P11.4 **COMPARISONS AND VERIFICATION OF AN AUTOMATED THUNDERSTORM POTENTIAL INDEX OUTPUT TO MANUAL PRODUCTS**

David Knapp *
U.S. Army Research Laboratory, White Sands Missile Range, New Mexico

Earl Barker
Harris Technical Services Corporation, Bellevue, Nebraska

Gordon Brooks
Scott Rentschler
Air Force Weather Agency, Offutt Air Force Base, Nebraska

1. INTRODUCTION

Research initiated in the 1990s focused on the development of a Thunderstorm Potential Index (TPI), which could be used to predict percent probabilities of local thunderstorm occurrence within a 100-km radius of a RAOB location for a 12-hour period (Knapp and Passner, 1998). The TPI was developed as an application for use in data-sparse and data-denied regions where a single RAOB might be the only data source in the local area. Using data collected at selected RAOB locations to represent the diversity of climatological areas across the continental U.S., a perfect prog approach was used to derive the TPI as a multivariate linear regression equation.

The TPI was applied to synoptic and mesoscale model data in an attempt to determine if the algorithm would be useful as a guidance product for forecasters producing graphical thunderstorm outlook areas valid across the continental U.S. Knapp and Brooks (2000), in their study comparing and verifying TPI output derived from the NCAR/Penn State Mesoscale Model (MM5) 36km resolution output at the Air Force Weather Agency (AFWA) to the Military Weather Advisory (MWA) output, describe the TPI as a product similar to the Convective Outlook produced by the NWS' Storm Prediction Center (SPC). Verification results proved the TPI produced accurate first guess non-severe thunderstorm forecasts comparable to military forecaster-derived products. However, weaknesses in the TPI were discovered for high elevation regions, and the TPI was found to overforecast convection during the Southwest U.S. monsoon and summer convection season.

Due to the shift of regional thunderstorm forecasting responsibilities from AFWA to the Air Force's Operational Weather Squadrons (OWS), the MWA was eliminated from production at AFWA. Thus, the original TPI and a high elevation version could not be further verified against the MWA. AFWA has incorporated the TPI, including the high elevation algorithm, into their MM5 precipitation type products to display areas of thunderstorm and severe thunderstorm potential (Brooks et al., 2004).

With the MWA no longer available, the SPC's Convective Outlook was used as the current forecast standard for a comparison and verification test against the TPI during the March-September 2005 convective season across the continental U.S. (CONUS). The purpose of this study was twofold:

1. To prove the TPI could add value as an automated first-guess tool to an accepted "standard" forecast product (the SPC Convective Outlook) and to other forecaster-derived thunderstorm products.
2. To determine if the TPI could be used as a stand-alone tool for forecasting non-severe and severe thunderstorms across a region in 12-24 hour forecast periods.

2. TPI EQUATIONS

The original TPI equation was derived based on three stability indices—K, Lifted Index (LI), and Severe Weather Threat (SWEAT)—out of 10 indices studied, because they were shown to be the most statistically significant in determining correct yes/no forecasts of thunderstorm occurrence during the 12-hour

* *Corresponding author address:* David Knapp, U.S. Army Research Laboratory, ATTN: AMSRD-ARL-CI-E, White Sands Missile Range, NM 88002-5501; email: dknapp@arl.army.mil.

period following RAOB report times (Knapp and Passner, 1998). The following equation, calculating a percent probability of occurrence, was applied to model grid point data:

$$TPI = [.1795 + .0073(K) - .0149(LI) + .0008(SWEAT)] \times 100 \quad (1)$$

A second equation, for model grid point locations with surface elevations higher than 850 mb, was derived using stability indices and other RAOB parameters for selected locations in the U.S. Rocky Mountains. For this perfect prog-derived equation, the LI and Precipitable Water (PW), in inches, correlated best to local area thunderstorm occurrence. This high elevation equation is expressed as follows:

$$TPI\ 850 = [.2101 + .7611(PW) - .054(LI)] \times 100 \quad (2)$$

For application of the TPI to forecast model data, equations (1) and (2) are applied, dependent upon the surface pressure calculated by the model at each grid point. For the purposes of this paper, the term TPI is used to describe the algorithm output at all model grid points, regardless of which equation was used to calculate the probability of thunderstorm occurrence.

During work leading to the original derivation of the TPI, the threshold value of 47% was determined to produce the best verification statistics when making simple yes/no thunderstorm occurrence predictions. This threshold value was determined based on verification for an independent set of RAOB data from across the U.S. Graphical depictions of the TPI output at AFWA using MM5 36-km gridded data to produce the TPI charts showed good agreement between the 73% TPI contour and the manually produced outlines of severe thunderstorms on the MWA charts. Thus, these two thresholds are used as the critical values for TPI output.

3. TPI COMPARISONS TO MANUAL PRODUCTS

3.1 *SPC and TPI Products*

The SPC's Day 1 and Day 2 Convective Outlook 24-hour forecast products were compared to corresponding TPI products created by the National Centers for Environmental Prediction North American Mesoscale (NAM) model 40-km output grids over the CONUS from March to September 2005. Only the initial Day 1 and Day 2 SPC products were used; daily updates/amendments to these products were not included in the comparison. TPI grids were created using NAM data, which would be available to SPC forecasters when creating their initial products. Typically, the Day 1 SPC Outlook data are distributed no later than 0600 UTC each day, valid for the 24-hour period starting at 1200 UTC. Thus, the 0000 UTC NAM forecasts were used to create the TPI grids.

Since the TPI was originally derived as a 12-hour thunderstorm forecast product and the SPC Outlook is a 24-hour product, TPI grids from more than one NAM valid time within a particular model run were meshed together to produce one TPI chart valid for 24 hours. Since NAM grids were available in 3-hour incremental valid times, TPI output from five valid times bounded within the SPC Outlook valid time for each NAM run were analyzed, and the highest TPI value was extracted at each grid point. Thus, the TPI was produced at each NAM grid point from the daily 0000 UTC model run at valid times of 1200, 1500, 1800, 2100, and 0000 UTC. Capturing the maximum TPI values at each grid point at the interim model valid times was important. Similar model data manipulations were performed to create the TPI output valid for the corresponding Day 2 Convective Outlook valid times, with 36-48 hour grids in 3-hour increments used from the same Day 1 0000 UTC NAM model run.

3.2 *SPC Outlook Image Processing*

Images of SPC's Day 1 Convective Outlook graphic were created from corresponding text bulletins disseminated by the National Weather Service. Using the latitude/longitude points defined in the bulletins, an automated Perl script parsed the Slight, Moderate, and High risk severe weather areas along with the general thunderstorm risk areas. The script drew lined outlook areas on a base map image created from the NAM 40-km grid for straightforward comparisons to corresponding TPI output from the NAM. The image of the SPC Convective Outlook area generated by the Perl script was then manually color-filled using Paint Shop Pro. Examples from 16 June 2005 showing the original graphic SPC Outlook and the corresponding comparison image for this study are shown in Figs. 1 and 2, respectively.

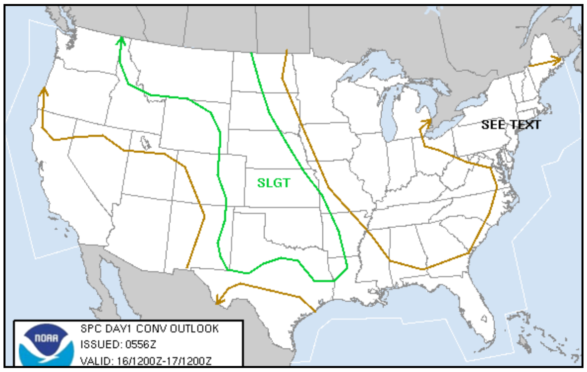

Fig.1. SPC Day 1 Convective Outlook forecast issued at 0556 UTC 16 June 2005 (from www.spc.noaa.gov). The forecast is valid from 16 June 2005 1200 UTC to 17 June 2005 1200 UTC.

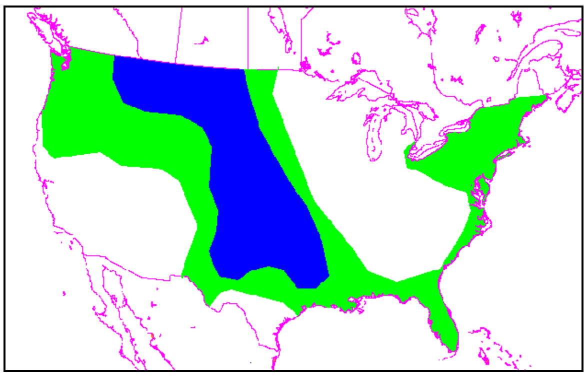

Fig. 2. Color-filled SPC Day 1 Convective Outlook from Fig. 1.

3.3 *TPI Image Processing*

The TPI grids, produced from the NAM output as gridded binary messages, were processed to generate TPI raster images to simulate the Day 1 and Day 2 Convective Outlooks. A cubic spline approach was used to create image pixels in 8-km resolution from the NAM 40-km source grids. The resultant images were color-coded so that pixel values of $\geq$73 (severe) were blue and values of $\geq$47 (general thunderstorms) were orange. A sample is shown in Fig. 3.

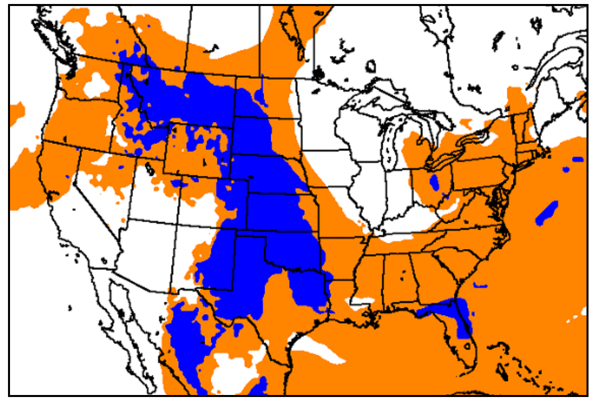

Fig. 3. TPI output from the 0000 UTC 16 June 2005 NAM model run. The valid time is the same as Fig. 1.

3.4 *Image Comparisons, Data Collection, and Statistics*

Data comparisons between the SPC Convective Outlook images and the corresponding TPI images (seen in Figs. 2 and 3, respectively) were accomplished using an automated Perl script that compared 8-km resolution pixel values between the product images. Pixels of "agreement" and "disagreement" were determined and then assigned to one of nine bins in the 3-by-3 matrix shown in Table 1.

Table 1. Contingency table showing bins comparing the SPC Outlook versus the TPI forecast for three categories of thunderstorm areas.

		TPI None	TPI Non-Severe	TPI Severe
SPC	None	a	b	c
SPC	Non-Severe	d	e	f
SPC	Severe	g	h	i

In addition to the data compiled in Table 1, a color-coded image graphically showing the bin totals was created for each comparison. An image based on comparing Figs. 2 and 3 is shown in Fig. 4. As listed in Table 1, the bins showing agreement between the SPC Outlook and the TPI are a, e, and i. These letters correspond to the white, orange, and dark blue areas, respectively, in Fig. 4. The other Table 1 bin letters that correspond to colors in Fig 4 are as follows: b = gray, c = black, d = red, f = yellow, g = light blue (none shown), and h = medium blue.

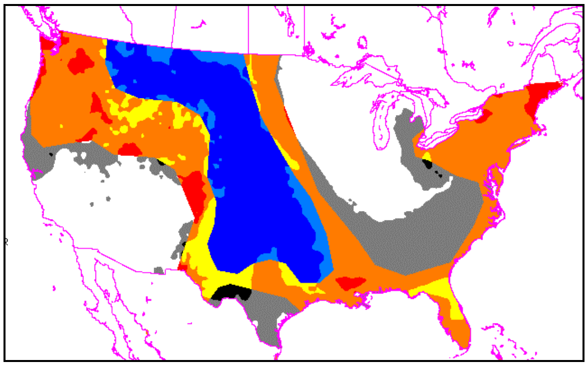

Fig 4. Color shading areas of SPC and TPI forecast agreement/disagreement derived from Figs. 2 and 3.

Using the data presented in Fig. 4, the pixel counts can be parsed into the bins shown in Table 1, resulting in the following totals for each bin:

a = 33091
b = 18325
c = 711
d = 4110
e = 33809
f = 6672
g = 0
h = 4248
i = 23600

From these bin totals, the following statistics were calculated showing the percent agreement between the SPC Outlook and corresponding TPI forecasts:

% Agree, All = (a+e+i)/(sum of all bins) x 100 = 73%
% Agree, None = (a)/(a+b+c+d+g) x 100 = 59%
% Agree, Non-Severe = (e)/(d+e+f+b+h) x 100 = 50%
% Agree, Severe = (i)/(g+h+i+c+f) x 100 = 67%

Comparisons of convection versus no convection forecast agreement can be extracted from the data in Table 1. These data were compiled into a 2-by-2 matrix, shown in Table 2.

Table 2. Same as Table 1, but for two categories of thunderstorm forecasts.

		TPI None	TPI Convection
SPC	None	a	b
	Convection	c	d

Using a similar method to the method used for the 3-by-3 table in Table 1, the following bin totals from Table 2 were taken from Fig. 4:

a = 33091, b = 19036, c = 4110, d = 68329

From these data, the "% Agree, Convection versus No Convection" statistic is calculated as

(a+d)/(a+b+c+d) x 100 = 81%

Percent Agree statistics were calculated for all the data collected during the period of this study. SPC Outlook versus TPI comparisons of convection/no convection forecasts are summarized in Table 3. The correlation coefficients showing the strength of the agreement between the two products for Day 1 and Day 2 were .65 and .61, respectively. Percent agreement was highest in March and lowest in September for the Day 1 charts. For Day 2, April showed the highest percent agreement, while September showed the lowest. The high average percent agreement values for both the Day 1 and Day 2 charts, combined with the high correlation coefficients, lead to the conclusion that the TPI does a good job of providing first-guess estimates of the SPC's general thunderstorm area contours throughout the convective season.

Table 3. Percent agreement between comparable SPC Outlook and TPI charts for areas of convection and no convection.

	Total	Mar	Apr	May	Jun	Jul	Aug	Sep
Day 1 Charts	82.2	86.3	83.7	82.0	80.9	82.4	83.6	77.5
Day 2 Charts	80.0	82.1	83.7	81.7	80.7	78.2	81.9	72.9

Statistics comparing the categorical breakout of percent agreement between the SPC Outlooks and the TPI for Day 1 and Day 2 products are presented in Table 4. When looking at more detail in the categorical forecasts, the TPI compared favorably overall with the SPC Outlooks through the spring season, with a drop off in percent agreement during the summer months. There is evidence of only a slight drop off in agreement between the two products from Day 1 to Day 2 across all categories. The Severe and Non-Severe statistics show significantly lower agreement between the areas predicted by both products, with the best agreement occurring in May for both Day 1 and Day 2 for the Non-Severe category. Agreement is strongest in June for Severe thunderstorm forecast areas in both Day 1 and Day 2 statistics.

Table 4. Percent agreement between comparable SPC Outlook and TPI charts for all categories of convection forecasts.

	Total	Mar	Apr	May	Jun	Jul	Aug	Sep
% Agree, Day 1 Charts								
All Categories	70.4	83.4	78.4	74.1	66.4	62.1	63.2	67.0
No Convection	68.9	84.4	79.1	69.2	53.8	53.5	55.0	64.8
Non-Severe	48.1	32.8	39.9	52.9	51.7	49.9	52.5	41.7
Severe	26.1	26.7	32.1	37.7	38.5	21.0	14.3	20.1
% Agree, Day 2 Charts								
All Categories	67.7	78.7	79.3	73.8	65.8	58.0	62.6	62.7
No Convection	64.5	79.0	79.6	68.4	52.8	46.6	55.3	60.5
Non-Severe	46.8	32.0	39.5	53.7	51.5	47.0	50.9	34.9
Severe	23.0	30.9	35.3	34.3	36.2	14.8	12.6	15.6

4. PRODUCT VERIFICATION VERSUS LIGHTNING AND SEVERE WEATHER REPORTS

4.1 *Data Collection and Setup*

The TPI and SPC Outlook charts were verified independently against archived lightning strike data for general thunderstorm areas and against severe weather reports as tabulated by the SPC. Lightning strike data were acquired for this study through the Air Force Combat Climatology Center, which archives the data recorded by the National Lightning Detection Network. Verification data were collected from April to September 2005, with the April-August data compiled as of this writing. The verification process was accomplished by binning the severe weather reports and lightning strikes to 40-km grid boxes that match the resolution of the model data used to create the TPI charts. Wherever a lightning strike occurred, the corresponding 40-km grid box was classified as a “hit” for a convection report. The same approach was used for classifying severe weather “hit” grid boxes when such a report was recorded.

4.2 *Verification Statistics*

Forecasts for both the SPC Outlook and TPI charts were compared to the lightning and severe weather reports, with correct and incorrect forecast areas divided into 40-km grid boxes and then counted for the statistical calculations. The grid box totals were then binned according to the 3-by-3 matrix, as seen in Table 5.

Table 5. Continuity table showing SPC Outlook or TPI forecast grid boxes compared to recorded thunderstorm severity and lightning reports.

		Lightning and Severe Reports		
		None	**Non-Severe**	**Severe**
SPC or TPI	**None**	a	b	c
	Non-Severe	d	e	f
	Severe	g	h	i

Again, a color-coded image graphically showing the bin totals was created for each product’s verification against the lightning and severe reports. Figs. 5 and 6 feature images showing the performance of the SPC Outlook and the corresponding TPI chart, respectively, as verified against the reports. The color coding in Figs. 5 and 6 and the corresponding bin letters in Table 5 are the same as those used in Table 1 and Fig. 4 in section 3.3. In Fig. 5, the percent correct forecast grid boxes for the SPC Outlook was 54%. For the TPI, shown in Fig. 6, the percent correct forecast was 41%.

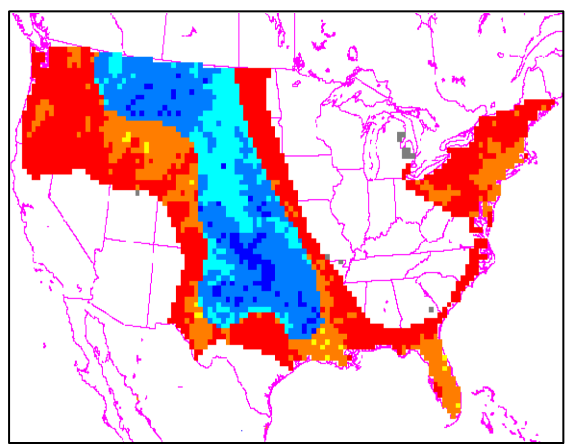

Fig. 5. Performance of the SPC Outlook versus verifying reports for the 16 June 2005 1200 UTC chart.

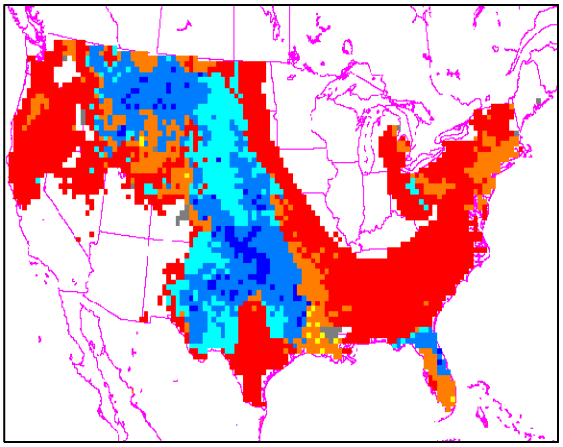

Fig 6. Performance of the TPI versus verifying reports for the 16 June 2005 1200 UTC chart.

Verification using raw reports, as were used in this case, was accomplished for each of the SPC Outlook and corresponding TPI charts. Thus far, only the Day 1 product statistics have been compiled. These statistics are presented in Table 6. Overall, the SPC Outlooks performed 10% better than the TPI. The TPI verified best, and verified better than the SPC Outlooks in April, with the two products separated by no more than 4% verification accuracy through the spring. The summer convection season showed the overforecasting tendency in the TPI as compared to the SPC Outlooks. This was probably due to the NAM's inability to consistently predict the vertical profile moisture data used in the TPI calculations as accurately as the SPC forecasters who have access to current evolving data and model tendencies. This TPI overforecasting confirmed similar findings as presented by Knapp and Brooks (2000), where MM5 data were used to calculate the TPI. The TPI overforecasting issue is being addressed by adding convective capping and other filters to the TPI output in an attempt to bring the forecast errors under control.

Table 6. Day 1 percent correct forecasts.

	Total	Apr	May	Jun	Jul	Aug
TPI Forecasts	52.4	77.5	58.5	42.8	41.2	42.2
SPC Outlooks	62.7	72.8	62.4	55.7	60.9	63.1

5. AIR FORCE OWS THUNDERSTORM FORECASTS

The nature of the TPI first-guess thunderstorm tool research and the reasoning behind comparing the TPI and SPC forecasts has parallel applications for current Air Force weather forecasting operations. In these challenging days, there are fewer Air Force forecasters producing an increasing number of forecaster-in-the-loop (FITL) products to support Department of Defense and humanitarian missions worldwide. Forecasters need access to quality first-look model products. Accordingly, AFWA has leveraged off the TPI-SPC Outlook comparisons to complete a very similar study using OWS thunderstorm forecast charts.

There are eight OWSs or "forecast production centers" in AFWA that provide worldwide FITL hazards charts, which include parameters such as thunderstorms, icing, and turbulence. For this study, "fused" thunderstorm forecast charts for the entire CONUS were created from four corresponding OWS regional charts. Because these fused FITL charts were point-in-time forecasts, one 24-hour composite OWS thunderstorm forecast chart was produced from the four projection times available within the appropriate 24-hour period corresponding to each SPC Outlook and TPI chart produced (Fig. 7). The OWS thunderstorm products include thunderstorm coverage (i.e., isolated, few, scattered, numerous), but do not attempt to forecast severity. Accordingly, the lightning data archive will be used to verify and compare the meshed TPI and composite OWS forecasts. The TPI and OWS comparison and statistical analysis work was still in progress at the time of this writing.

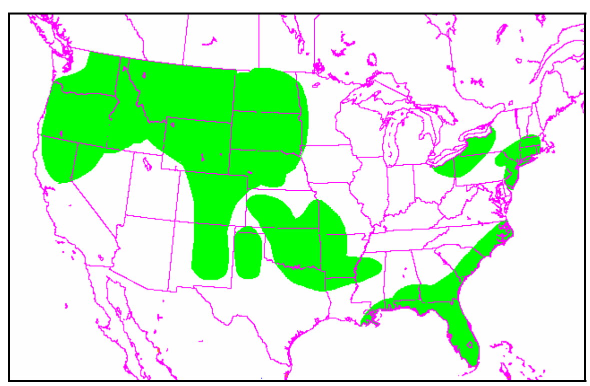

Fig. 7. An example of a 24-hour composite OWS thunderstorm forecast chart, with a valid time the same as Fig. 1. Green indicates forecast coverage that featured "isolated or few thunderstorms."

6. CONTINUING WORK

In addition to the OWS work described in section 5, TPI and SPC Outlook verification against reported lighting and severe weather observations will be expanded to include such statistics as Probability of Detection, False Alarm Rate, True Skill Statistic, Heidke Skill Score, and Bias for the forecast categories of No Convection, Non-Severe, Severe, and Convection/No Convection. Verification will be completed for the entire data collection period ending in September 2005. Selected months will be examined using the TPI calculated with a convective inhibition filter to see whether it reduces the overforecasting tendency evident during the summer convection season.

The TPI was also run using Rapid Update Cycle (RUC) model output during the 2005 verification period, but objective statistical data were not collected to complete comparisons. Subjective visual evaluations of the RUC TPI output indicated significantly less overforecasting than the TPI. The RUC TPI output will be compared against the SPC Outlooks in a similar study during 2006.

7. CONCLUSIONS

This study has proven that automatically model-derived TPI can add value as a first-guess tool to the accepted “standard” manually derived SPC Outlook. With an 80% agreement in determining between the Outlook and TPI convective boundary predictions for both Day 1 and Day 2 products, the TPI was shown to be a more than adequate first-guess tool. Verifying both against reported lightning and severe weather occurrences showed an overall 10% difference in accuracy between the two products. However, more work on this part of the verification needs to be completed before confident conclusions can be reached.

Can the TPI be used as a stand-alone tool for general and severe thunderstorm forecast guidance? Yes, for outlining general convection forecast areas. Statistics are still being worked on to determine if the TPI can produce severe thunderstorm forecast area boundaries that can be used with little or no forecaster modifications.

REFERENCES

Brooks, G.R., M.J. Noehrenberg, and D.I. Knapp, 2004: Development of a thunderstorm algorithm from very high resolution AFWA MM5 data. *11th Conf. on Aviation, Range, and Aerospace Meteorology*.

Knapp, D.I., and G.R. Brooks, 2000: Use of a new thunderstorm potential index for 12-hour forecasts using mesoscale model data. Preprint, *9th Conf. on Aviation, Range, and Aerospace Meteorology*, J54–58.

Knapp, D.I., and J.E. Passner, 1998: Development of a local 12-hour general thunderstorm forecasting technique for use with RAOB or mesoscale model vertical profile data. Preprint, *16th Conf. Wea. Analysis and Forecasting*, 115–117.