias values increase dramatically for all of the forecast types. This result suggests that the lightning observations are much sparser than the radar observations and probably should not be used as a separate verification dataset. Thus, the remainder of the discussion in this section focuses on the NCDP results presented in Table 3.

Comparisons between the results for the 1-h NCWFP98 and NCWFP forecasts suggest that the NCWFP provides an improvement over the NCWFP98. In particular, the NCWFP skill scores (HSS and GSS) are larger than the skill scores for the NCWFP98 forecasts. In addition, the PODy values are somewhat larger and the FAR values are somewhat smaller for the NCWFP forecasts than for the NCWFP98 forecasts. However, it is important to note that at least some of these differences may be due to the slightly different datasets used to generate the NCWFP98 forecasts (see Section 2).

The largest differences between the NCWFP verification statistics and the statistics for the C-SIGMETs are associated with the Bias and % Area values. In particular, the 1-hr NCWFP statistics have a very small overall Bias, which indicates that the NCWFP consistently underforecast convective regions. In contrast, the 1-h C-SIGMETs have a Bias of about 1 for the unfiltered observations, and a larger Bias (3.1) for the filtered observations. Thus, the C-SIGMETs, evaluated as point forecasts, forecast approximately the correct amount of convection when the data are not filtered, and they forecast an area that is too large by a factor of three when the smaller convective areas are removed. Correspondingly, the FAR values are 50-100% larger for the C-SIGMETs than for the NCWFP forecasts.

Due to the small areas covered by the NCWFP, and the low Bias values associated with the NCWFP forecasts, the PODy values for the NCWFP forecasts also are relatively small. The PODy values for the C-SIGMETs are considerably larger. In contrast, the PODn values for both types of forecasts are quite close to 1. The summary and skill statistics (CSI, HSS, and GSS) all indicate that both the C-SIGMETs and the NCWFP have some skill in forecasting convective areas, with greater skill associated with the 1-h forecasts and less skill associated with the 2-h forecasts. Differences between the skill scores for the C-SIGMETs and NCWFP forecasts are relatively small when the large differences in the Bias values are taken into account. In particular, most skill scores for forecasts of rare events can be improved by overforecasting (Marzban 1998).

Table 3. Overall verification results, stratified by observation type, forecast product, and forecast length. Statistics were computed by combining counts for all days and hours.

Product	Fcst. length (h)	Filter	PODy	PODn	FAR	CSI	HSS	GSS	Bias	%Area
NCDP										
NCWFP	1	None	0.13	1.00	0.32	0.12	0.21	0.12	0.2	0.51
NCWFP98	1	None	0.06	1.00	0.31	0.06	0.11	0.06	0.1	0.22
C-SIGMETs	1	None	0.33	0.98	0.67	0.20	0.31	0.19	1.0	2.30
NCWFP	1	12	0.25	1.00	0.61	0.18	0.29	0.17	0.6	0.74
NCWFP98	1	12	0.12	1.00	0.58	0.10	0.19	0.10	0.3	0.22
C-SIGMETs	1	12	0.47	0.98	0.85	0.13	0.22	0.12	3.1	2.29
NCWFP	2	None	0.07	1.00	0.50	0.06	0.11	0.06	0.1	0.31
C-SIGMETs	2	None	0.25	0.98	0.75	0.14	0.23	0.13	1.0	2.29
C-SIGMETs	0-2	None	0.24	0.99	0.37	0.21	0.32	0.19	0.4	2.30
NCWFP	2	12	0.10	1.00	0.74	0.08	0.14	0.08	0.4	0.42
C-SIGMETs	2	12	0.32	0.98	0.90	0.09	0.15	0.08	3.1	2.29
C-SIGMETs	0-2	12	0.33	0.99	0.55	0.23	0.36	0.22	0.7	2.30
Outlooks	2-6	None	0.47	0.89	0.70	0.22	0.28	0.17	1.6	14.19
Outlooks	6	None	0.43	0.87	0.93	0.06	0.08	0.04	6.1	14.11
Outlooks	2-6	12	0.52	0.88	0.81	0.16	0.22	0.12	2.8	14.25
Outlooks	6	12	0.48	0.86	0.98	0.02	0.03	0.02	19.3	14.08
Lightning										
NCWFP	1	None	0.27	1.00	0.83	0.11	0.20	0.11	1.7	0.51
NCWFP98	1	None	0.15	1.00	0.83	0.09	0.16	0.09	0.9	0.21
C-SIGMETs	1	None	0.53	0.98	0.94	0.05	0.10	0.05	9.4	2.28
NCWFP	2	None	0.14	1.00	0.89	0.06	0.12	0.06	1.2	0.31
C-SIGMETs	2	None	0.36	0.98	0.96	0.04	0.06	0.03	9.4	2.27
C-SIGMETs	0-2	None	0.42	0.98	0.73	0.20	0.32	0.19	1.6	2.30
Outlooks	2-6	None	0.56	0.87	0.90	0.09	0.13	0.07	5.5	14.25
Outlooks	6	None	0.48	0.86	0.99	0.01	0.01	0.01	59.2	14.25
Radar										
NCWFP	1	None	0.13	1.00	0.34	0.12	0.21	0.12	0.2	0.50
NCWFP98	1	None	0.06	1.00	0.29	0.06	0.11	0.06	0.1	0.21
C-SIGMETs	1	None	0.32	0.98	0.67	0.19	0.31	0.18	1.0	2.27
NCWFP	2	None	0.07	1.00	0.47	0.07	0.12	0.06	0.1	0.31
C-SIGMETs	2	None	0.25	0.98	0.75	0.14	0.24	0.13	1.0	2.28
C-SIGMETs	0-2	None	0.23	0.99	0.40	0.20	0.31	0.18	0.4	2.31
Outlooks	2-6	None	0.47	0.89	0.71	0.22	0.28	0.16	1.6	14.30
Outlooks	6	None	0.43	0.87	0.93	0.06	0.08	0.04	6.2	14.14

Filtering the observations has a large impact on the values of the basic verification statistics in Table 3 (note that the filtering stratification of the results is only relevant when the NCDP is used as the verification data). The NCWFP results for the 1-h forecasts have an overall PODy value of 0.13 when the observations are *not* filtered, and a value of 0.25 when the data *are* filtered. Similarly, the FAR value for 1-h NCWFP forecasts is 0.32 when the observations are unfiltered and 0.61 when they are filtered. The skill scores are influenced somewhat less by the filtering; for example, the HSS for the NCWFP 1-h forecasts changes from 0.21 when the data are not filtered to 0.29 when they are filtered. These relatively small changes in the skill scores reflect the trade-off between PODy and FAR.

The results in Table 3 also suggest that the 2-h forecasts have relatively skill, compared to the 1-h forecasts; however, both the C-SIGMETs and NCWFP exhibit positive skill at this lead time. When the C-SIGMETs are evaluated over the entire 2-h period, they have a smaller FAR and Bias than the 2-h SIGMETs. Finally, the verification statistics for the C-SIGMET Outlooks indicate (not surprisingly) that these forecasts cover much broader areas than the C-SIGMETs and the NCWFP, and that they have much larger Bias values.

5.2 Hourly and daily results

Tables 4 and 5 contain the overall verification statistics, by hour of day, for the NCWFP and the C-SIGMETs, respectively. These tables only include results for the 1-h forecasts, using filtered observations. Results for the 2-h forecasts and for the unfiltered observations are presented in Appendix A.

The results in Table 4 suggest that the performance of the NCWFP is best, overall, in the period between 0000 and 0600 UTC. In particular, the overall PODy values and skill scores generally are largest during this period. The statistics for the C-SIGMETs (Table 5) do not exhibit this diurnal variability to the same extent as the statistics for the NCWFP. Comparing the NCWFP results with the results for the C-SIGMETs suggests that, for most hours, the NCWFP forecasts have somewhat larger skill scores (CSI, HSS, and GSS), and much smaller FAR, Bias, and % Area. In contrast, the C-SIGMETs have larger PODy values, by nearly a factor of two.

Day-to-day variations in the verification statistics are illustrated in Figure 1, which shows the daily statistics for the 1- and 2-h NCWFP forecasts valid at 0000 UTC, using unfiltered observations. The diagrams in Figure 1 demonstrate that the PODy and FAR values vary from day-to-day, with greater variations in FAR than in PODy. In contrast, the PODn values consistently are close to or equal to one. The lack of variability in PODn, relative to the much greater variability in PODy, is clearly illustrated in the PODy vs. 1-PODn diagram (second diagram in the righthand column). The nearly constant nature of PODn results from the fact that the NCWFP predicts convection over such a small proportion of the country. Due to the consistently large values of PODy, the

Table 4. Overall results by issue time for the 1-h NCWFP forecasts with verification based on the filtered observations.

Issue time (UTC)	Fcst. length (h)	Valid time (UTC)	PODy	PODn	FAR	CSI	TSS	HSS	GSS	Bias	% Area
0	1	1	0.25	1.00	0.57	0.19	0.25	0.31	0.18	0.58	0.65
1	1	2	0.25	1.00	0.57	0.19	0.24	0.31	0.18	0.57	0.53
2	1	3	0.26	1.00	0.53	0.20	0.26	0.33	0.20	0.57	0.42
3	1	4	0.27	1.00	0.51	0.21	0.27	0.35	0.21	0.56	0.37
4	1	5	0.27	1.00	0.52	0.21	0.27	0.34	0.21	0.55	0.33
5	1	6	0.27	1.00	0.48	0.21	0.26	0.35	0.21	0.51	0.29
6	1	7	0.25	1.00	0.49	0.20	0.24	0.33	0.20	0.49	0.26
7	1	8	0.22	1.00	0.51	0.18	0.21	0.30	0.17	0.44	0.22
8	1	9	0.22	1.00	0.48	0.18	0.22	0.31	0.18	0.42	0.21
9	1	10	0.22	1.00	0.48	0.19	0.22	0.31	0.18	0.42	0.21
10	1	11	0.20	1.00	0.51	0.16	0.20	0.28	0.16	0.40	0.21
11	1	12	0.22	1.00	0.51	0.18	0.22	0.30	0.18	0.44	0.21
12	1	13	0.20	1.00	0.51	0.17	0.20	0.28	0.17	0.41	0.19
13	1	14	0.20	1.00	0.52	0.16	0.20	0.28	0.16	0.41	0.17
14	1	15	0.14	1.00	0.56	0.12	0.14	0.21	0.12	0.33	0.12
15	1	16	0.13	1.00	0.57	0.11	0.13	0.20	0.11	0.31	0.12
16	1	17	0.14	1.00	0.53	0.12	0.13	0.21	0.12	0.28	0.14
17	1	18	0.13	1.00	0.51	0.12	0.13	0.20	0.11	0.27	0.19
18	1	19	0.14	1.00	0.52	0.12	0.14	0.22	0.12	0.30	0.29
19	1	20	0.17	1.00	0.58	0.14	0.16	0.23	0.13	0.39	0.47
20	1	21	0.19	1.00	0.60	0.15	0.18	0.25	0.14	0.46	0.56
21	1	22	0.20	1.00	0.60	0.15	0.19	0.26	0.15	0.48	0.60
22	1	23	0.21	1.00	0.58	0.16	0.21	0.28	0.16	0.51	0.72
23	1	0	0.25	1.00	0.61	0.18	0.24	0.29	0.17	0.63	0.74
Total			0.24	1.00	0.57	0.18	0.23	0.30	.018	0.54	0.51

Table 5. As in Table 4, for the 1-h C-SIGMETs.

Issue time (UTC)	Fcst. length (h)	Valid time (UTC)	PODy	PODn	FAR	CSI	TSS	HSS	GSS	Bias	% Area
0	1	1	0.49	0.97	0.87	0.12	0.45	0.20	0.11	3.67	3.92
1	1	2	0.48	0.97	0.86	0.12	0.45	0.21	0.12	3.41	3.11
2	1	3	0.48	0.98	0.84	0.14	0.46	0.23	0.13	2.97	2.25
3	1	4	0.46	0.98	0.83	0.14	0.44	0.24	0.14	2.73	1.84
4	1	5	0.49	0.99	0.82	0.15	0.47	0.26	0.15	2.62	1.60
5	1	6	0.48	0.99	0.81	0.16	0.47	0.27	0.15	2.53	1.44
6	1	7	0.50	0.99	0.80	0.16	0.49	0.28	0.16	2.53	1.37
7	1	8	0.49	0.99	0.81	0.16	0.48	0.27	0.15	2.57	1.30
8	1	9	0.48	0.99	0.80	0.17	0.47	0.28	0.16	2.39	1.18
9	1	10	0.45	0.99	0.79	0.16	0.45	0.28	0.16	2.21	1.12
10	1	11	0.45	0.99	0.81	0.15	0.44	0.26	0.15	2.4	1.20
11	1	12	0.41	0.99	0.83	0.14	0.40	0.24	0.13	2.37	1.13
12	1	13	0.41	0.99	0.84	0.13	0.40	0.22	0.12	2.64	1.20
13	1	14	0.40	0.99	0.86	0.12	0.39	0.20	0.11	2.87	1.14
14	1	15	0.38	0.99	0.87	0.11	0.37	0.19	0.10	3.00	1.12
15	1	16	0.38	0.99	0.87	0.11	0.38	0.20	0.11	2.85	1.11
16	1	17	0.39	0.99	0.84	0.13	0.38	0.22	0.12	2.49	1.20
17	1	18	0.41	0.99	0.83	0.14	0.40	0.24	0.13	2.38	1.61
18	1	19	0.42	0.98	0.83	0.14	0.40	0.23	0.13	2.53	2.42
19	1	20	0.47	0.97	0.85	0.13	0.44	0.22	0.12	3.07	3.39
20	1	21	0.50	0.96	0.86	0.12	0.46	0.20	0.11	3.67	4.43
21	1	22	0.50	0.96	0.88	0.11	0.45	0.18	0.10	4.00	5.02
22	1	23	0.50	0.96	0.87	0.12	0.45	0.19	0.11	3.77	5.07
23	1	0	0.51	0.96	0.88	0.11	0.46	0.18	0.10	4.24	4.83
Totals	1		0.47	0.98	0.85	0.13	0.45	0.22	0.12	3.12	2.29

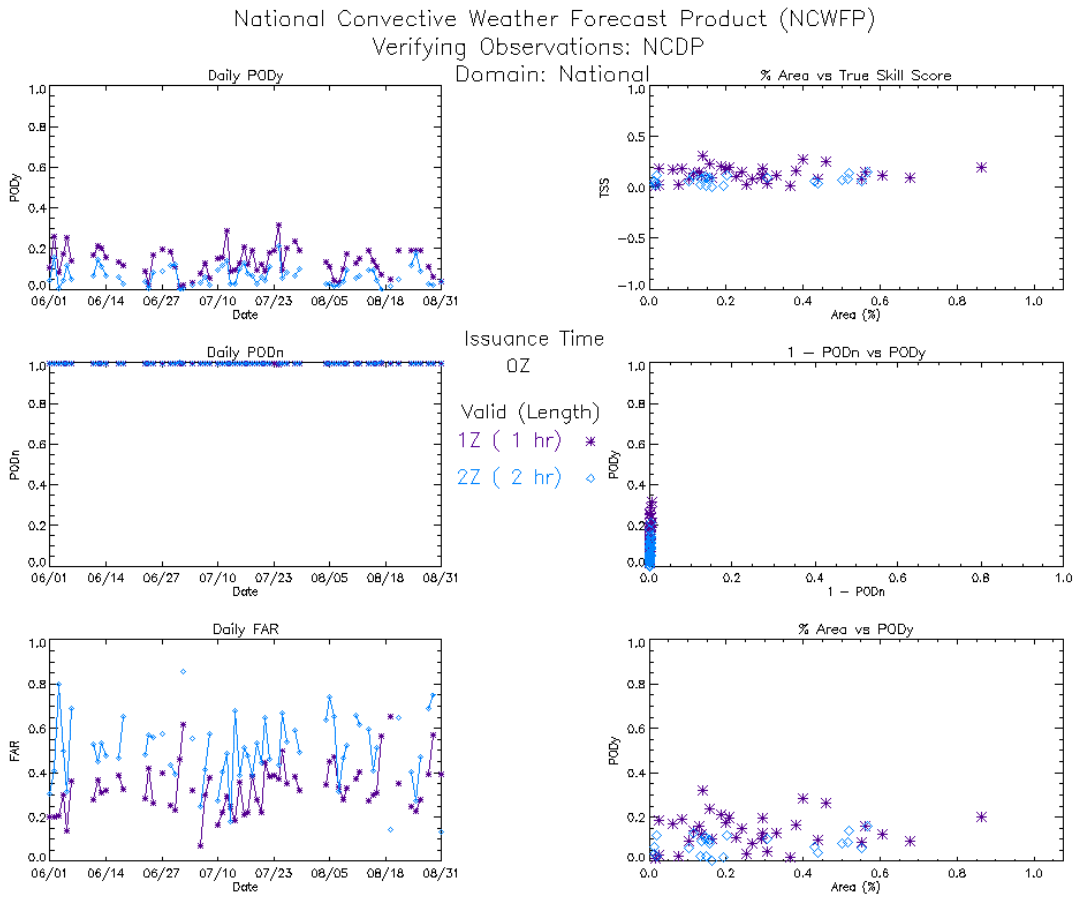


Figure 1. Daily variations in verification statistics for NCWFP forecasts issued at 0000 UTC.

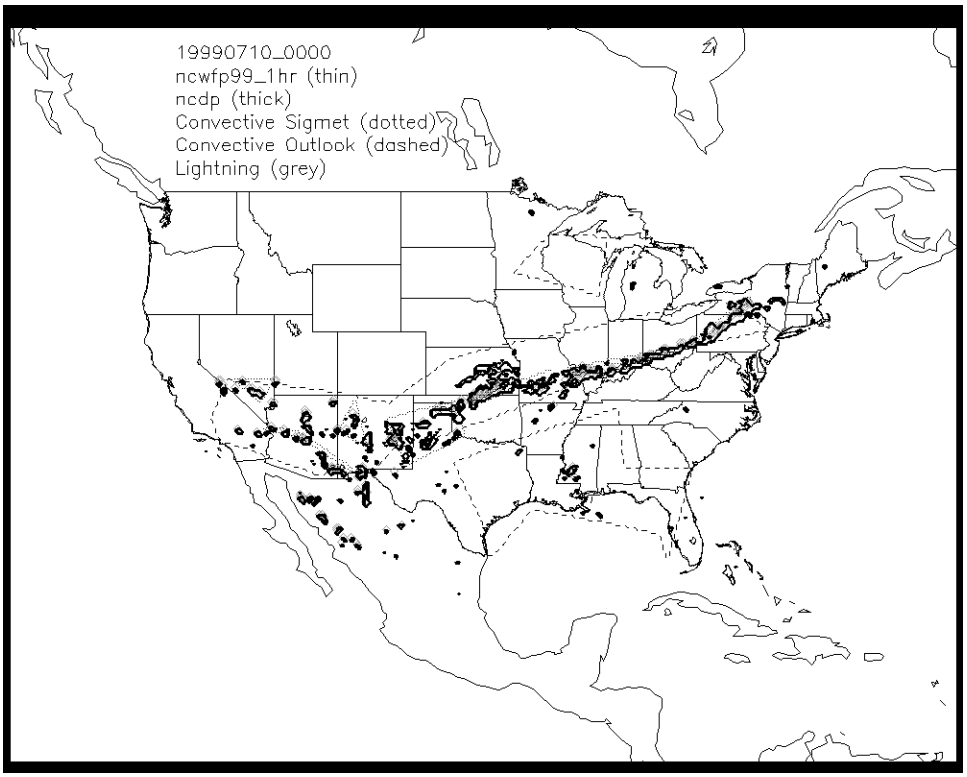
TSS values are essentially equivalent to PODy, and TSS provides little additional information, as shown in Figure 1 and in Tables 4 and 5.

5.3 *Example cases*

In this section, the verification results for several specific cases (days and hours) are shown, to illustrate some of the variability in the statistics. In addition, the effect of the filter on the observations and forecasts is demonstrated for one case.

Filter effects. The maps in Figure 2 illustrate the effects of the observation filter for a particular case (10 July at 0000 UTC). The two diagrams in this figure show the actual 4-km NCDP observations that were included both without and with application of the filter, along with outlines of the forecasts that were valid at this time. As shown in Figure 2, a large number of very small convective areas were removed by application of the filter.

(a)



(b)

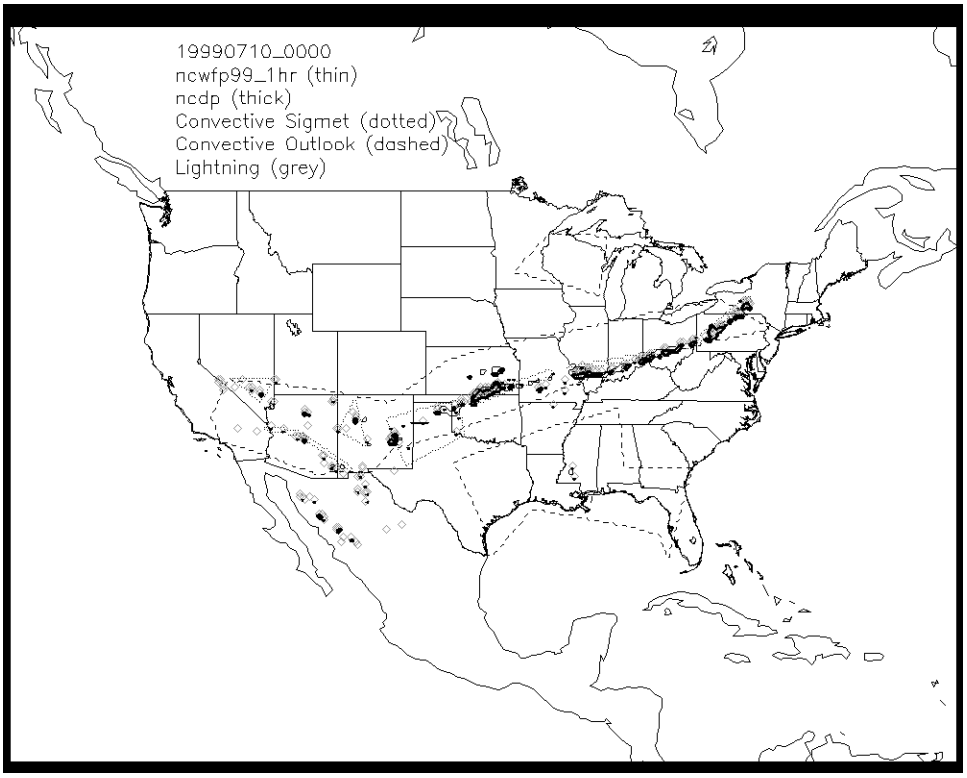


Figure 2. Maps demonstrating the impact of the filter on the observations, for forecasts valid at 0000 UTC on 10 July 1999. In (a) the NCDP values are not filtered and in (b) the observation filter has been applied.

9-10 July 1999. The maps in Figures 3-5 show the 20-km forecast and observation boxes that were included in the evaluation of the NCWFP and the C-SIGMETs at three different times on the 9th and 10th of July, 1999, with the filter applied to the observations. Table 6 presents verification statistics for these times. As shown in Table 6, the verification statistics change notably through this period, especially for the NCWFP. In fact, at the latest time, the PODy value for the NCWFP is about the same as the PODy value for the C-SIGMETs, and the NCWFP Bias is approaching 1. Even at 0000 UTC on 10 July, the NCWFP statistics have improved, whereas the major change in the C-SIGMETs is an increase in FAR. By 0200 UTC on 10 July, statistics for both the NCWFP and the C-SIGMETs are improved. Skill scores for both types of forecasts increase through the period. The good verification statistics attained by the NCWFP for this case probably can be attributed to the highly-organized and long-lived nature of the convection on this day. The larger Bias associated with the NCWFP at 0200 UTC on 10 July is likely to be a result of the fact that the convection was decaying during the latter part of the period (Cindy Mueller, personal communication).

30 June 1999. The forecast map for 30 June at 2000 UTC is shown in Figure 6. Table 7 contains verification statistics for this case. On this day, the convection was much less well-organized than the convection in the 9-10 July case. However, at the selected time, the verification statistics for the NCWFP are about the same as the overall values shown in Table 3, and they are better than the values for 2000 UTC shown in Table 4. In contrast, the C-SIGMET statistics for this case are somewhat better than the overall values in Table 3 and the statistics for hour 2000 in Table 5.

24 August 1999. The forecast map for 24 August at 1600 UTC is shown in Figure 7, and the verification results for this case are shown in Table 8. For this example, the Bias and PODy values for the NCWF are smaller than usual, as represented by the numbers in Table 3, but they are close to the typical values for NCWFP forecasts valid at 1600 UTC (Table 4). In contrast, the C-SIGMET statistics for this case were better, in all cases, than the overall C-SIGMET statistics, as well as the C-SIGMET statistics for 1600 UTC (Table 5). In subsequent hours on this day, the statistics for the NCWFP improved.

19 August 1999. The forecast map for 19 August at 2200 UTC is shown in Figure 8. Table 9 contains verification statistics for this case. For this day and time, the NCWFP again performs about the same as average (as represented by the statistics in Tables 3 and 4), with a slightly larger PODy, smaller FAR and somewhat larger skill scores than the overall values. As in some of the other examples, the convection in this case was less well-organized, particularly compared to the 9-10 July case.

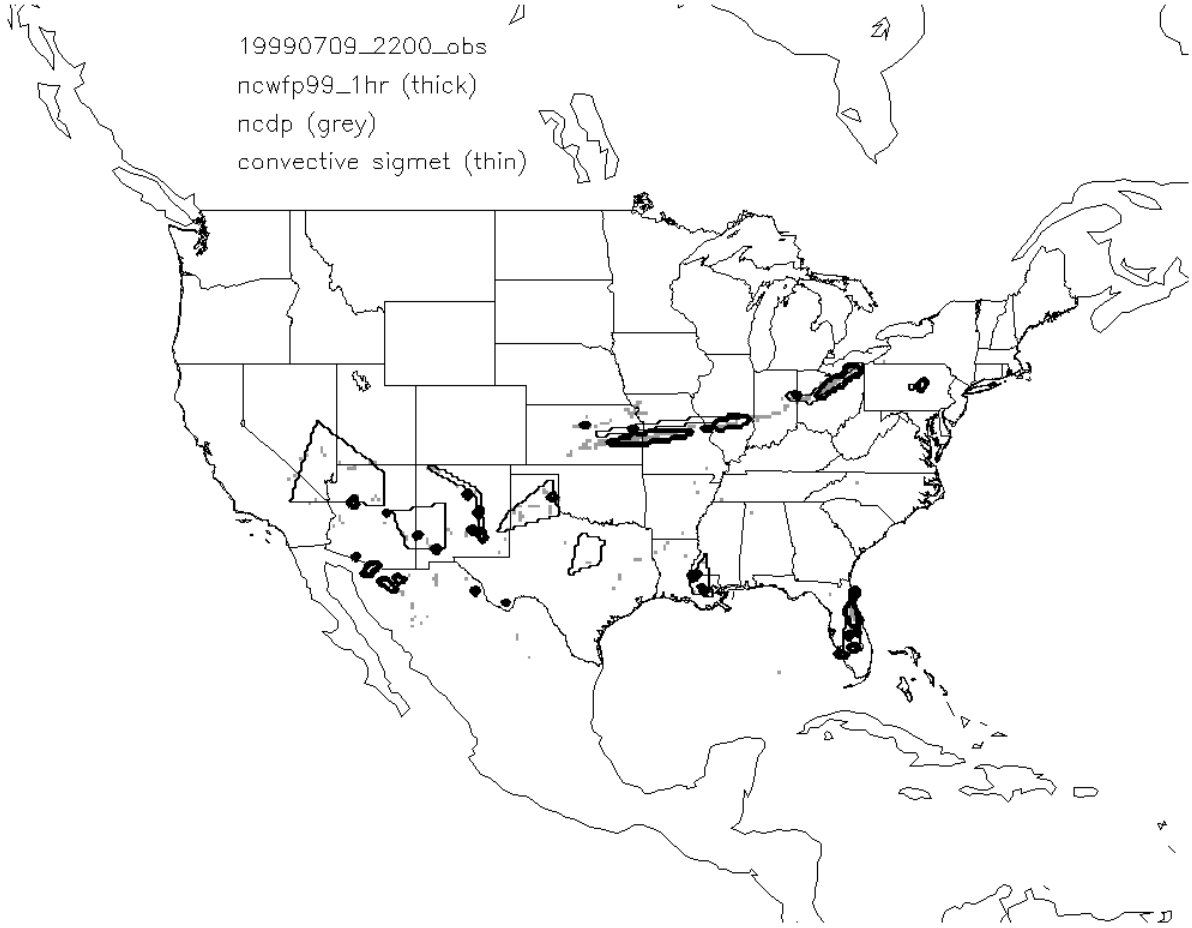


Figure 3. Map showing forecasts and filtered observations used in the verification analysis for forecasts valid at 2200 UTC on 9 July 1999. Forecast regions and observations shown are based on the actual 20-km boxes included in the verification analysis.

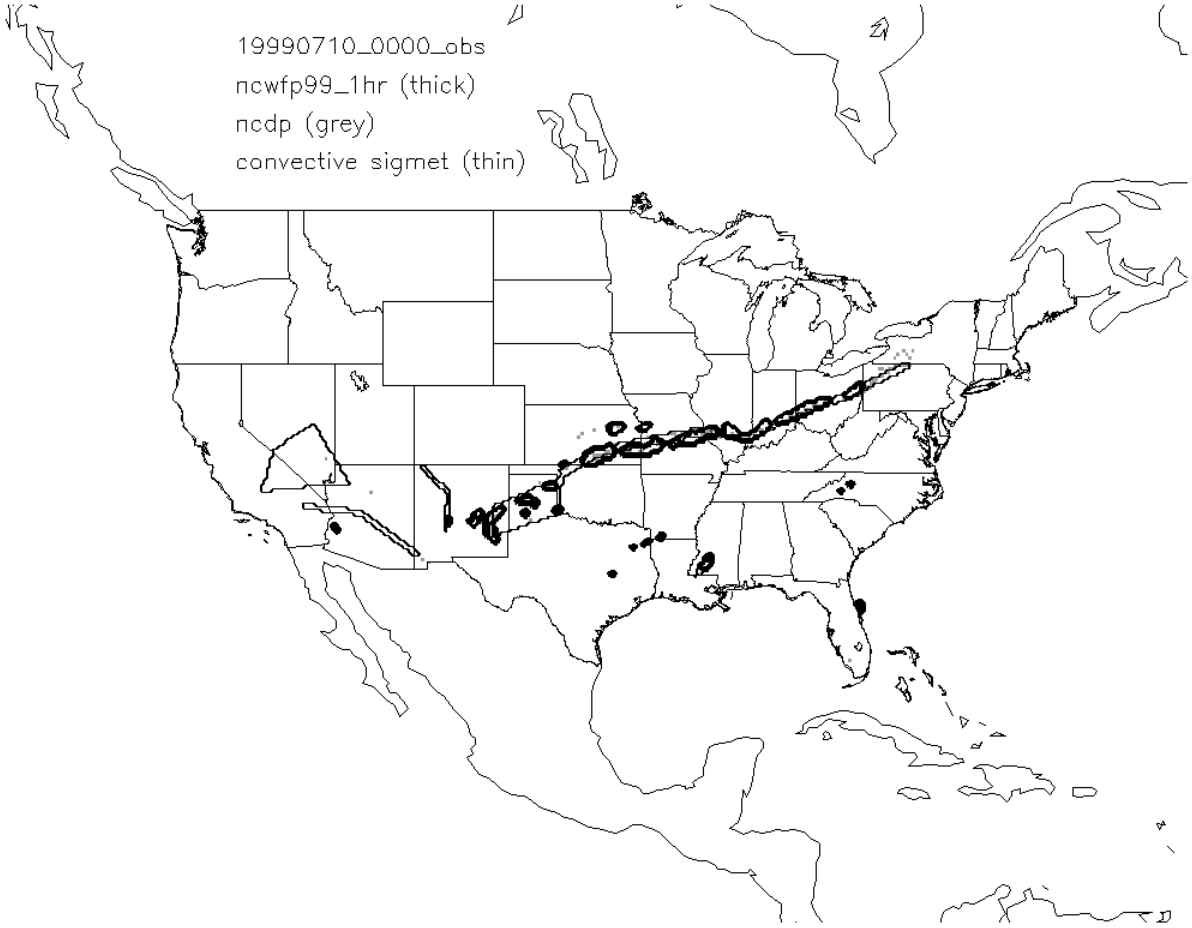


Figure 4. As in Figure 3, for forecasts valid at 0000 UTC on 10 July 1999.

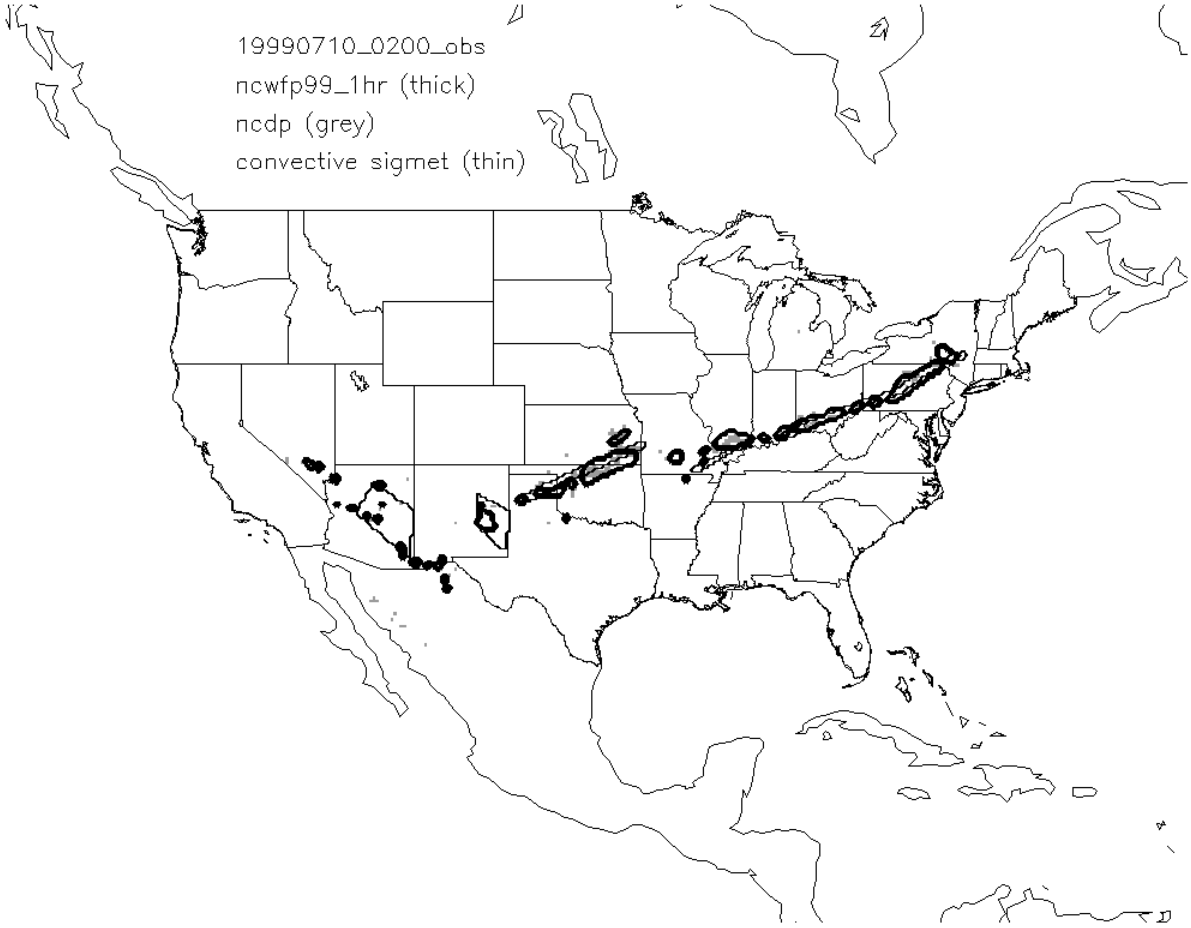


Figure 5. As in Figure 3, for forecasts valid at 0200 UTC on 10 July 1999.

Table 6. Verification statistics for selected hours on 9-10 July 1999, computed using filtered observations.

Product	PODy	PODn	FAR	CSI	HSS	GSS	Bias	%Area
9 July 1999, 2200 UTC								
NCWFP	0.20	1.00	0.46	0.17	0.28	0.17	0.4	0.97
C-SIGMETs	0.36	0.96	0.60	0.15	0.23	0.13	1.8	4.63
10 July 1999, 0000 UTC								
NCWFP	0.37	0.99	0.54	0.26	0.40	0.25	0.6	1.28
C-SIGMETs	0.39	0.97	0.81	0.14	0.24	0.13	2.1	3.37
10 July 1999, 0200 UTC								
NCWFP	0.45	0.99	0.50	0.31	0.46	0.30	0.89	1.55
C-SIGMETs	0.52	0.98	0.63	0.27	0.42	0.26	1.4	2.54

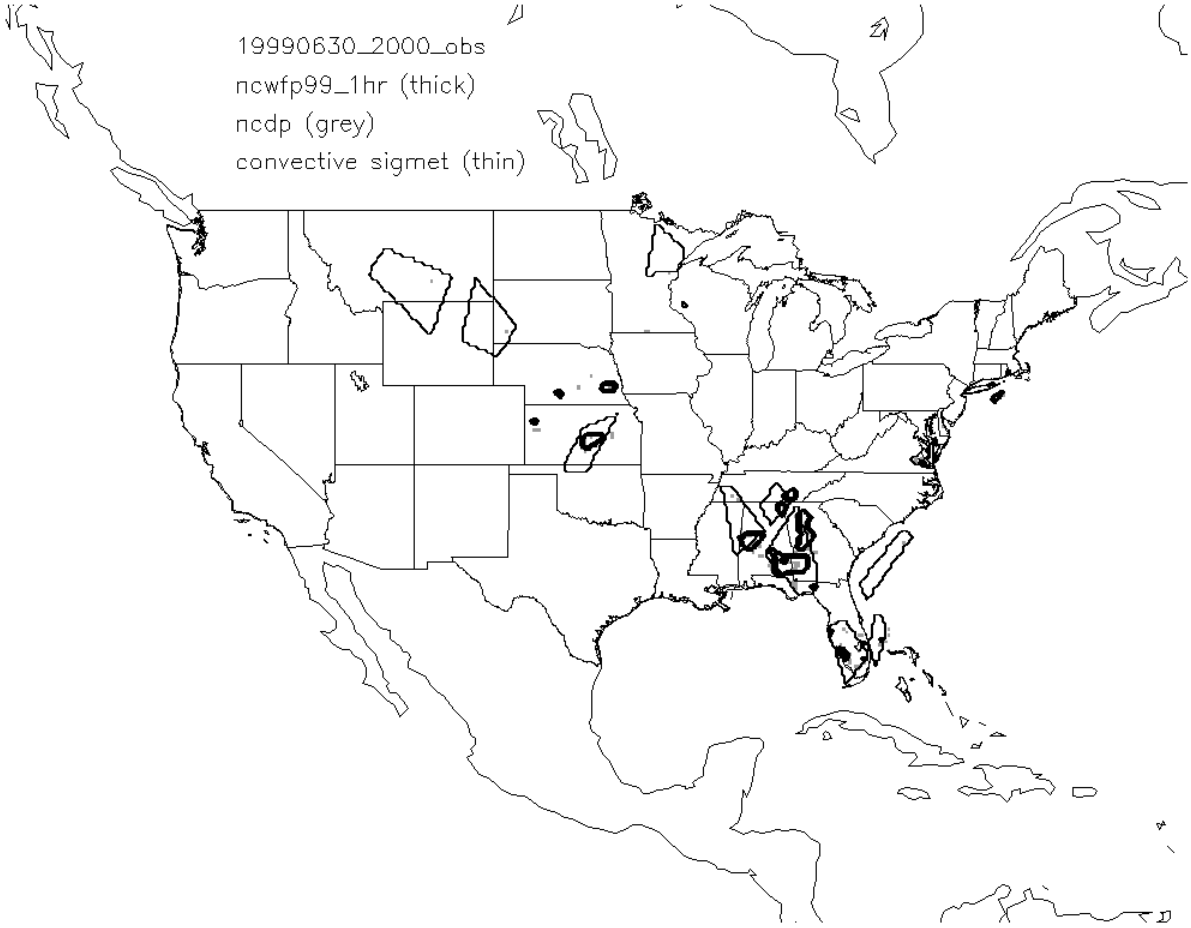


Figure 6. As in Figure 3, for forecasts valid at 2000 UTC on 30 June 1999.

Table 7. Verification statistics for 2000 UTC on 30 June 1999, computed using filtered observations.

Product	PODy	PODn	FAR	CSI	HSS	GSS	Bias	%Area
NCWFP	0.24	1.00	0.52	0.19	0.31	0.18	0.5	0.7
C-SIGMETs	0.53	0.95	0.87	0.12	0.19	0.10	4.1	9.2

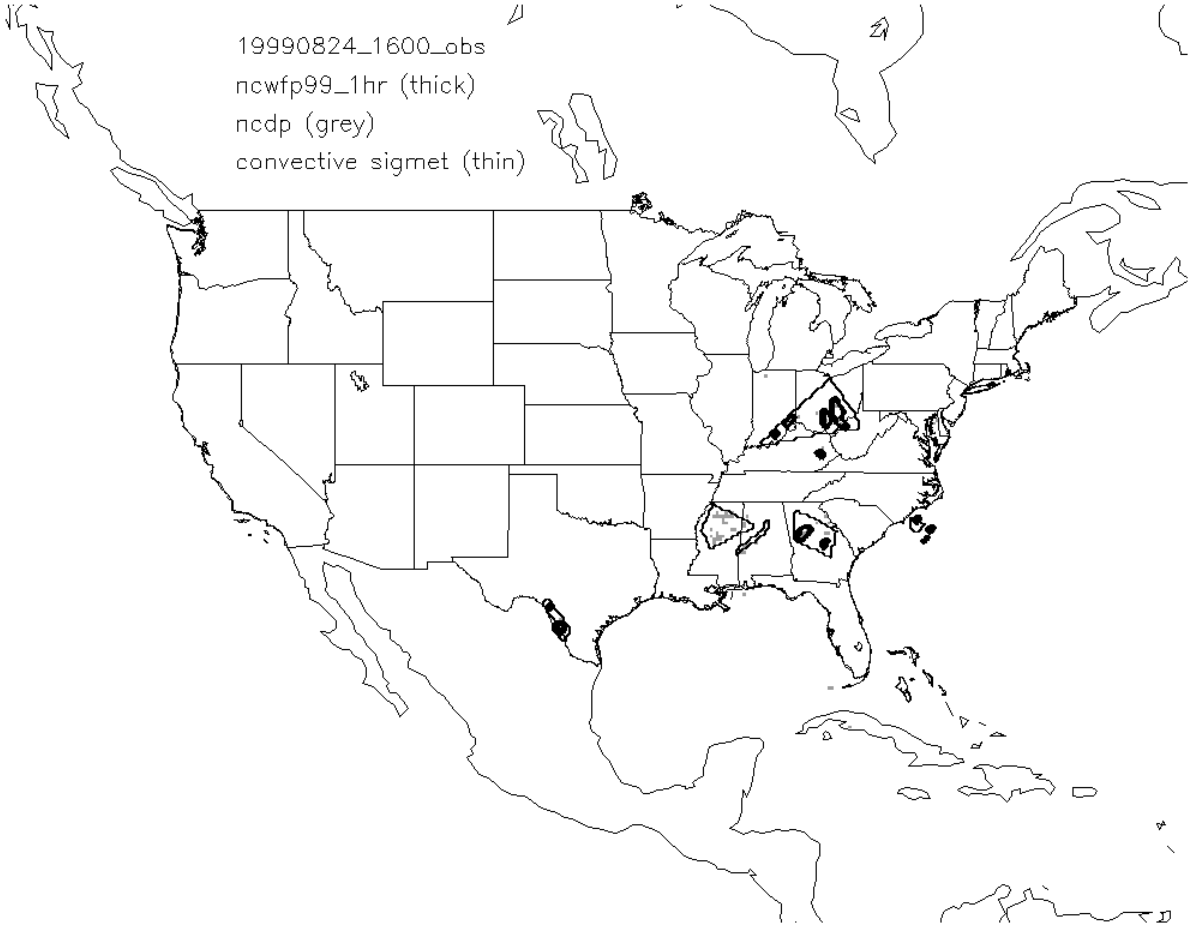


Figure 7. As in Figure 3, for forecasts valid at 1600 UTC on 24 August 1999.

Table 8. Verification statistics for 1600 UTC on 24 August 1999, computed using filtered observations.

Product	PODy	PODn	FAR	CSI	HSS	GSS	Bias	%Area
NCWFP	0.17	1.00	0.46	0.15	0.26	0.15	0.3	0.3
C-SIGMETs	0.62	0.98	0.74	0.23	0.36	0.22	2.3	2.2

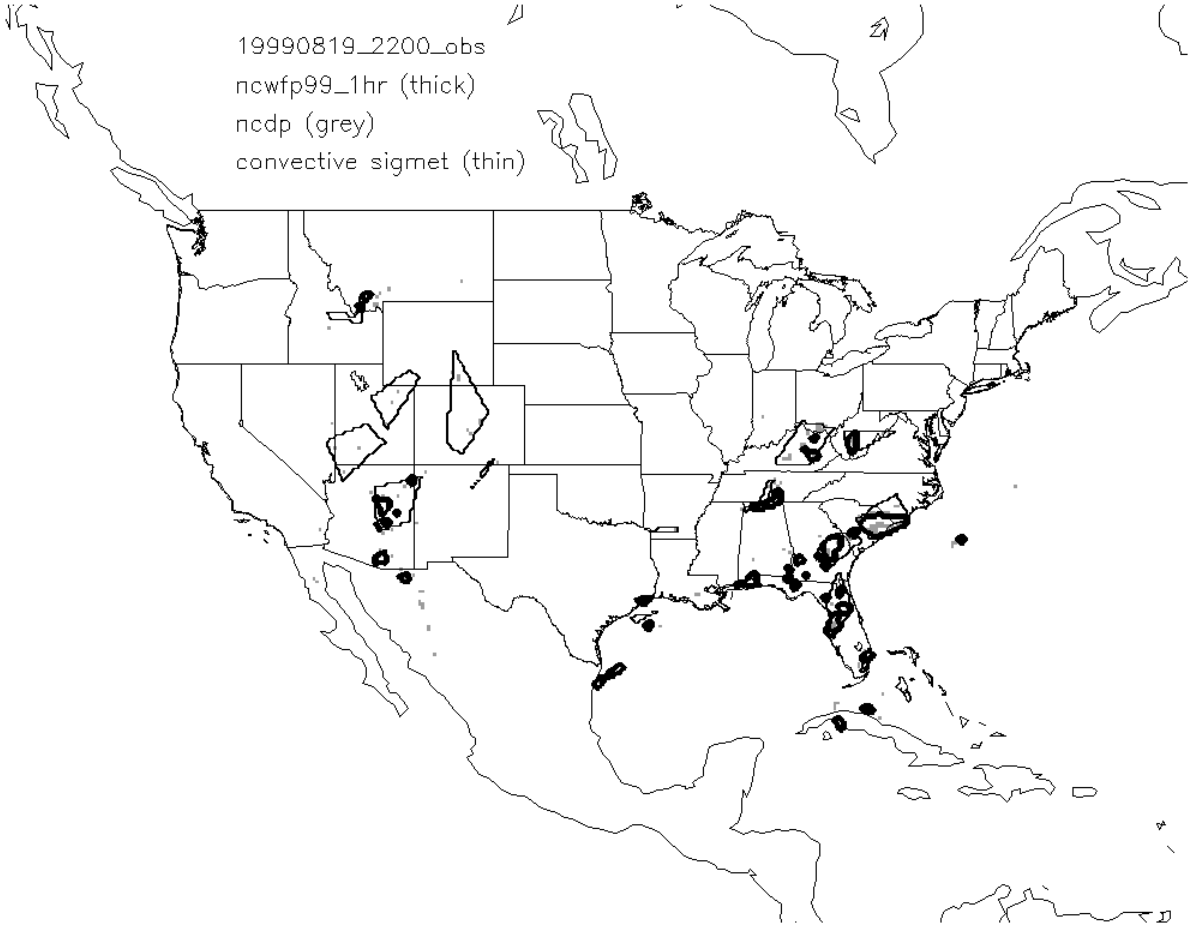


Figure 8. As in Figure 3, for forecasts valid at 2200 UTC on 19 August 1999.

Table 9. Verification statistics for 2200 UTC on 19 August 1999, computed using filtered data.

Product	PODy	PODn	FAR	CSI	HSS	GSS	Bias	%Area
NCWFP	0.28	0.99	0.56	0.21	0.33	0.20	0.6	1.1
C-SIGMETs	0.59	0.96	0.79	0.19	0.30	0.17	2.7	4.8

3. Discussion and conclusions

The TEGO Quality Assessment Plan for the NCWFP (January 2000) indicated that the required TEGO quality assessment should include the following aspects:

> *1) evaluate the NCWFP over at least one convective season; 2) compare the quality of the NCWFP to the quality of other relevant products (e.g., convective SIGMETs); 3) include all relevant forecast lengths and issue times in the evaluation; and 4) consider the day-to-day variations in the verification statistics.*

All of these aspects, as well as some others, were included in this evaluation of the NCWFP. However, a deeper analysis of the results would have been facilitated by the availability of additional resources, and should be pursued in the future. In particular, it would be desirable to develop and test additional methods for verifying these forecasts; additional methods and analyses would allow us to tell a more complete story regarding the quality of the NCWFP and other convective forecast products.

As noted earlier, the method used by the CWPDT to match forecasts and observations is very different from the method utilized here. In addition to differences in the matching methods, small areas (<500 km^2) were eliminated from the CWPDT analyses. Thus, it is not surprising that the verification statistics presented in Section 5 differ a great deal from the statistics reported by the CWPDT (Convective Weather PDT 1999). In particular, the CWPDT reported PODy values ranging from 0.45 to 0.53; FAR values ranging from 0.37 to 0.57; CSI values ranging from 0.30 to 0.35; and Bias values ranging from 0.7 to 1.2 for a set of about 7 days in the summer of 1998. These numbers generally represent better verification scores than those obtained in this study; however, the fact that they are "better" does not imply that they are also more "correct," since they were derived for a different purpose using a very different approach. Other verification analyses of forecasts generated by storm-motion extrapolation algorithms, in which direct grid-to-grid comparisons were made for 30-min forecasts, attained verification statistics that were intermediate in magnitude between those obtained here and those reported by the CWPDT (Brown and Brandes 1997).

It is important to note that C-SIGMETs (used for most of the comparisons reported here) and the NCWFP are actually very different types of forecasts, with different objectives. While the NCWFP provides an instantaneous snapshot, updated every 5 min, the C-SIGMETs are valid over a 2-h period. Moreover, the NCWFP is designed to function on areas of active convection that are expected to persist, and it does not focus on isolated convective cells. In contrast, the SIGMETs are expected to incorporate regions of isolated and developing convection (NWS 1991). Thus, the C-SIGMETs naturally would be expected to forecast over larger areas, in response to the time element and the need to include isolated cells that may develop during the valid period. The NCWFP does not attempt to forecast the development of new convective regions; rather, it focuses on frequent updates of the movement of existing convective areas.

In conclusion, the results presented here suggest that the NCWFP forecasts have positive skill, and that this skill is comparable to the skill associated with other types of convective forecasts (e.g., C-SIGMETs). PODy values for the NCWFP forecasts are relatively small, but these values increase when small convective areas are removed from the evaluation via the observation filter. Daily statistics show some variability, and the statistics appear to be better in cases when the convection is well organized. Positive attributes of the NCWFP, in comparison to other convective forecasts, include a relatively small FAR and small areal coverage by the forecasts. Because of the nature of the NCWFP, the Bias values obtained in this evaluation suggest that the NCWFP generally underforecasts convective areas. This underforecasting results from the fact that the NCWFP is an extrapolation algorithm, without the ability to "grow" or initiate convection. Finally, differences in the nature and objectives of the NCWFP and the C-SIGMETs suggest that these two products could be used in a complementary manner, and that the NCWFP would be able to provide informative forecasts to users who are interested in anticipating the movement of existing convective regions.

Acknowledgments

This research is in response to requirements and funding by the Federal Aviation Administration. The views expressed are those of the authors and do not necessarily represent the official policy and position of the U.S. Government.

We would like to express our appreciation to Joan Hart (CIRES/FSL) for her dedication to this project and her help in obtaining the necessary data, tables, and figures, and Chris Fischer (CIRES/FSL) for assisting with many of the diagrams and figures.

References

Brown, B.G., and E. Brandes, 1997: An intercommparison of 2-D storm-motion extrapolation algorithms. *Preprints, 28th Conference on Radar Meteorology*, 7-12 September, Austin TX, American Meteorological Society (Boston), 495-496.

Brown, B.G., G. Thompson, R.T. Bruintjes, R. Bullock, and T. Kane, 1997: Intercomparison of in-flight icing algorithms. Part II: Statistical verification results. *Wea. and Forec*., **12,** 890-914.

Convective Weather PDT, 1999: NCAR Convective Weather FY99 End-of-Year Report. Available from Cindy Mueller, Research Applications Program, National Center for Atmospheric Research.

Doswell, C.A., R. Davies-Jones, and David L. Keller, 1990: On summary measures of skill in rare event forecasting based on contingency tables. *Wea. and Forec.*, **5**, 576-585.

Mahoney, J.L., B.G. Brown, and J. Hart, 2000: Statistical Verification Results for the Collaborative Convective Forecast Product. NOAA Technical Report OAR 457-FSL 6, U.S. Department of Commerce, National Oceanic and Atmospheric Administration, Forecast Systems Laboratory, 30 pp.

Mahoney, J.L., J.K. Henderson, and P.A. Miller, 1997: A description of the Forecast Systems Laboratory's Real-Time Verification System (RTVS). *Preprints, 7th Conference on Aviation, Range, and Aerospace Meteorology*, Long Beach, American Meteorological Society, J26-J31.

Marzban, C., 1998: Scalar measures of performance in rare-event situations. *Weather and Forecasting*, **13**, 753-763.

Mueller, C.K., C.B. Fidalgo, D.W. McCann, D. Meganhart, N. Rehak, and T. Carty, 1999: National Convective Weather Forecast Product. *Preprints, 8th Conference on Aviation Range, and Aerospace Meteorology*, American Meteorological Society (Boston), 230-234.

NWS, 1991: National Weather Service Operations Manual, D-22. National Weather Service. (Available at this web sight http://www.nws.noaa.gov).

Orville, R.E., 1991: Lightning ground flash density in the contiguous United States-1989. *Mon. Wea. Rev*., **119**, 573-577.

Schaefer, J.T., 1990: The Critical Success Index as an indicator of warning skill. *Wea. and Forec*., **5**, 570-575.

Wilks, D.S., 1995: *Statistical Methods in the Atmospheric Sciences*. Academic Press, 467 pp.

Appendix A

The tables in this appendix show the hourly results for the NCWFP and the C-SIGMETs, as a supplement to the tables presented in Section 5.2.

Table A1. Overall results by issue time for the 1-h NCWFP forecasts with verification based on the unfiltered observations.

Issue time (UTC)	Fcst. length (h)	Valid time (UTC)	PODy	PODn	FAR	CSI	TSS	HSS	GSS	Bias	% Area
0	1	1	0.14	1.00	0.32	0.13	0.14	0.22	0.13	0.21	0.65
1	1	2	0.14	1.00	0.32	0.13	0.13	0.22	0.12	0.20	0.53
2	1	3	0.13	1.00	0.29	0.13	0.13	0.22	0.12	0.19	0.43
3	1	4	0.13	1.00	0.26	0.13	0.13	0.22	0.13	0.18	0.37
4	1	5	0.13	1.00	0.25	0.13	0.13	0.22	0.12	0.18	0.33
5	1	6	0.13	1.00	0.24	0.12	0.12	0.21	0.12	0.17	0.29
6	1	7	0.12	1.00	0.25	0.11	0.12	0.20	0.11	0.15	0.26
7	1	8	0.10	1.00	0.24	0.10	0.10	0.17	0.10	0.13	0.22
8	1	9	0.10	1.00	0.24	0.09	0.10	0.17	0.09	0.13	0.21
9	1	10	0.10	1.00	0.24	0.09	0.10	0.17	0.09	0.13	0.21
10	1	11	0.09	1.00	0.26	0.09	0.09	0.15	0.08	0.12	0.21
11	1	12	0.09	1.00	0.26	0.09	0.09	0.16	0.09	0.13	0.21
12	1	13	0.08	1.00	0.25	0.08	0.08	0.15	0.08	0.11	0.19
13	1	14	0.08	1.00	0.25	0.08	0.08	0.15	0.08	0.11	0.17
14	1	15	0.06	1.00	0.29	0.06	0.06	0.11	0.06	0.08	0.13
15	1	16	0.05	1.00	0.30	0.05	0.05	0.09	0.05	0.07	0.12
16	1	17	0.06	1.00	0.24	0.06	0.06	0.10	0.05	0.08	0.14
17	1	18	0.06	1.00	0.26	0.06	0.06	0.10	0.05	0.08	0.19
18	1	19	0.07	1.00	0.27	0.07	0.07	0.12	0.06	0.10	0.29
19	1	20	0.09	1.00	0.32	0.08	0.09	0.15	0.08	0.13	0.47
20	1	21	0.1	1.00	0.37	0.09	0.09	0.16	0.09	0.15	0.56
21	1	22	0.11	1.00	0.33	0.10	0.10	0.17	0.09	0.16	0.60
22	1	23	0.12	1.00	0.34	0.11	0.12	0.19	0.11	0.18	0.72
23	1	0	0.14	1.00	0.35	0.13	0.14	0.22	0.12	0.21	0.74
Total			0.13	1.00	0.32	0.12	0.13	0.21	0.12	0.19	0.51

Table A2. Overall results by issue time for the 1-h C-SIGMETs with verification based on the unfiltered observations.

Issue time (UTC)	Fcst. length (h)	Valid time (UTC)	PODy	PODn	FAR	CSI	TSS	HSS	GSS	Bias	% Area
0	1	1	0.38	0.97	0.71	0.20	0.35	0.31	0.18	1.31	3.95
1	1	2	0.36	0.98	0.70	0.19	0.34	0.31	0.18	1.21	3.15
2	1	3	0.34	0.98	0.67	0.20	0.32	0.32	0.19	1.02	2.27
3	1	4	0.32	0.99	0.65	0.20	0.31	0.32	0.19	0.91	1.85
4	1	5	0.32	0.99	0.62	0.21	0.31	0.34	0.20	0.85	1.60
5	1	6	0.32	0.99	0.62	0.21	0.31	0.34	0.20	0.82	1.44
6	1	7	0.33	0.99	0.60	0.22	0.32	0.35	0.21	0.80	1.36
7	1	8	0.32	0.99	0.60	0.21	0.31	0.34	0.21	0.79	1.30
8	1	9	0.30	0.99	0.58	0.21	0.30	0.34	0.21	0.72	1.18
9	1	10	0.29	0.99	0.57	0.21	0.28	0.34	0.20	0.67	1.12
10	1	11	0.29	0.99	0.60	0.20	0.28	0.33	0.20	0.73	1.20
11	1	12	0.27	0.99	0.61	0.19	0.26	0.31	0.18	0.69	1.13
12	1	13	0.26	0.99	0.64	0.18	0.25	0.29	0.17	0.73	1.20
13	1	14	0.25	0.99	0.67	0.17	0.24	0.28	0.16	0.76	1.14
14	1	15	0.25	0.99	0.68	0.16	0.24	0.27	0.16	0.77	1.11
15	1	16	0.23	0.99	0.67	0.16	0.23	0.26	0.15	0.72	1.09
16	1	17	0.24	0.99	0.63	0.17	0.24	0.28	0.16	0.67	1.20
17	1	18	0.26	0.99	0.62	0.19	0.25	0.30	0.18	0.69	1.61
18	1	19	0.29	0.98	0.63	0.19	0.28	0.31	0.18	0.80	2.40
19	1	20	0.34	0.98	0.65	0.21	0.32	0.33	0.19	0.98	3.39
20	1	21	0.38	0.97	0.69	0.21	0.35	0.31	0.19	1.21	4.44
21	1	22	0.39	0.96	0.70	0.21	0.36	0.31	0.18	1.31	5.05
22	1	23	0.39	0.96	0.71	0.20	0.35	0.30	0.18	1.33	5.13
23	1	0	0.40	0.96	0.72	0.19	0.36	0.30	0.17	1.44	4.83
Totals	1		0.33	0.98	0.67	0.20	0.31	0.31	0.19	0.99	2.30

Table A3. Overall results by issue time for the 2-h NCWFP forecasts with verification based on the unfiltered observations.

Issue time (UTC)	Fcst. length (h)	Valid time (UTC)	PODy	PODn	FAR	CSI	TSS	HSS	GSS	Bias	% Area
0	2	2	0.07	1.00	0.49	0.07	0.07	0.12	0.06	0.14	0.37
1	2	3	0.07	1.00	0.50	0.07	0.07	0.12	0.07	0.14	0.33
2	2	4	0.07	1.00	0.43	0.07	0.07	0.13	0.07	0.13	0.28
3	2	5	0.06	1.00	0.47	0.06	0.06	0.11	0.06	0.12	0.23
4	2	6	0.07	1.00	0.48	0.07	0.07	0.12	0.06	0.13	0.23
5	2	7	0.07	1.00	0.41	0.07	0.07	0.12	0.06	0.12	0.20
6	2	8	0.06	1.00	0.40	0.06	0.06	0.11	0.06	0.11	0.18
7	2	9	0.06	1.00	0.40	0.06	0.06	0.11	0.06	0.10	0.16
8	2	10	0.05	1.00	0.37	0.05	0.05	0.10	0.05	0.09	0.14
9	2	11	0.05	1.00	0.42	0.05	0.05	0.09	0.05	0.09	0.15
10	2	12	0.04	1.00	0.43	0.04	0.04	0.08	0.04	0.08	0.13
11	2	13	0.05	1.00	0.47	0.05	0.05	0.08	0.04	0.09	0.15
12	2	14	0.05	1.00	0.45	0.05	0.05	0.08	0.04	0.09	0.13
13	2	15	0.04	1.00	0.46	0.04	0.04	0.07	0.04	0.08	0.11
14	2	16	0.03	1.00	0.47	0.02	0.02	0.05	0.02	0.05	0.08
15	2	17	0.02	1.00	0.45	0.02	0.02	0.04	0.02	0.04	0.08
16	2	18	0.02	1.00	0.51	0.02	0.02	0.04	0.02	0.04	0.10
17	2	19	0.02	1.00	0.42	0.02	0.02	0.03	0.02	0.03	0.10
18	2	20	0.02	1.00	0.46	0.02	0.02	0.04	0.02	0.04	0.16
19	2	21	0.03	1.00	0.52	0.03	0.03	0.06	0.03	0.07	0.26
20	2	22	0.04	1.00	0.48	0.04	0.04	0.07	0.03	0.07	0.29
21	2	23	0.04	1.00	0.50	0.04	0.04	0.07	0.04	0.08	0.32
22	2	0	0.06	1.00	0.53	0.05	0.06	0.10	0.05	0.12	0.43
23	2	1	0.07	1.00	0.51	0.06	0.07	0.11	0.06	0.14	0.43
Total			0.07	1.00	0.50	0.06	0.06	0.11	0.06	0.13	0.31

Table A4. Overall results by issue time for the 2-h C-SIGMETs with verification based on the unfiltered observations.

Issue time (UTC)	Fcst. length (h)	Valid time (UTC)	PODy	PODn	FAR	CSI	TSS	HSS	GSS	Bias	% Area
0	2	2	0.31	0.97	0.79	0.14	0.28	0.23	0.13	1.52	3.95
1	2	3	0.29	0.97	0.80	0.14	0.26	0.22	0.12	1.42	3.17
2	2	4	0.26	0.98	0.76	0.14	0.25	0.23	0.13	1.12	2.27
3	2	5	0.24	0.99	0.75	0.14	0.23	0.23	0.13	0.97	1.82
4	2	6	0.26	0.99	0.72	0.16	0.25	0.26	0.15	0.92	1.60
5	2	7	0.24	0.99	0.70	0.16	0.23	0.26	0.15	0.81	1.37
6	2	8	0.26	0.99	0.69	0.16	0.25	0.27	0.16	0.85	1.40
7	2	9	0.23	0.99	0.70	0.15	0.22	0.25	0.14	0.79	1.29
8	2	10	0.23	0.99	0.68	0.15	0.22	0.26	0.15	0.71	1.18
9	2	11	0.22	0.99	0.68	0.15	0.21	0.25	0.14	0.68	1.13
10	2	12	0.22	0.99	0.69	0.15	0.22	0.25	0.14	0.72	1.19
11	2	13	0.20	0.99	0.71	0.13	0.19	0.23	0.13	0.69	1.12
12	2	14	0.21	0.99	0.74	0.13	0.20	0.22	0.12	0.80	1.19
13	2	15	0.20	0.99	0.75	0.12	0.19	0.21	0.12	0.81	1.17
14	2	16	0.17	0.99	0.76	0.11	0.16	0.19	0.10	0.71	1.08
15	2	17	0.16	0.99	0.74	0.11	0.15	0.19	0.10	0.62	1.12
16	2	18	0.15	0.99	0.70	0.11	0.14	0.19	0.10	0.50	1.17
17	2	19	0.18	0.99	0.67	0.13	0.17	0.21	0.12	0.54	1.63
18	2	20	0.22	0.98	0.69	0.15	0.20	0.23	0.13	0.69	2.39
19	2	21	0.26	0.97	0.72	0.15	0.23	0.24	0.14	0.93	3.41
20	2	22	0.29	0.97	0.74	0.16	0.26	0.24	0.14	1.14	4.41
21	2	23	0.31	0.96	0.77	0.15	0.27	0.23	0.13	1.32	5.09
22	2	0	0.32	0.96	0.79	0.14	0.28	0.22	0.12	1.53	5.15
23	2	1	0.33	0.96	0.80	0.14	0.29	0.22	0.13	1.60	4.85
Totals	2		0.25	0.98	0.75	0.14	0.23	0.23	0.13	0.99	2.29

Table A5. Overall results by issue time for the 2-h NCWFP forecasts with verification based on the filtered observations.

Issue time (UTC)	Fcst. length (h)	Valid time (UTC)	PODy	PODn	FAR	CSI	TSS	HSS	GSS	Bias	% Area
0	2	2	0.11	1.00	0.72	0.09	0.11	0.15	0.08	0.40	0.37
1	2	3	0.12	1.00	0.73	0.09	0.11	0.16	0.09	0.42	0.33
2	2	4	0.13	1.00	0.67	0.10	0.13	0.18	0.10	0.40	0.28
3	2	5	0.11	1.00	0.70	0.09	0.11	0.16	0.09	0.38	0.23
4	2	6	0.13	1.00	0.70	0.10	0.12	0.17	0.10	0.42	0.22
5	2	7	0.13	1.00	0.64	0.11	0.13	0.19	0.11	0.37	0.20
6	2	8	0.13	1.00	0.64	0.10	0.12	0.18	0.10	0.35	0.18
7	2	9	0.12	1.00	0.65	0.10	0.12	0.17	0.09	0.34	0.16
8	2	10	0.11	1.00	0.62	0.09	0.11	0.17	0.09	0.29	0.14
9	2	11	0.11	1.00	0.65	0.09	0.11	0.16	0.09	0.30	0.15
10	2	12	0.09	1.00	0.67	0.07	0.09	0.13	0.07	0.26	0.13
11	2	13	0.10	1.00	0.69	0.08	0.10	0.15	0.08	0.32	0.15
12	2	14	0.10	1.00	0.70	0.08	0.10	0.14	0.08	0.33	0.13
13	2	15	0.09	1.00	0.71	0.07	0.09	0.13	0.07	0.30	0.11
14	2	16	0.06	1.00	0.72	0.05	0.05	0.09	0.05	0.20	0.08
15	2	17	0.05	1.00	0.69	0.04	0.05	0.08	0.04	0.15	0.08
16	2	18	0.04	1.00	0.73	0.04	0.04	0.07	0.03	0.15	0.10
17	2	19	0.03	1.00	0.66	0.03	0.03	0.06	0.03	0.10	0.10
18	2	20	0.04	1.00	0.69	0.04	0.04	0.07	0.04	0.14	0.16
19	2	21	0.06	1.00	0.71	0.05	0.06	0.09	0.05	0.21	0.26
20	2	22	0.06	1.00	0.72	0.06	0.06	0.10	0.05	0.23	0.29
21	2	23	0.07	1.00	0.72	0.06	0.07	0.10	0.06	0.24	0.32
22	2	0	0.09	1.00	0.75	0.07	0.08	0.12	0.07	0.36	0.43
23	2	1	0.10	1.00	0.74	0.08	0.10	0.14	0.08	0.39	0.42
Total			0.10	1.00	0.74	0.08	0.10	0.14	0.08	0.39	0.42

Table A6. Overall results by issue time for the 2-h C-SIGMETs with verification based on the filtered observations.

Issue time (UTC)	Fcst. length (h)	Valid time (UTC)	PODy	PODn	FAR	CSI	TSS	HSS	GSS	Bias	% Area
0	2	2	0.37	0.96	0.91	0.07	0.33	0.13	0.07	4.29	3.92
1	2	3	0.35	0.97	0.92	0.07	0.32	0.13	0.07	4.15	3.14
2	2	4	0.34	0.98	0.90	0.09	0.32	0.15	0.08	3.34	2.25
3	2	5	0.31	0.98	0.90	0.08	0.30	0.15	0.08	2.98	1.81
4	2	6	0.36	0.99	0.87	0.10	0.35	0.18	0.10	2.82	1.60
5	2	7	0.34	0.99	0.87	0.10	0.32	0.18	0.10	2.54	1.37
6	2	8	0.36	0.99	0.87	0.11	0.35	0.19	0.10	2.77	1.40
7	2	9	0.33	0.99	0.87	0.10	0.32	0.18	0.10	2.62	1.29
8	2	10	0.32	0.99	0.86	0.11	0.31	0.19	0.10	2.35	1.18
9	2	11	0.31	0.99	0.86	0.11	0.30	0.19	0.10	2.28	1.13
10	2	12	0.32	0.99	0.87	0.10	0.31	0.18	0.10	2.49	1.19
11	2	13	0.28	0.99	0.89	0.09	0.27	0.15	0.08	2.48	1.12
12	2	14	0.31	0.99	0.90	0.08	0.30	0.15	0.08	3.02	1.19
13	2	15	0.29	0.99	0.91	0.07	0.28	0.13	0.07	3.14	1.17
14	2	16	0.25	0.99	0.91	0.07	0.24	0.12	0.07	2.82	1.10
15	2	17	0.24	0.99	0.90	0.08	0.23	0.14	0.07	2.31	1.13
16	2	18	0.22	0.99	0.87	0.09	0.21	0.15	0.08	1.73	1.17
17	2	19	0.25	0.99	0.85	0.10	0.24	0.17	0.10	1.72	1.64
18	2	20	0.29	0.98	0.87	0.10	0.27	0.17	0.09	2.17	2.39
19	2	21	0.32	0.97	0.89	0.09	0.29	0.15	0.08	2.83	3.42
20	2	22	0.35	0.96	0.90	0.08	0.31	0.14	0.07	3.52	4.43
21	2	23	0.36	0.95	0.90	0.08	0.31	0.13	0.07	3.73	5.02
22	2	0	0.37	0.95	0.92	0.07	0.32	0.12	0.06	4.52	5.15
23	2	1	0.38	0.96	0.92	0.07	0.33	0.12	0.06	4.53	4.84
Totals	2		0.32	0.98	0.90	0.09	0.30	0.15	0.08	3.12	2.29