



# AMERICAN METEOROLOGICAL SOCIETY

*Journal of Atmospheric and Oceanic Technology*

## EARLY ONLINE RELEASE

This is a preliminary PDF of the author-produced manuscript that has been peer-reviewed and accepted for publication. Since it is being posted so soon after acceptance, it has not yet been copyedited, formatted, or processed by AMS Publications. This preliminary version of the manuscript may be downloaded, distributed, and cited, but please be aware that there will be visual differences and possibly some content differences between this version and the final published version.

The DOI for this manuscript is doi: 10.1175/2009JTECHA1279.1

The final published version of this manuscript will replace the preliminary version at the above DOI once it is available.



## Thunderstorm strike probability nowcasting

SANDY DANCE \*

ELIZABETH EBERT

DAVID SCURRAH

*Centre for Australian Weather and Climate Research* †

---

\* *Corresponding author address:* Sandy Dance, CAWCR, GPO Box 1289K, VIC, 3001, Australia

E-mail: s.dance@bom.gov.au

† *A partnership between the Australian Bureau of Meteorology and CSIRO.*

## ABSTRACT

To assist in thunderstorm warning, automated nowcasting systems have been developed that detect thunderstorm cells in radar images and propagate them forward in time to generate forecasted threat areas. Current methods, however, fail to quantify the probabilistic nature of the error structure of such forecasts. This paper introduces the Thunderstorm Environment Strike Probability Algorithm (THESPA), which forecasters can use to provide probabilistic thunderstorm nowcasts for risk assessment and emergency decision making. This method accounts for the prediction error by transforming thunderstorm nowcasts into a strike probability, or the probability that a given location will be impacted by a thunderstorm in a given period, by specifying a bivariate Gaussian distribution of speed and direction errors. This paper presents the development and analysis of the THESPA method and verifies performance using experimental data. Results from a statistical analysis of Thunderstorm Identification, Tracking, Analysis, and Nowcasting (TITAN) tracking errors of nowcasts made near Sydney, Australia were used to specify the distribution, which is then applied to data collected from the World Weather Research Programme (WWRP) Beijing 2008 Forecast Demonstration Project. The results are encouraging and show Brier skill scores between 0.36 and 0.44 with respect to a deterministic advected threat area forecast.

# 1. Introduction

Providing public warnings of impending thunderstorms is an important role of weather services. By observing the motion and evolution of thunderstorms using radar data, forecasters can reasonably predict the location and severity of thunderstorms up to one hour into the future. This is an example of what is widely known as “nowcasting”.

To help objectify and automate the nowcasting process, many radar-based systems have been developed in recent years. Among these are Thunderstorm Identification, Tracking, Analysis and Nowcasting (TITAN; Dixon and Wiener (1993)), Warning Decision Support System (WDSS; Eilts et al. (1996)), Short-Term Ensemble Prediction System (STEPS; Bowler et al. (2006)), Short-range Warnings of Intense Rainstorms in Localized Systems (SWIRLS; Wong et al. (2006)), CANadian Radar Decision System (CARDS; Joe et al. (2003)), Global/Regional Assimilation and PrEdiction System (GRAPES-SWIFT; Feng et al. (2007); Hu et al. (2007)), Nimrod (Golding 1998), National Convective Weather Forecast version 2 (NCWF-2; Megenhardt et al. (2004)) and the system described by Schmeits et al. (2008).

Two of these systems, TITAN and CARDS, classify thunderstorms in radar data by detecting connected regions of high-intensity reflectivity then track these regions through time to generate size, location, and velocity estimates of the thunderstorm. WDSS is similar, but also uses Doppler radar data to detect regions of vorticity associated with tornadoes. The nowcast that is generated from these systems is usually of a deterministic nature, often with an implicit threat area indicated by velocity vectors and detected storm shapes. Forecasters must then interpret these indicators to make public warnings, acknowledging that the now-

casts usually contain significant errors in position, size, and intensity, owing to uncertainties in predicted storm motion and evolution. The usual approach is to hand-draw a threat area that is somewhat broader than the implicit threat area.

STEPS, GRAPES-SWIFT and NIMROD all blend radar-based nowcasting with numerical weather prediction (NWP), and in general forecast rainfall rather than track thunderstorms. STEPS is an ensemble precipitation forecasting system that combines advection forecasts with downscaled NWP rainfall forecasts. Similarly, Nimrod also uses advection in its nowcasting component. SWIRLS makes use of both radar and raingauge data to monitor and predict local rainfall distribution trends within the next couple of hours, using correlation to track radar echoes. The system described by Schmeits et al. (2008) makes probabilistic thunderstorm forecasts over a 6 hour period based on blending NWP and advected lightning and radar storm detections, but unlike the thunderstorm nowcasting systems described earlier, tracks thunderstorm prone regions rather than individual storms.

Another system that generates probabilistic thunderstorm location output is NCWF-2. This system uses both advection and extrapolation of filtered radar reflectivity fields, combined with a probability distribution function (PDF) based on a “local Lagrangian” probability distribution (Germann and Zawadzki 2004) around a point. This system generates probability fields based on the radar reflectivity field, in contrast to our system (described shortly) which generates probability fields from individual thunderstorms detected by other systems, e.g., TITAN.

In this paper we present an algorithm, Thunderstorm Environment Strike Probability Algorithm (THESPA), to generate probabilistic nowcasts. THESPA uses the concept of a “strike probability” originally used in cyclone forecasting, which gives the probability a given

location will be affected by a storm over a nominated period. The two main approaches for computing cyclone strike probability are based on NWP which models the cyclones explicitly. One approach employs an ensemble of numerical models or runs from a single model, where the weighted percentage of the ensemble members that put the cyclone within a set distance of a given point is the strike probability for that point (Weber 2003; van der Grijn 2002). Another approach (Templeton and Keenan 1982) calculates the strike probability by time-interpolating the forecast cyclone position in a single NWP model run, assuming the cyclone path is given by a polynomial, the shape of the cyclone is circular, the probability distribution at each step is given by an a priori bivariate normal distribution, and there is statistical independence between the time steps. The THESPA method computes strike probabilities also using an assumed bivariate normal distribution, but here the approach assumes a linear path, a storm shape given by the initial detection, and complete dependence along potential storm tracks with mutual exclusivity between tracks.

THESPA can improve the capabilities of current nowcasting systems by producing automated objective probabilistic threat areas based on the error characteristics of the estimated storm motion from one or more sources. This probabilistic threat area can be combined with probabilistic products from other automated systems. Additionally, forecasters can manipulate storm motion, size, and other properties to issue better warnings to provide more effective risk assessment and decision-making. Such an approach is being pursued by the U.S. National Weather Service (Kuhlman et al. 2008).

In Section 2 we explore the errors in predicted thunderstorm motion that lead to an empirical strike probability. An analysis of the variability of thunderstorm motion dependent on the nowcast lead time and other storm properties is included. In Section 3 we develop the

THESPA algorithm for generating thunderstorm strike probability. Section 4 shows examples of THESPA strike probability nowcasts generated from TITAN nowcasts, and verifies their accuracy using data collected during the World Weather Research Programme (WWRP) Beijing 2008 Forecast Demonstration Project (B08FDP) (Yu et al. 2005). In Section 5 we give an example of the use of THESPA in B08FDP, where strike probabilities from three independent nowcast systems were combined to create a consensus strike probability. Conclusions are given in Section 6.

## 2. Errors in nowcast thunderstorm motion

Automated nowcast systems such as TITAN attempt to provide an optimal estimate for thunderstorm motion, based on the most recent set of detections for individual storms. The usual approach is to detect thunderstorm cells in 2D radar Constant Altitude Plan Position Indicator (CAPPI) images or 3D radar volumes using specified reflectivity and size thresholds, then “track” them in subsequent radar images using cross-correlation techniques. Each storm’s speed and direction can be estimated from its position in two or more recent detections. Some algorithms adjust the storm motion using information about the motion of surrounding storms, or model-based wind fields (Feng et al. 2007; Hu et al. 2007). Even when using such sophisticated techniques, errors in the predicted storm location may be 5–10km after 30 minutes (Ebert et al. 2004). If the track errors can be estimated, they can be exploited to construct a probabilistic storm nowcast that accounts for the uncertainty in storm motion.

To investigate the statistics of predicted thunderstorm motion a database of 1682 thun-

derstorm tracks was collected from the Kurnell radar near Sydney, Australia. This included 8302 TITAN nowcasts up to one hour, which were verified against actual motion (Ebert et al. 2005). The goal was to determine how the errors in nowcast thunderstorm position (determined by speed and direction) vary as a function of lead time, storm speed, storm strength, and other factors. These TITAN storm tracks were produced in an operational setting in the Sydney office of the Bureau of Meteorology, and used various reflectivity thresholds to detect the thunderstorms. The 35 dBZ threshold was used most often, based on experimentation (J. Bally, personal communication), although thresholds of 40 and 45 dBZ were also used. The radar scan interval was five minutes.

The first step in developing THESPA was understanding what the strike probability should look like. Intuitively, we expect it to be high near the detected storm and in the direction of the predicted motion, and fan out with decreasing probability as the distance from the detected storm increases. In order to make an “empirical” thunderstorm strike probability chart from the Sydney data, we use information about the storm’s predicted trajectory during a period of time, e.g., 1 hour, as well as the error in its predicted motion. Each storm’s velocity vector was normalized by aligning its predicted hourly trajectory onto the x-axis with a (non-dimensional) length of 1, and plotting its subsequent detected positions relative to the predicted trajectory. For a 5-minute scan interval a perfectly predicted storm would yield 12 detections evenly spaced along the x-axis between 0 and 1; imperfectly predicted tracks will generally lie off the x-axis (due to errors in predicted direction) and may have unevenly spaced detections (due to errors in predicted speed). The detected relative tracks were accumulated for all nowcasts to build a two-dimensional histogram representing the empirical strike probability, shown in Figure 1.

It can be seen in Figure 1 that the empirical strike probability has an elliptical core and roughly triangular envelope, in agreement with our expectations. Few thunderstorms lasted the full hour; in fact most lasted only a short time or were slow compared with their initial speed, as evident from the large spike in the histogram less than 0.1 from the origin. This may be related to the use of a 35 dBZ detection threshold, which would include weak storms with a short life.

To build the error information into a mathematical model for strike probability, we next investigated the statistics of the nowcast position errors as a function of lead time and detected storm properties. The position errors were separated into speed and direction components, as earlier studies have indicated that the error in predicted storm direction generally outweighs the error in predicted speed (Bellon and Austin 1978).

Figure 2 shows distance ( $s$ ), direction ( $\theta$ ), and speed ( $V$ ) error standard deviations as a function of nowcast lead time. Standard deviations of the error will be used to specify the bivariate normal probability distributions that underpin the THESPA algorithm. Fig. 2a shows a roughly linear increase in the standard deviation of the Cartesian distance errors with time. When viewed in polar coordinates, Fig. 2b shows that the standard deviation for both speed and direction errors ( $\sigma_V$  and  $\sigma_\theta$ ) are roughly constant with time, at around  $10 \text{ km h}^{-1}$  and  $30^\circ$ , respectively. This is consistent with the findings of Bellon and Austin (1978), and suggests that we can indeed use these standard deviations for predicting the variability of thunderstorm motion.

The results in Figure 2 were for all storms considered together. There may be further

dependencies of speed and direction error on other storm characteristics such as the height of maximum radar reflectivity, vertically integrated liquid (VIL), and detected speed. Figures 3 and 4 show that for the detected storm height of maximum reflectivity and VIL, there is rough consistency of  $\sigma_V$  and  $\sigma_\theta$  through the lead times (the curves overlay each other). Moreover, no functional dependence was found on lead time. Stronger storms, i.e., those with larger values of height of maximum reflectivity and VIL, tended to have smaller errors in storm direction. The speed errors were not sensitive to storm intensity except for very low values of VIL (Fig. 4a).

Figure 5 shows  $\sigma_V$  and  $\sigma_\theta$  versus initial detection speed. Figure 5a shows somewhat erratic behaviour of  $\sigma_V$  at low speed, but becoming fairly consistent at a detected speed greater than about  $15 \text{ km h}^{-1}$ , with a steady increase of  $\sigma_V$  from  $6 \text{ km h}^{-1}$  to  $10 \text{ km h}^{-1}$  as the detected speed increases to  $50 \text{ km h}^{-1}$ . The direction can have quite large uncertainties ( $\sigma_\theta \geq 50^\circ$ ) for the slowest thunderstorms, decreasing to about  $15^\circ$  for storms moving at  $30 \text{ km h}^{-1}$  or more, as one might expect for faster storms whose motion is easier to diagnose (Fig. 5b).

In summary, the speed and direction errors of estimated thunderstorm motion appear to be nearly constant with lead time, which simplifies the mathematical model used by THESPA. Of the thunderstorm measures analyzed, detected speed appears to be the best candidate for functional dependence of speed and direction error standard deviations, compared with height of maximum reflectivity and VIL. However, initial tests indicated that little improvement was likely to be gained from functional dependence. In this study, we assumed constant error statistics for all thunderstorms.

### 3. The THESPA algorithm

We wish to calculate the probability that a thunderstorm passes over some point  $(x, y)$  at some time  $(t)$  in the nowcast period, given an existing thunderstorm defined by its shape, size, speed  $(V)$ , and direction  $(\theta)$  at time  $t = 0$ . Here, without loss of generality, we assume its velocity is oriented along the x axis. It is assumed that the thunderstorm moves at constant velocity, lasts for the full nowcast period, and does not change size or shape. Of course, these are not completely realistic assumptions, but are necessary so the assumption of Lagrangian steady state is justifiable over the relatively short lead times we are considering.

In Section 2 it was shown that the distribution of thunderstorm speed and direction errors had standard deviations  $\sigma_V$  and  $\sigma_\theta$  that were roughly constant over the nowcast period. We assume that at any instant in time  $t$  the central position, which we will call the center of gravity (CoG), of the thunderstorm is given by a bivariate normal (2D Gaussian) distribution in Cartesian space centered at  $x = Vt$ , and whose standard deviations are  $\sigma_x = t\sigma_V$  and  $\sigma_y = Vt \tan(\sigma_\theta)$ . The formula for the 2D Gaussian distribution of thunderstorm CoG positions is:

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\left(\left(\frac{x - Vt}{\sqrt{2}\sigma_x}\right)^2 + \left(\frac{y}{\sqrt{2}\sigma_y}\right)^2\right)} \quad (1)$$

which has a dimension of  $area^{-1}$ , and is normalized so that integrating over all  $x$  and  $y$  gives a probability of 1. Figure 6 illustrates the ellipse of one standard deviation for the position of the CoG after a time  $t$ .

In order to derive the probability that a point  $(x, y)$  will be affected by any part of a thunderstorm over the nowcast period  $T$ , we start by computing the probability that  $(x, y)$  will be affected by the storm's CoG, then expand the scope to include any part of the storm. This probability involves summing the probability that the point coincides with the CoG at time  $T$  together with all CoGs that coincide with the point at earlier times, i.e., faster storms. This is equivalent to summing along a radial from the point to infinity due to the assumption of thunderstorm linear motion. To make the derivation more tractable, we first convert the Cartesian coordinate variables  $(x, y)$  in Eq 1 to polar coordinate variables  $(r, \theta)$ ,<sup>1</sup>. We then set  $t = T$  to use the distribution at the end of the nowcast period, and integrate over radial distances from  $r$  to infinity with each point weighted by the area  $rdrd\theta$ . Thus the probability that a point  $(r, \theta)$  will coincide with the CoG of the thunderstorm over the nowcast period is given by,

$$probCoG(r, \theta) = \int_{s=r}^{+\infty} \frac{1}{2\pi\sigma_x\sigma_y} e^{-\left(\left(\frac{s \cos \theta - VT}{\sqrt{2}\sigma_x}\right)^2 + \left(\frac{s \sin \theta}{\sqrt{2}\sigma_y}\right)^2\right)} s ds \quad (2)$$

where  $s$  and  $ds$  are radial distance measures. Note that  $probCoG(r, \theta)$  has the dimension of

---

<sup>1</sup>To convert from Cartesian to polar coordinates, use  $x = r \cos \theta$  and  $y = r \sin \theta$ .

$angle^{-1}$ . Completing the integration gives:

$$\begin{aligned}
probCoG(r, \theta) = & \\
& \frac{\sigma_y}{4\pi} e^{-\frac{T^2 V^2}{2\sigma_x^2}} \left( \sqrt{2\pi} \sigma_y TV \cos \theta e^{-\frac{\cos^2 \theta \sigma_y^2 T^2 V^2}{2\sigma_x^2 \eta}} \left( 1 - \operatorname{erf} \left( \frac{r\eta - \cos \theta \sigma_y^2 TV}{\sqrt{2\eta} \sigma_x \sigma_y} \right) \right) / \eta^{3/2} \right. \\
& \left. + \frac{2\sigma_x e^{\frac{r}{2} \left( 2 \cos \theta TV / \sigma_x^2 - \frac{r\eta}{\sigma_x^2 \sigma_y^2} \right)}}{\eta} \right) \quad (3)
\end{aligned}$$

where  $\eta = \sin^2 \theta \sigma_x^2 + \cos^2 \theta \sigma_y^2$  and erf is the error function (Fristedt and Gray 1997).

The next step in the development of the algorithm is to derive a distribution which characterizes the locations that will be affected by any part of the thunderstorm over the nowcast period, assuming the size and shape of the storm remains unchanged during this period. To compute the required probability given a specific thunderstorm, we need for each point  $(x, y)$  to combine the probabilities of all possible future positions of the thunderstorm that overlay  $(x, y)$ . This set of possible future thunderstorms is dubbed the thunderstorm overlay set, and the set of CoGs of these possible storms is dubbed the thunderstorm area or TSA, as shown in Figure 7. The CoGs for all possible thunderstorms overlaying  $(x, y)$  encompass the TSA, whose shape is a 180 degree rotation of the original thunderstorm centred at  $(x, y)$ . Thus for each such possible thunderstorm position overlaying  $(x, y)$ , we find the probability at its CoG using Eq 3, and to account for the whole TSA, we combine such probabilities for all storm footprints within the thunderstorm overlay set.

Since the probabilities over the TSA are spatially correlated, they cannot simply be added. The correlation results from the fact that some potential positions of the CoG encompass other positions, namely those of thunderstorm trajectories in the same direction but faster, i.e., those which have passed over the same point earlier, and are now further

from the origin. These “downstream” points should not be included in the combination. Since the probability density of CoG positions are in units of  $angle^{-1}$ , we can find the total probability ( $strikeProb(x, y)$ ) that the point  $(x, y)$  was affected by any part of the thunderstorm by integrating along the “trailing edge” of the TSA for that point. This trailing edge is seen as points of the TSA closest to the origin in Figure 8. Let  $TE_{xy}(\theta)$  (a mapping from polar angle to radial distance) describe the shape of the trailing edge for the given TSA with  $\theta_1$  to  $\theta_2$  being the angular limits of the trailing edge. Computing the strike probability involves integrating  $probCoG(r, \theta)$  (Eq. 3) along the trailing edge giving:

$$strikeProb(x, y) = \int_{\theta_1}^{\theta_2} probCoG(TE_{xy}(\theta), \theta) d\theta \quad (4)$$

where  $r$  has been replaced by  $TE_{xy}(\theta)$ . Note that the relatively small number of points along the trailing edge compared with the full area of the TSA leads to more efficient computation. One then repeats this operation for each point  $(x, y)$  in the geographic region to generate the complete probability field.

To deal with the case of multiple thunderstorms in the same region, the strike probability distribution can be calculated for each thunderstorm separately, and then pointwise combined with the usual rules of probability, assuming independence:

$$strikeProb(x, y) = 1 - \prod_{j=1}^J \left(1 - strikeProb_j(x, y)\right) \quad (5)$$

for  $J$  thunderstorms affecting a point  $(x, y)$ . Of course, storms are not necessarily statistically independent (Grecu and Krajewski 2000); however, by assuming they are we create probabilities that are upper bounds of the real values. An alternative approach would be to take the average probability, or the maximum. In a thunderstorm warning environment, potentially overstating the threat by using an upper bound may be prudent.

THESPA strike probability nowcasts for some typical thunderstorm parameters are shown in Figure 9. The original thunderstorm shape is shown schematically as an ellipse, whose interior of course has a probability of 1 since those points have already been affected by the thunderstorm. For high speed thunderstorms relative to speed standard deviation ( $\sigma_V$ ), the probability field shows a smoothed “triangular” envelope centered around the velocity vector (Fig. 9a). For slower moving thunderstorms there is a non-negligible chance that the thunderstorm will move opposite to its initial velocity vector, which shows up as finite probability on the back side of the thunderstorm (with respect to velocity) (Fig. 9b).

## 4. Accuracy of strike probability nowcasts

We assess the quality of the strike probabilities by examining their reliability, that is, the agreement between the predicted probabilities and the corresponding observed detection frequency.

For each storm track, a 60-minute strike probability nowcast at 1 *km* spatial resolution was generated from the first detection that included a nowcast velocity (this is the second

thunderstorm detection, as TITAN requires two detections to compute velocity). The values for speed and direction uncertainty were set to  $\sigma_V = 10 \text{ km h}^{-1}$  and  $\sigma_\theta = 30^\circ$  in accordance with the results from Section 2. For verification, the observed area affected by the storm was defined as the convex hull of each storm observation (which were at 5 or 6 minute intervals) and its predecessor in its track (if any), as shown in Figure 10. For all forecast pixels with a given probability value the number of pixels containing observed thunderstorm occurrences and non-occurrences were counted, giving an observed relative frequency. These observed relative frequencies were then plotted against probability value giving a reliability diagram (Wilks 1995). Since the algorithm will return non-zero probability values for points far from the initially detected storm, a bounding box is drawn around the storm based on its size and velocity. Not calculating a probability beyond this box improves computational efficiency of the algorithm. The ignored points had less than 0.030, and usually less than 0.005, probability.

We also calculate the Brier skill score (BSS) (Wilks 1995) as an overall measure of the system’s performance. The Brier score (BS) gives the magnitude of the probability forecast errors, and is defined as:

$$BS = \frac{1}{N} \sum_{i=1}^N (p_i - o_i)^2 \quad (6)$$

where  $p_i$  is the  $i$ th probability forecast of some event occurring in a series of  $N$  trials, and  $o_i$  is the subsequent observation of that event, set to 1 if it occurs and zero otherwise. The Brier score measures the mean squared probability error, and ranges from 0 to 1 with a perfect score at 0. The BSS measures the relative skill of a probabilistic forecast over that of some

reference forecast, and is defined as:

$$BSS = 1 - \frac{BS}{BS_{reference}} \quad (7)$$

and ranges from minus infinity to 1. A BSS of 0 indicates no skill when compared to the reference forecast, and a perfect score is 1.

For a reference forecast we use a variation of the current manual forecast method, which is to project the detected storm forward at its detected velocity for the nowcast period, and to define the probability of the swept out area to be one, and the rest zero (Figure 11). We refer to this as the “advected threat area forecast”.

To get an initial assessment of the performance of THESPA to forecast thunderstorm strike probability, we examined the aforementioned Sydney data set. Since this dataset was also used to determine the speed and direction standard deviations for the THESPA algorithm, this initial study is not as useful a guide to the performance as an independent dataset, but can be viewed as an upper bound on system performance. Overall, higher nowcast strike probabilities were associated with higher observed frequency of thunderstorm detection, which means THESPA was working correctly. When verified for all storms THESPA had a tendency to overestimate the probabilities, particularly at the highest values (Fig. 12a). These higher probabilities, in general over 0.8, were for regions close to the original storm. The drop in observed frequency for these regions was due to weak storms which were not subsequently detected, i.e., having a storm lifetime of one radar scan cycle (subsequently called a “storm lifetime of one step”). When these weak storms were filtered out using VIL

greater than  $30 \text{ kg m}^{-2}$  as a criterion, the reliability improved substantially (Fig. 12b). To confirm this explanation, it was found that the number of storms with a lifetime of one step in the dataset was 175, or about 10% of the total, but when storms with VIL under  $30 \text{ kg m}^{-2}$  were filtered out, the number of storms with a lifetime of one step was 21, or about 1.2%. The smaller number of cases at these higher probabilities (around 100 pixels per probability increment) also gave noisy or erratic results due to limited statistical sampling. Interestingly these results are not reflected in the BSS, which shows little change between filtered and non-filtered data. Even though the BS is halved, i.e., improved, for the filtered data (Table 1), filtering the data also improved the advected forecast.

Next, THESPA was evaluated using TITAN thunderstorm nowcasts from the B08FDP. In this experiment forecasters used a storm detection threshold of  $45 \text{ dBZ}$ , somewhat greater than was used in the Sydney calibration dataset. This provided a dataset of 2394 tracks, which was processed with the same standard deviations as determined from the Sydney data. The climate of Beijing is different to that of Sydney, but one would expect the higher threshold to select more stable, predictable storms, and hence THESPA should be more reliable, especially at higher probability, than for the Sydney data. This was found to be the case as will be shown next.

The reliability for the THESPA strike probabilities in B08FDP can be seen in Figure 13. Fig. 13a shows results for all storms, and Fig. 13b shows results for the same storms also filtered on VIL greater than  $30 \text{ kg m}^{-2}$ . Here we also include the “no resolution” (climatology) and the “no skill” (halfway between climatology and a perfect forecast) lines in the reliability diagram. The climatological strike frequency was calculated from the 51 days of the B08FDP, 1 August - 20 September 2008. This value for the B08FDP at  $45 \text{ dBZ}$  is 0.003. The

THESPA strike probabilities were quite reliable for the Beijing data, with better reliability for high probability forecasts, compared with Sydney. The BSS for Beijing varied between 0.36 and 0.44, depending on the storm detection threshold and whether or not the storms were filtered by VIL (Table 1). The improved behaviour of the VIL filtered data (Fig. 13b) at high probability can be explained in terms of filtering out of numerous short-lived low-VIL “pulse” storms observed during B08FDP (P. Joe, personal communication). We verified this hypothesis by counting the number of short-lived storms (storm lifetimes of 3 or fewer steps, or 12 minutes, which is 20% of the nowcast period) in the database, which amounted to 830 or about 33% of all storms. These short-lived storms would mostly affect higher strike probability regions. This was then compared with the same count after filtering out storms with VIL less than  $30 \text{ kg m}^{-2}$ , and indeed, the number of short-lived storms dropped to 228 (or about 9%).

To test the sensitivity of THESPA performance to the assigned parameters, namely thunderstorm detection threshold,  $\sigma_V$  and  $\sigma_\theta$ , the B08FDP radar data were reprocessed to generate new TITAN nowcasts and new THESPA strike probabilities. Fig. 14 shows the reliability when the detection threshold was 45 dBZ, the standard deviation for speed was decreased to  $7 \text{ km h}^{-1}$  (Fig. 14a), and the standard deviation for direction was reduced to  $20^\circ$  (Fig. 14b). Reducing the specified uncertainty in the direction error reduced the angular spread of the probability forecasts so that the probabilities were increased overall. This increase in over-prediction was slightly detrimental to the forecast. On the other hand, the reduction in speed error uncertainty had less effect on the reliability. This is due to the speed error uncertainty mainly affecting the “leading edge” of the probability distribution, with fewer storms surviving long enough to influence the statistics there, whereas the direction

error uncertainty affects the distribution for storms of any lifetime. Neither of these effects was reflected in the BSS, suggesting that THESPA is not very sensitive to  $\sigma_V$  and  $\sigma_\theta$ .

Using a lower detection threshold of 40 dBZ (generating a database of 4244 storm tracks) appeared to have little effect on the reliability of the strike probabilities, as seen in Figure 15. The BSS values were also little effected, with 0.38 for 40dBZ data, and 0.39 for 45dBZ data. Similarly, when filtered on VIL there was little or no difference, with BSS of 0.41 for both datasets. This is likely due to the small change in the dataset size when filtered on 40 vs 45 dBZ, whereas the next step in dBZ, explained next, appears to have a larger effect.

Using a 35 dBZ threshold (resulting in a database of 12597 storm tracks) had a negative impact with a BSS of 0.36. And unlike the 40 and 45 dBZ TITAN nowcasts, restricting the storms to high-VIL cases improved the reliability for the 35 dBZ threshold case (Fig. 16b), with an improvement in the BSS to 0.44. This is due to filtering out short-lived storms, with total short-lived storms in the database numbering 5521 or about 36%, and short-lived storms with VIL greater than  $30 \text{ kg m}^{-2}$  numbering 376 or about 2.5%. Table 1 gives a a summary of these results. In fact, the TITAN storm detection threshold was raised from 35 dBZ to 40 dBZ early in the B08FDP when it was seen that heavy stratiform rain was being mis-diagnosed as thunderstorms (J. Bally, personal communication). In general, the smaller the percentage of short lived storms, the better the BSS, suggesting that the THESPA algorithm could be improved if it took into account the distribution of storm lifetimes for the location and season.

## 5. Real time example

The computational efficiency of the THESPA algorithm allowed it to be run in the Thunderstorm Interactive Forecast System (TIFS; Bally (2004)) in real time during B08FDP, as strike probability calculations for several storms in a  $500km$  square domain took only 1-2 seconds when run on a Linux workstation. In real time operation in B08FDP strike probabilities were calculated not only for TITAN thunderstorm nowcasts but also from SWIRLS (Wong et al. 2006) and CARDS (Joe et al. 2003) nowcasts. In the absence of advance knowledge of typical track errors for SWIRLS and CARDS nowcasts, the strike probability calculations for these nowcasts used the same values for  $\sigma_V$  and  $\sigma_\theta$  as were determined for the Sydney TITAN cases. The strike probability grids from each individual nowcast system were then averaged to create consensus strike probability nowcasts. In principle the input nowcasts could be weighted according to their recent accuracy; however, this was not done during the B08FDP. These consensus strike probability grids formed the basis for integrated objective thunderstorm guidance used by operational forecasters at the Beijing Meteorological Bureau (BMB) during the Olympic Games (Bally et al. 2009).

A sample strike probability chart is shown in Figure 17 for nowcasts generated on 6 September 2008 when thunderstorms were moving through Beijing from the southwest. TITAN, SWIRLS, and CARDS all detected the presence of a large thunderstorm over the city center, but estimated somewhat different sizes and velocities for the storm. The consensus strike probability map shows a strong chance of the thunderstorm affecting a SW-NE region across central Beijing, with probabilities falling off toward the northeast. The two smaller target regions in the north-eastern part of the strike probability map reflect the prediction of thunderstorms by only one of the three input systems (SWIRLS), and therefore have a central probability of 0.33. These display the characteristic fan shaped probability pattern predicted by the THESPA algorithm (see Figure 9).

This is the first time probabilistic thunderstorm predictions from different nowcasting systems have been combined. BMB forecasters found them very useful, according to surveys conducted during the Games (Y. Duan, personal communication). Verification of the real-time consensus strike probability nowcasts produced during the B08FDP showed a high degree of reliability and skill (Bally et al. 2009).

## 6. Conclusions

In this paper we have presented a new algorithm for producing thunderstorm strike probabilities for thunderstorms detected with nowcast systems. A brief review of existing thunderstorm detection and nowcast systems showed that most such systems generate output consisting of shape, size and velocity of thunderstorms, but not probabilities that can be used by other systems. Some nowcast systems do generate thunderstorm probabilities,

but these in general do not nowcast individual storms, but rather the likelihood of thunderstorms occurring over regions, often blending NWP output with radar or other observational techniques.

We next investigated the statistics of thunderstorm motion using a dataset from a radar near Sydney, Australia, and found that the standard deviation of errors in speed and direction of motion were roughly constant over the lifetime of the storms. This provided the basis for the subsequent development of the THESPA algorithm.

The THESPA algorithm was designed for mathematical tractability and algorithmic simplicity. Verifying its output using datasets from the original Sydney radar, and also data from the Beijing 2008 Forecast Demonstration Project (B08FDP) showed the reliability to be quite good, and with BSS between 0.36 and 0.44 with respect to an advected threat area forecast. The results were not very sensitive to the chosen values of the parameters, i.e., detection threshold,  $\sigma_V$  and  $\sigma_\theta$ . Likewise, favorable performance was shown when the algorithm parameters determined using the Sydney data were applied to data collected during the B08FDP. The hypothesis that these parameters provide similar results in other locations should be substantiated in a future study. Looking more closely at the results, we also found that the smaller the percentage of short lived storms, i.e., those with a lifetime of 3 or fewer steps, the better the skill. This indicates that a fruitful line of enquiry would be to further develop the THESPA algorithm to take into account the statistics of storm lifetime.

Some additional calibration of the probabilities, e.g., as a post-processing step, could improve their reliability still further. Including a functional dependence of the speed and direction error standard deviations on detected storm properties, namely VIL and thunderstorm speed, may also be tested.

The algorithm was embedded in the Thunderstorm Interactive Thunderstorm Forecasting System used by forecasters in the B08FDP, who found THESPA strike probabilities to be very useful. THESPA will be migrated into an operational environment in the Australian Bureau of Meteorology in the near future.

## 7. Acknowledgments

Tony Bannister and John Bally provided many useful comments during the course of this work. The authors would like to thank the anonymous reviewers, especially reviewer A, for their very useful contributions to this paper.

## REFERENCES

- Bally, J., 2004: The thunderstorm interactive forecast system: Turning automated thunderstorm tracks into severe weather warnings. *Weather Forecasting*, **19**, 64–72.
- Bally, J., D. Scurrah, E. Ebert, and D. Su, 2009: Composite products and nowcast decision support for the Beijing 2008 forecast demonstration project. *WMO Symposium on Nowcasting*, Whistler, B.C., Canada.
- Bellon, A. and G. L. Austin, 1978: The evaluation of two years of real-time operation of a

- short-term precipitation forecasting procedure (SHARP). *Journal of Applied Meteorology*, **17 (12)**, 1778–1787.
- Bowler, N., C. Pierce, and A. Seed, 2006: STEPS: A probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP. *Quarterly Journal of the Royal Meteorological Society*, **132**, 2127–2155.
- Dixon, M. and G. Wiener, 1993: TITAN: thunderstorm identification, tracking, analysis and nowcasting - a radar based methodology. *J. Atmos. Oceanic Tech.*, **10 (6)**, 785–797.
- Ebert, E., T. Keenan, J. Bally, and S. Dance, 2005: Verification of operational thunderstorm nowcasts. *World Weather Research Program Symposium on Nowcasting and Very Short Range Forecasting*, Toulouse, France, 8.12.
- Ebert, E., L. Wilson, B. Brown, P. Nurmi, H. Brooks, J. Bally, and M. Jaeneke, 2004: Verification of nowcasts from the WWRP sydney 2000 forecast demonstration project. *Weather Forecasting*, **19**, 73–96.
- Eilts, M. D., et al., 1996: Severe weather warning decision support system. *Preprints, 18th Conf. on Severe Local Storms*, San Francisco, CA, USA, 536–540.
- Feng, Y., Y. Wang, T. Peng, and J. Yan, 2007: An algorithm on convective weather potential in the early rainy season over the Pearl River Delta in China. *Advances in Atmospheric Sciences*, **24 (1)**, 101–110.
- Fristedt, B. and L. Gray, 1997: *A Modern Approach to Probability Theory*. Birkhaeuser Boston, 756 pp.

- Germann, U. and I. Zawadzki, 2004: Scale dependence of the predictability of precipitation from continental radar images. Part II: Probability forecasts. *Journal of Applied Meteorology*, **43** (1), 74–89.
- Golding, B., 1998: Nimrod: a system for generating automated very short range forecasts. *Meteorological Applications*, **5**, 1–16.
- Greco, M. and W. F. Krajewski, 2000: A large-sample investigation of statistical procedures for radar-based short-term quantitative precipitation forecasting. *J. Hydrol*, **239**, 69–84.
- Hu, S., S. Gu, X. Zhuang, and H. Luo, 2007: Automatic identification of storm cells using Doppler radars. *Acta Meteorologica Sinica*, **21** (3), 353–365.
- Joe, P., M. Falla, P. V. Rijn, L. Stamadianos, T. Falla, D. Magosse, L. Ing, and J. Dobson, 2003: Radar data processing for severe weather in the national radar project of Canada. *21st Conf. Severe Local Storms*, San Antonio, TX, 221–224.
- Kuhlman, K. M., T. M. Smith, G. J. Stumpf, K. L. Ortega, and K. L. Manross, 2008: Experimental probabilistic hazard information in practice: Results from the 2008 EWP spring program. *Preprints, 24th Conf. Severe Local Storms*, Savannah, GA. USA.
- Megenhardt, D. L., C. Mueller, S. Trier, D. Ahijevych, , and N. Rehak, 2004: Nwcf-2 probabilistic forecasts. *11th Conference on Aviation, Range, and Aerospace*, Hyannis, MA, 5.2.
- Schmeits, M. J., K. J. Kok, D. H. P. Vogelesang, and R. M. V. Westrhenen, 2008: Probabilistic forecasts of (severe) thunderstorms for the purpose of issuing a weather alarm in the Netherlands. *Weather and Forecasting*, **23**, 1253–1267.

- Templeton, J. I. and T. D. Keenan, 1982: Tropical cyclone strike probability forecasting in the Australian region. Tech. Rep. 49, Bureau of Meteorology, GPO Box 1289K, Melbourne, VIC, 3001, Australia.
- van der Grijn, G., 2002: Tropical cyclone forecasting at ECMWF: New products and validation. Tech. Rep. 386, ECMWF Technical Memorandum.
- Weber, H. C., 2003: Hurricane track prediction using a statistical ensemble of numerical models. *Monthly Weather Review*, **131** (5), 749–770.
- Wilks, D. S., 1995: *Statistical Methods in Atmospheric Sciences*. Academic Press, 467 pp.
- Wong, M., W. Wong, and E. Lai, 2006: From SWIRLS to RAPIDS: Nowcast applications development in Hong Kong. *WMO PWS Workshop on Warnings of Real-Time Hazards by Using Nowcasting Technology*, Sydney, Australia, 9–13.
- Yu, X., T. Keenan, P. Joe, J. Wilson, J. Bally, , and Y. C. Wang, 2005: Overview of Beijing 2008 Olympics WWRP nowcast FDP scientific and implementation plan. *World Weather Research Program Symposium on Nowcasting and Very Short Range Forecasting*, Toulouse, France.

## List of Tables

- 1 BS for Thespa (BS Th), BS for advected forecast (BS ad), resulting BSS and percentage of short lived storms, i.e., lifetimes of fewer than 4 steps, for cases of filtering and not filtering on VIL less than  $30 \text{ kg m}^{-2}$  for datasets of Beijing data thresholded on 45, 40 and 35dBZ and the Sydney data. In general, the smaller the percentage of short lived storms, the better the BSS. 27

dataset	no VIL filter				VIL filter			
	BS Th	BS ad	BSS	% short life	BS Th	BS ad	BSS	% short life
45dBZ	0.00009	0.00016	0.39	33	0.00004	0.00007	0.41	9
40dBZ	0.0002	0.00031	0.38	33	0.00005	0.00008	0.41	3
35dBZ	0.0008	0.0012	0.36	36	0.00013	0.00023	0.44	2.5
Sydney	0.00020	0.00029	0.31	40	0.00010	0.00014	0.29	12

TABLE 1. BS for Thespa (BS Th), BS for advected forecast (BS ad), resulting BSS and percentage of short lived storms, i.e., lifetimes of fewer than 4 steps, for cases of filtering and not filtering on VIL less than  $30 \text{ kg m}^{-2}$  for datasets of Beijing data thresholded on 45, 40 and 35dBZ and the Sydney data. In general, the smaller the percentage of short lived storms, the better the BSS.

## List of Figures

- 1 Histogram of normalized detected thunderstorm positions (from 1682 Kurnell thunderstorm tracks) relative to their nowcast trajectories. Each detected thunderstorm position was recorded in polar coordinates  $(r_d/(t_n s_n), \theta_d - \theta_n)$  where  $\theta_d$  = detected angular position,  $\theta_n$  = nowcast angular position,  $r_d$  = detected radial position,  $t_n$  = nowcast period (lead time) and  $s_n$  = nowcast speed, and the result converted to (dimensionless) Cartesian coordinates. This value was then used to increment a corresponding bin in this 2 dimensional histogram. Counts for a given bin are indicated by the scale on the right. 32
- 2 Thunderstorm nowcast verification statistics (from 1682 Kurnell thunderstorm tracks) as a function of lead time  $t$  (minutes) showing standard deviations of errors in, (a),  $s$  distance in  $km$ ,  $s/t$  and  $s/\sqrt{t}$ , and (b),  $V$  speed ( $km\ h^{-1}$ ) and  $\theta$  direction (*degrees*). 33
- 3 Thunderstorm nowcast verification statistics as a function of height of maximum reflectivity showing (a) standard deviations of errors in speed ( $\sigma_V$ ,  $km\ h^{-1}$ ), and (b), standard deviations of errors in direction ( $\sigma_\theta$ , *degrees*), with lead times 5 to 60 minutes shown with grey lines, and the average over all lead times shown in black. No systematic variation was found with respect to lead time, the grey lines are shown to indicate the scatter about the mean. 34
- 4 As in Figure 3, for verification statistics as a function of VIL. Not all nowcasts in the Sydney dataset reported VIL, so here we show only the 10 minute lead times (grey lines), average over all lead times (black). 35

- 5 As in Fig. 3, for verification statistics as a function of initial detected speed  $V$ . Lead times 5 to 60 minutes (grey), average over all lead times (black). 36
- 6 Schematic showing the original thunderstorm (TS) and the 1 standard deviation ellipse for its CoG after time  $t$  given speed  $V$ , with velocity and direction std dev of  $\sigma_V, \sigma_\theta$ . 37
- 7 A set of possible future positions of the original thunderstorm (shaded) that overlay a specific point  $(x, y)$  (thunderstorm overlay set). Note that those shown (light curves) are examples of the extremes of the thunderstorm overlay set, with  $(x, y)$  being right on their boundaries, and the collection of their CoGs laying on the dark solid curve (the thunderstorm area (TSA) boundary). The thunderstorm overlay set also includes all those potential thunderstorms whose CoGs lie within the TSA boundary. 38
- 8 Plan in Cartesian coordinates of the original thunderstorm (in grey), the thunderstorm area (TSA) for a position  $(x, y)$ , and the “trailing edge” (thick line) that contributes all the probability to the sum for that position.  $\theta_1$  and  $\theta_2$  are the angular limits of the trailing edge. 39
- 9 Examples of output from THESPA with feasible storm tracking values. Figure (a),  $\theta = 115^\circ, \sigma_\theta = 30^\circ, V = 75 \text{ km h}^{-1}, \sigma_V = 10 \text{ km h}^{-1}$ , figure (b),  $\theta = 115^\circ, \sigma_\theta = 30^\circ, V = 4 \text{ km h}^{-1}, \sigma_V = 10 \text{ km h}^{-1}$ , with probability contours shown at 0.1 intervals. The original storms are shown as ellipses with nowcast velocity as an arrow. A scale marker shows  $25 \text{ km}$ . Black shading represents a strike probability of 1, white shading represents a strike probability of zero, as indicated by the grey-scale bar on the right. 40

- 10 Schematic diagram showing a pair of successive storm detections (dark grey) at time steps  $t = n$  and  $t = n + 1$ . For verification we wish to include not only the points under detected storms (dark grey), but also the points between detections that were presumably under the storm between radar scans (light grey). Hence for verification we include all points in the convex hull (shown in both shades of grey) of the two detections. 41
- 11 Schematic diagram showing the advected threat area forecast modelled on current manual practice which is used as the reference forecast in calculating the Brier skill score (BSS). Here a detected thunderstorm is swept forward along the detected velocity vector to a new position at the end of the now-cast period, with the swept out area defined as having a thunderstorm strike probability of one, and the rest zero. 42
- 12 Reliability diagrams for thunderstorm strike probability computed from the original 1682 Sydney TITAN storm tracks with  $\sigma_V = 10 \text{ km h}^{-1}$  and  $\sigma_\theta = 30^\circ$ , for (a) all thunderstorms, and (b) only those storms with VIL greater than  $30 \text{ kg m}^{-2}$ . The diagonal line indicates perfect reliability. The sharpness diagram at upper left on each chart shows the number of times each probability value (horizontal axis, linear scale) was predicted (vertical axis, logarithmic scale). The BSS for (a) was 0.31 and for (b) was 0.29, with respect to the advected threat area forecast. 43

- 13 As in Fig. 12, for THESPA strike probabilities generated from TITAN nowcasts collected during the WWRP Beijing 2008 Forecast Demonstration Project using a storm detection threshold of 45 dBZ (2394 thunderstorm tracks). This diagram includes the “no skill” and “no resolution” lines calculated from the climatology of thunderstorm occurrence. The BSS for (a) was 0.39 and for (b) was 0.41, with respect to the advected threat area forecast. 44
- 14 As in Fig. 13a, with smaller assumed (a) speed error  $7km\ h^{-1}$ , and (b) direction error  $20^\circ$ . The BSS for (a) was 0.40 and for (b) was 0.41, with respect to the advected threat area forecast. 45
- 15 As in Fig. 13a, with a storm detection threshold set to 40 dBZ (4244 storm tracks), for (a) all storms and (b) only those storms with VIL greater than  $30kg\ m^{-2}$ . The BSS for (a) was 0.38 and for (b) was 0.41, with respect to the advected threat area forecast. 46
- 16 As in Fig. 15, with a storm detection threshold set to 35 dBZ (12597 storm tracks). The BSS for (a) was 0.36 and for (b) was 0.44, with respect to the advected threat area forecast. 47
- 17 Thunderstorm nowcasts from the TITAN, SWIRLS, and CARDS nowcasting systems, and consensus strike probability (lower right panel) for the hour beginning at 23:24 UTC on 6 September 2008. The arrows show the predicted movement of the thunderstorms during the next 60 minutes. 48

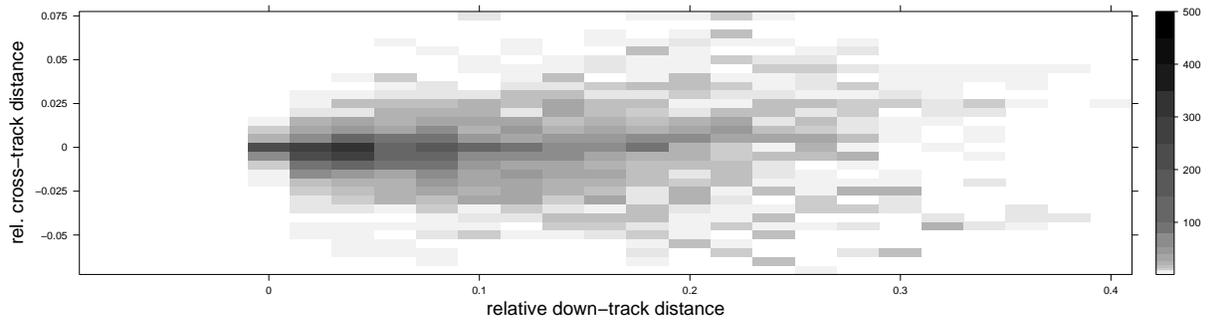


FIG. 1. Histogram of normalized detected thunderstorm positions (from 1682 Kurnell thunderstorm tracks) relative to their nowcast trajectories. Each detected thunderstorm position was recorded in polar coordinates  $(r_d/(t_n s_n), \theta_d - \theta_n)$  where  $\theta_d =$  detected angular position,  $\theta_n =$  nowcast angular position,  $r_d =$  detected radial position,  $t_n =$  nowcast period (lead time) and  $s_n =$  nowcast speed, and the result converted to (dimensionless) Cartesian coordinates. This value was then used to increment a corresponding bin in this 2 dimensional histogram. Counts for a given bin are indicated by the scale on the right.

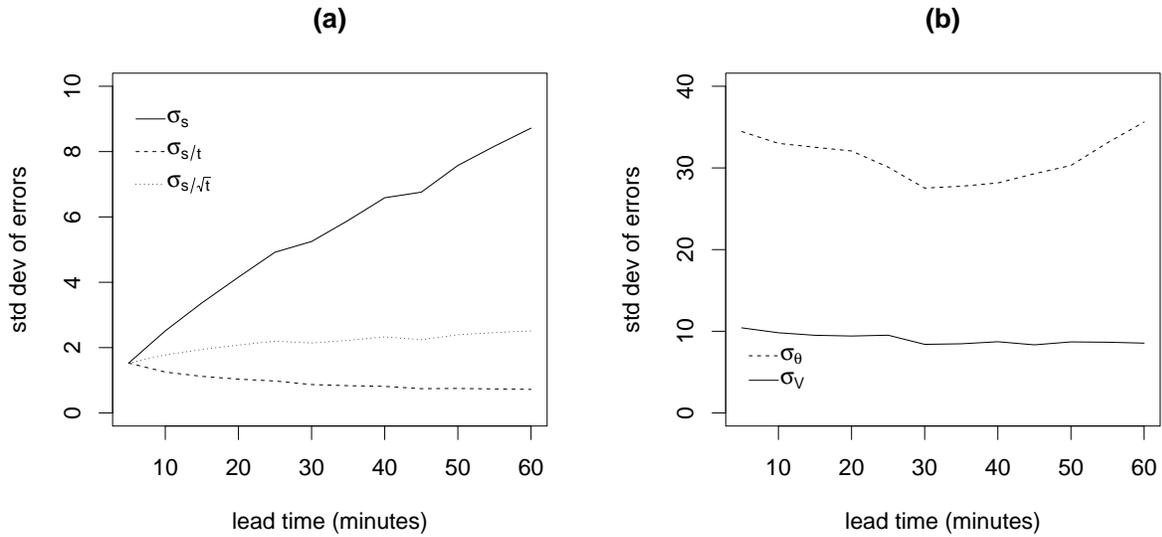


FIG. 2. Thunderstorm nowcast verification statistics (from 1682 Kurnell thunderstorm tracks) as a function of lead time  $t$  (minutes) showing standard deviations of errors in, (a),  $s$  distance in  $km$ ,  $s/t$  and  $s/\sqrt{t}$ , and (b),  $V$  speed ( $km\ h^{-1}$ ) and  $\theta$  direction (*degrees*).

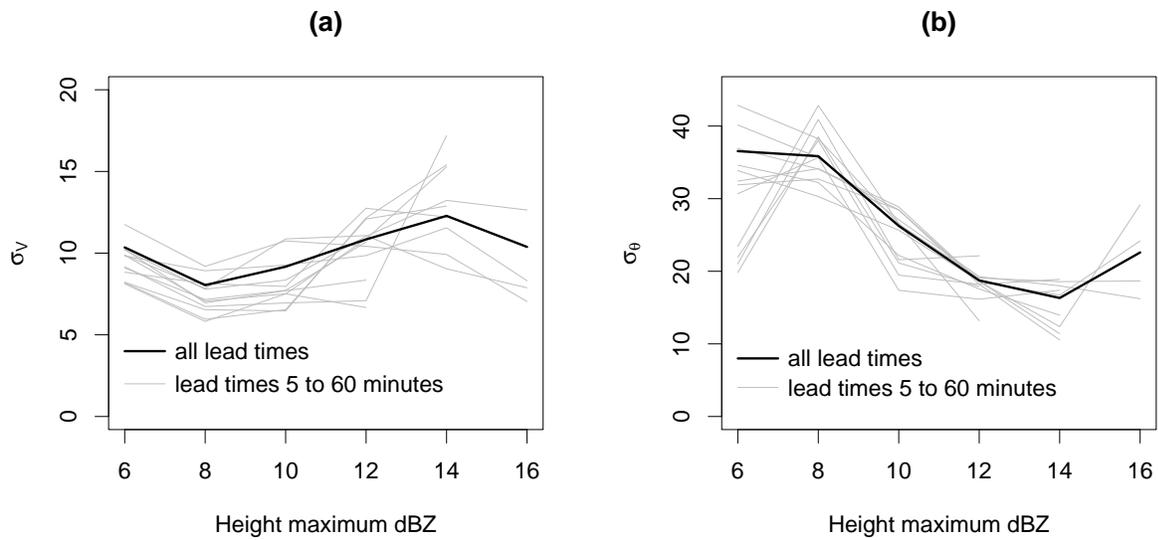


FIG. 3. Thunderstorm nowcast verification statistics as a function of height of maximum reflectivity showing (a) standard deviations of errors in speed ( $\sigma_V$ ,  $km\ h^{-1}$ ), and (b), standard deviations of errors in direction ( $\sigma_\theta$ , *degrees*), with lead times 5 to 60 minutes shown with grey lines, and the average over all lead times shown in black. No systematic variation was found with respect to lead time, the grey lines are shown to indicate the scatter about the mean.

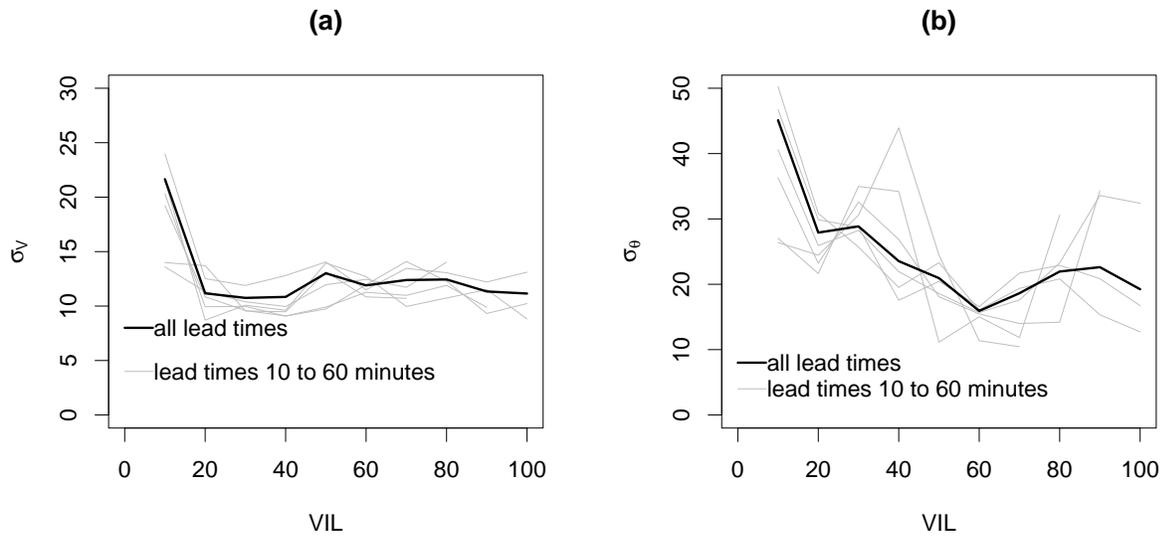


FIG. 4. As in Figure 3, for verification statistics as a function of VIL. Not all nowcasts in the Sydney dataset reported VIL, so here we show only the 10 minute lead times (grey lines), average over all lead times (black).

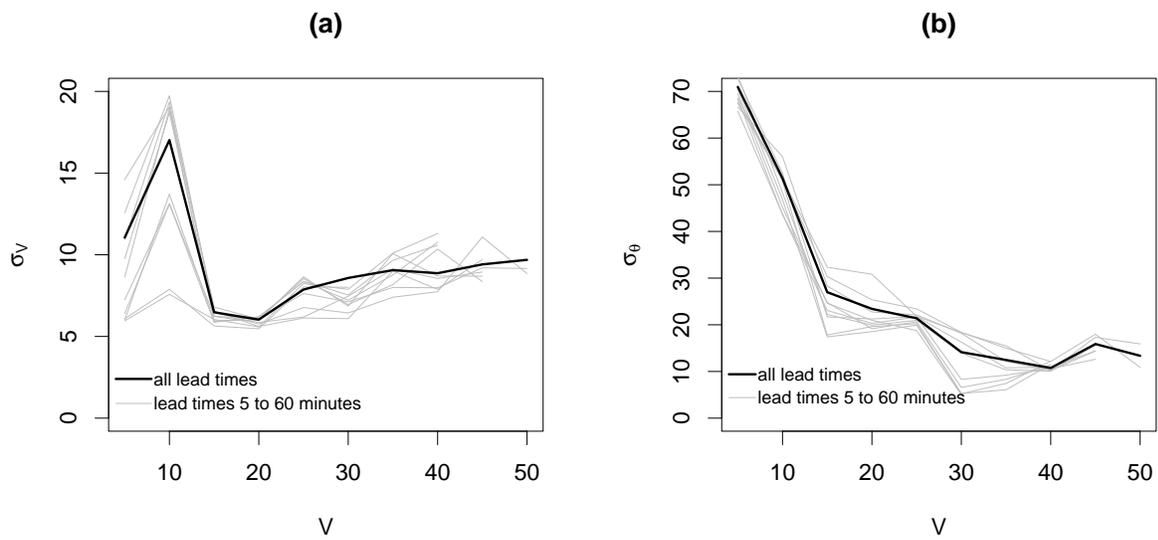
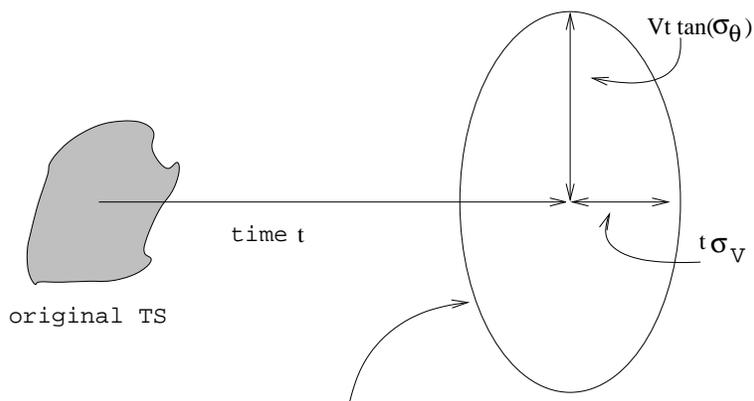


FIG. 5. As in Fig. 3, for verification statistics as a function of initial detected speed  $V$ . Lead times 5 to 60 minutes (grey), average over all lead times (black).



1 std dev region for future TS CoG in Cartesian coordinate system.

FIG. 6. Schematic showing the original thunderstorm (TS) and the 1 standard deviation ellipse for its CoG after time  $t$  given speed  $V$ , with velocity and direction std dev of  $\sigma_V$ ,  $\sigma_\theta$ .

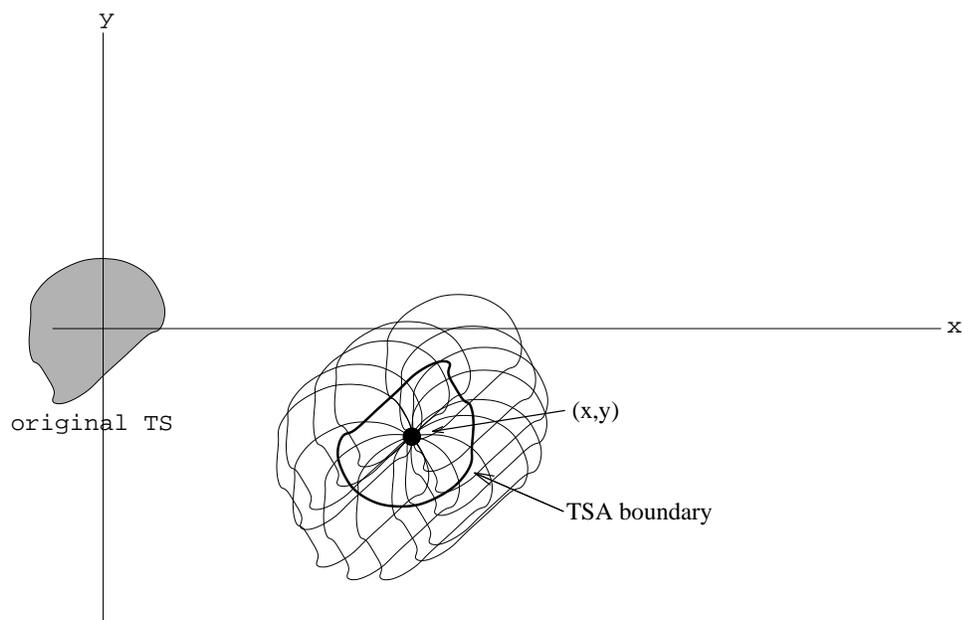


FIG. 7. A set of possible future positions of the original thunderstorm (shaded) that overlay a specific point  $(x, y)$  (thunderstorm overlay set). Note that those shown (light curves) are examples of the extremes of the thunderstorm overlay set, with  $(x, y)$  being right on their boundaries, and the collection of their CoGs laying on the dark solid curve (the thunderstorm area (TSA) boundary). The thunderstorm overlay set also includes all those potential thunderstorms whose CoGs lie within the TSA boundary.

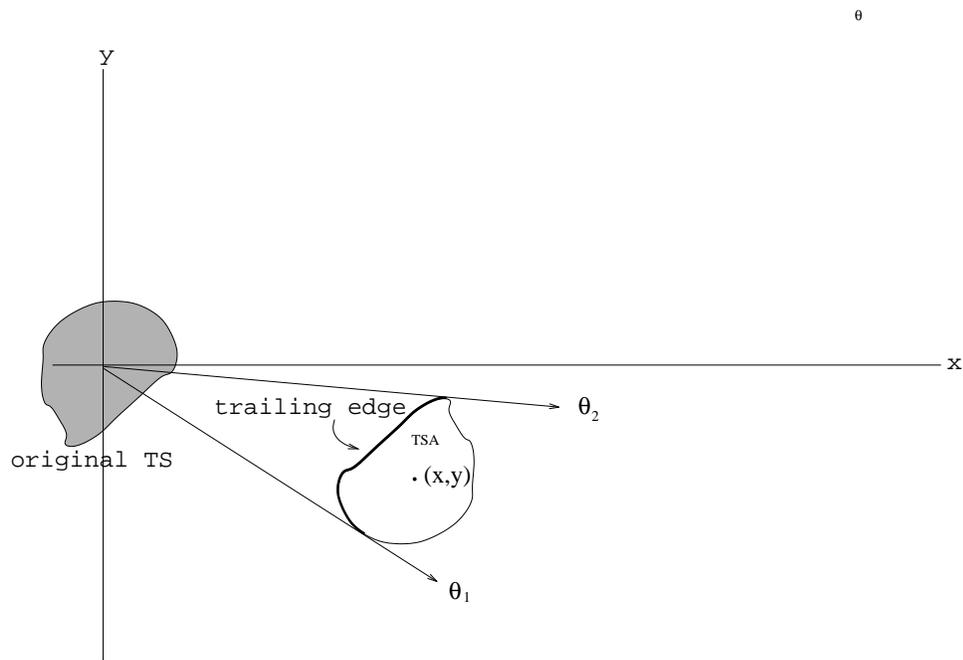


FIG. 8. Plan in Cartesian coordinates of the original thunderstorm (in grey), the thunderstorm area (TSA) for a position  $(x, y)$ , and the “trailing edge” (thick line) that contributes all the probability to the sum for that position.  $\theta_1$  and  $\theta_2$  are the angular limits of the trailing edge.

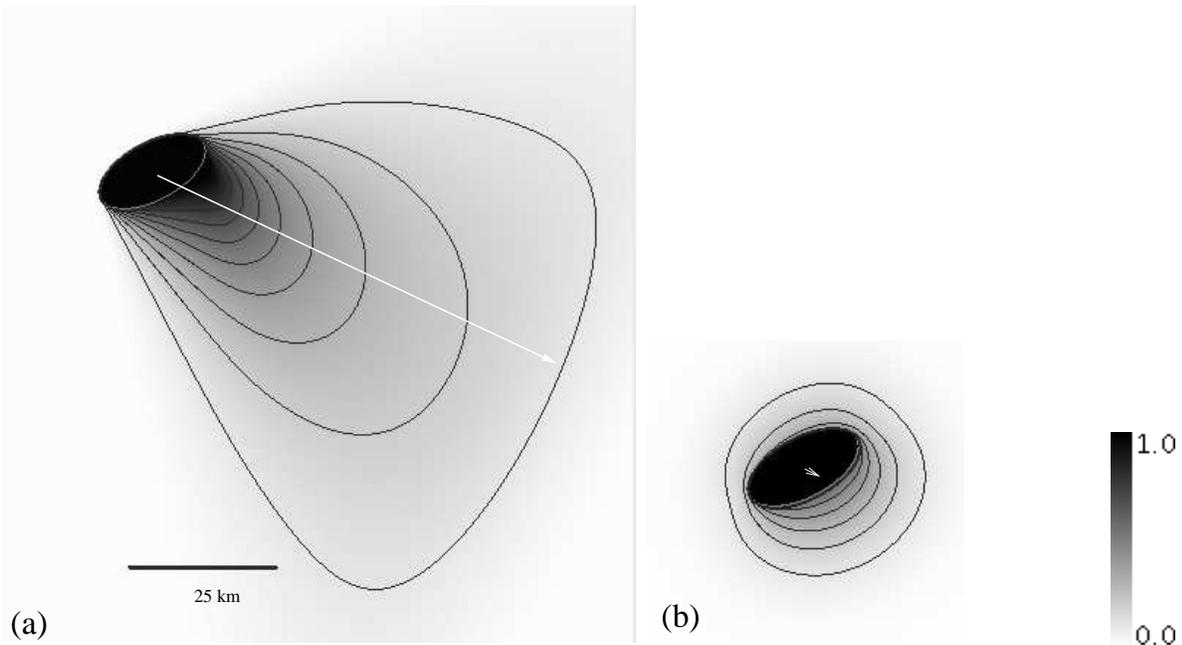


FIG. 9. Examples of output from THESPA with feasible storm tracking values. Figure (a),  $\theta = 115^\circ$ ,  $\sigma_\theta = 30^\circ$ ,  $V = 75 \text{ km h}^{-1}$ ,  $\sigma_V = 10 \text{ km h}^{-1}$ , figure (b),  $\theta = 115^\circ$ ,  $\sigma_\theta = 30^\circ$ ,  $V = 4 \text{ km h}^{-1}$ ,  $\sigma_V = 10 \text{ km h}^{-1}$ , with probability contours shown at 0.1 intervals. The original storms are shown as ellipses with nowcast velocity as an arrow. A scale marker shows  $25 \text{ km}$ . Black shading represents a strike probability of 1, white shading represents a strike probability of zero, as indicated by the grey-scale bar on the right.

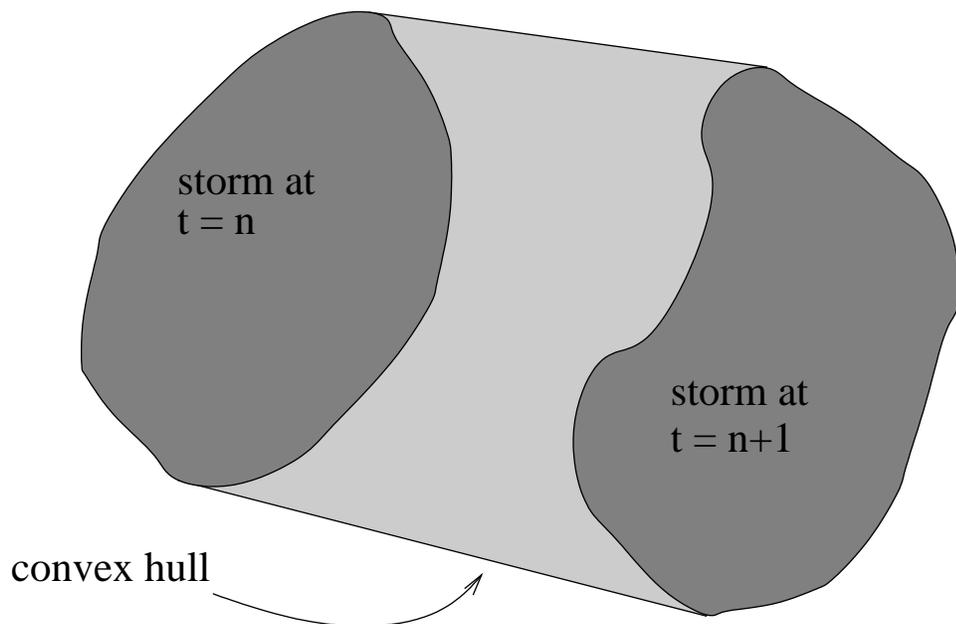


FIG. 10. Schematic diagram showing a pair of successive storm detections (dark grey) at time steps  $t = n$  and  $t = n + 1$ . For verification we wish to include not only the points under detected storms (dark grey), but also the points between detections that were presumably under the storm between radar scans (light grey). Hence for verification we include all points in the convex hull (shown in both shades of grey) of the two detections.

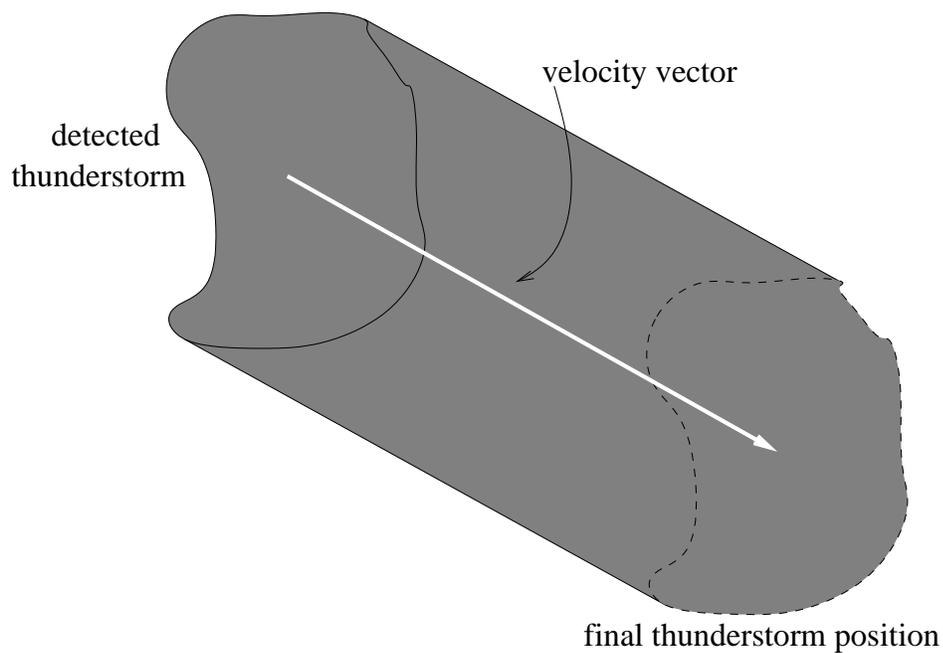


FIG. 11. Schematic diagram showing the advected threat area forecast modelled on current manual practice which is used as the reference forecast in calculating the Brier skill score (BSS). Here a detected thunderstorm is swept forward along the detected velocity vector to a new position at the end of the nowcast period, with the swept out area defined as having a thunderstorm strike probability of one, and the rest zero.

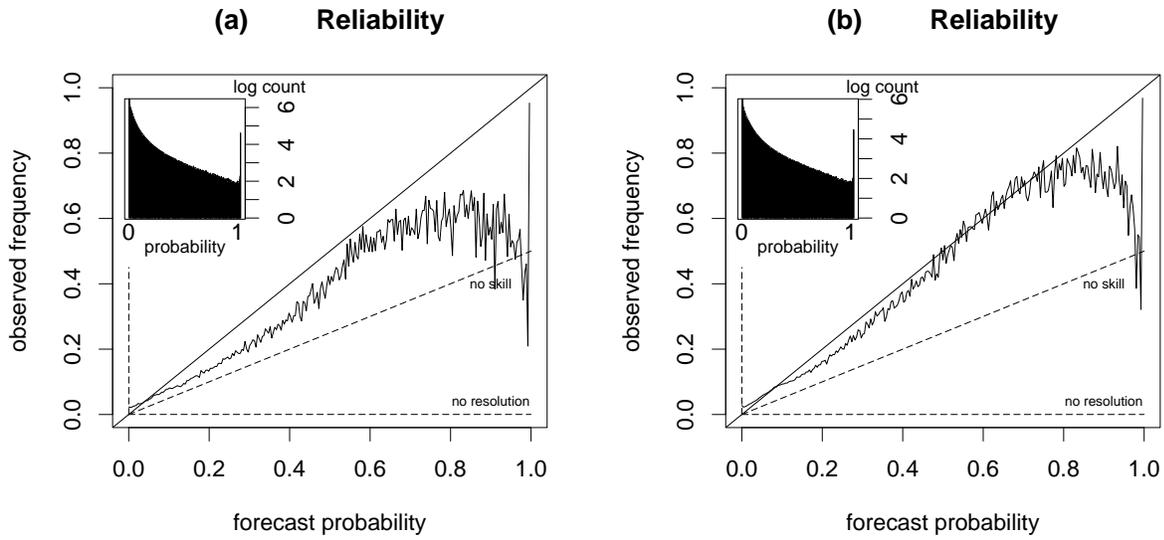


FIG. 12. Reliability diagrams for thunderstorm strike probability computed from the original 1682 Sydney TITAN storm tracks with  $\sigma_V = 10 \text{ km h}^{-1}$  and  $\sigma_\theta = 30^\circ$ , for (a) all thunderstorms, and (b) only those storms with VIL greater than  $30 \text{ kg m}^{-2}$ . The diagonal line indicates perfect reliability. The sharpness diagram at upper left on each chart shows the number of times each probability value (horizontal axis, linear scale) was predicted (vertical axis, logarithmic scale). The BSS for (a) was 0.31 and for (b) was 0.29, with respect to the advected threat area forecast.

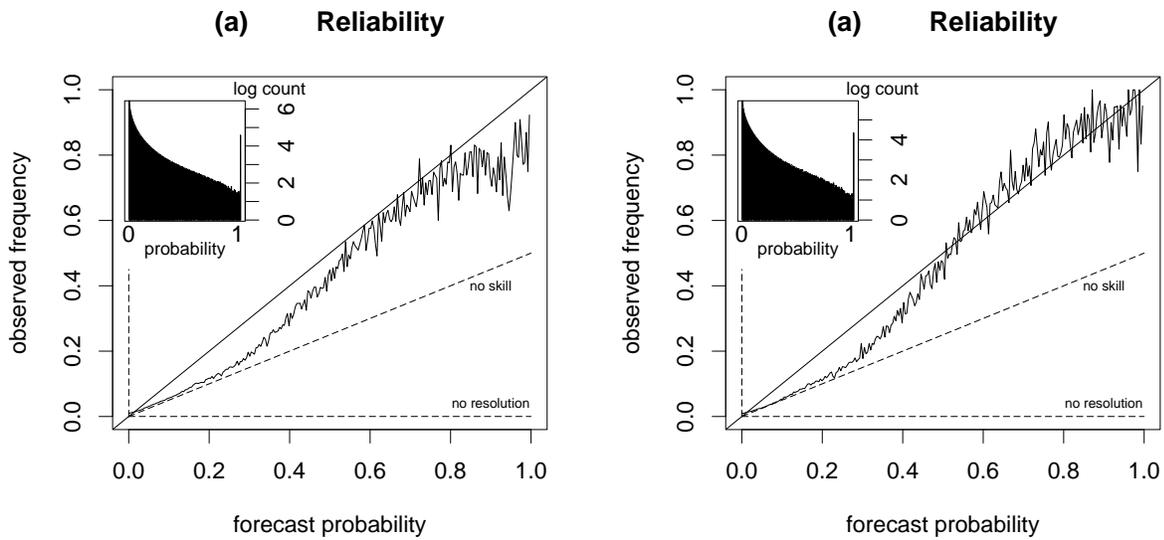


FIG. 13. As in Fig. 12, for THESPA strike probabilities generated from TITAN nowcasts collected during the WWRP Beijing 2008 Forecast Demonstration Project using a storm detection threshold of 45 dBZ (2394 thunderstorm tracks). This diagram includes the “no skill” and “no resolution” lines calculated from the climatology of thunderstorm occurrence. The BSS for (a) was 0.39 and for (b) was 0.41, with respect to the advected threat area forecast.

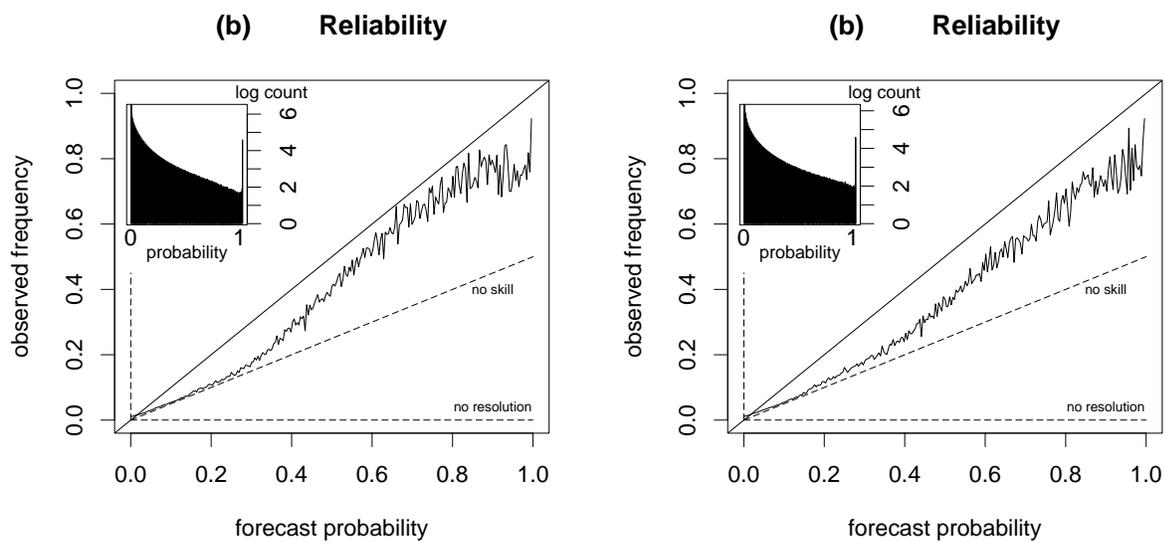


FIG. 14. As in Fig. 13a, with smaller assumed (a) speed error  $7\text{km h}^{-1}$ , and (b) direction error  $20^\circ$ . The BSS for (a) was 0.40 and for (b) was 0.41, with respect to the advected threat area forecast.

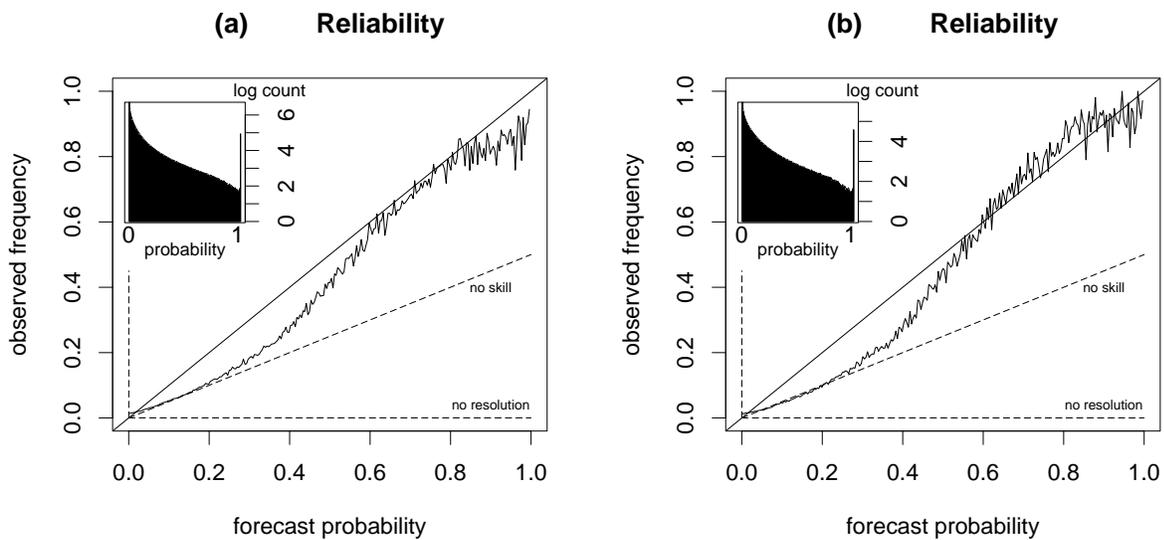


FIG. 15. As in Fig. 13a, with a storm detection threshold set to 40 dBZ (4244 storm tracks), for (a) all storms and (b) only those storms with VIL greater than  $30 \text{ kg m}^{-2}$ . The BSS for (a) was 0.38 and for (b) was 0.41, with respect to the advected threat area forecast.

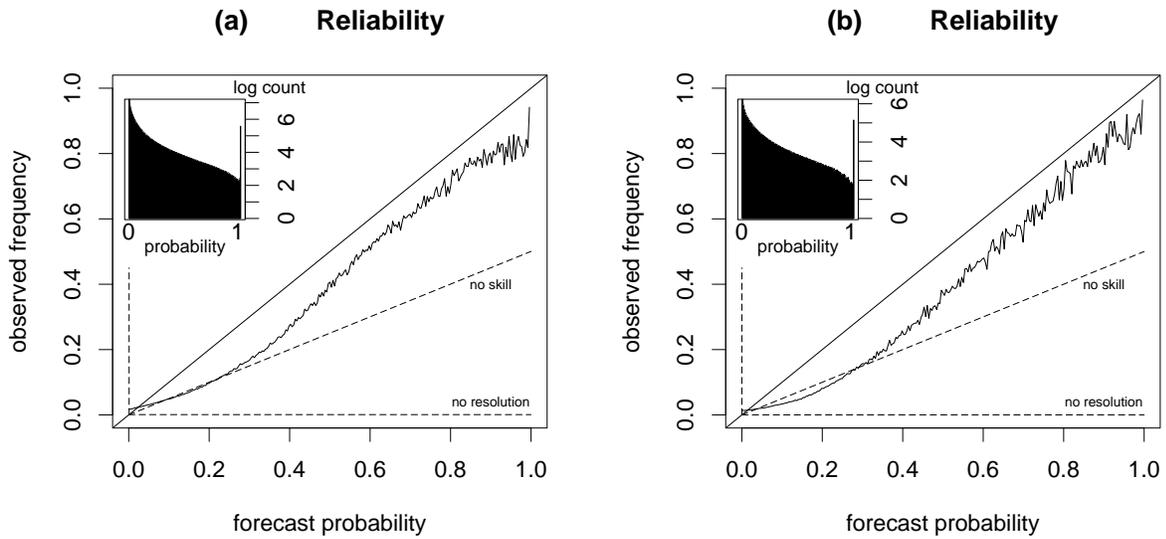


FIG. 16. As in Fig. 15, with a storm detection threshold set to 35 dBZ (12597 storm tracks). The BSS for (a) was 0.36 and for (b) was 0.44, with respect to the advected threat area forecast.

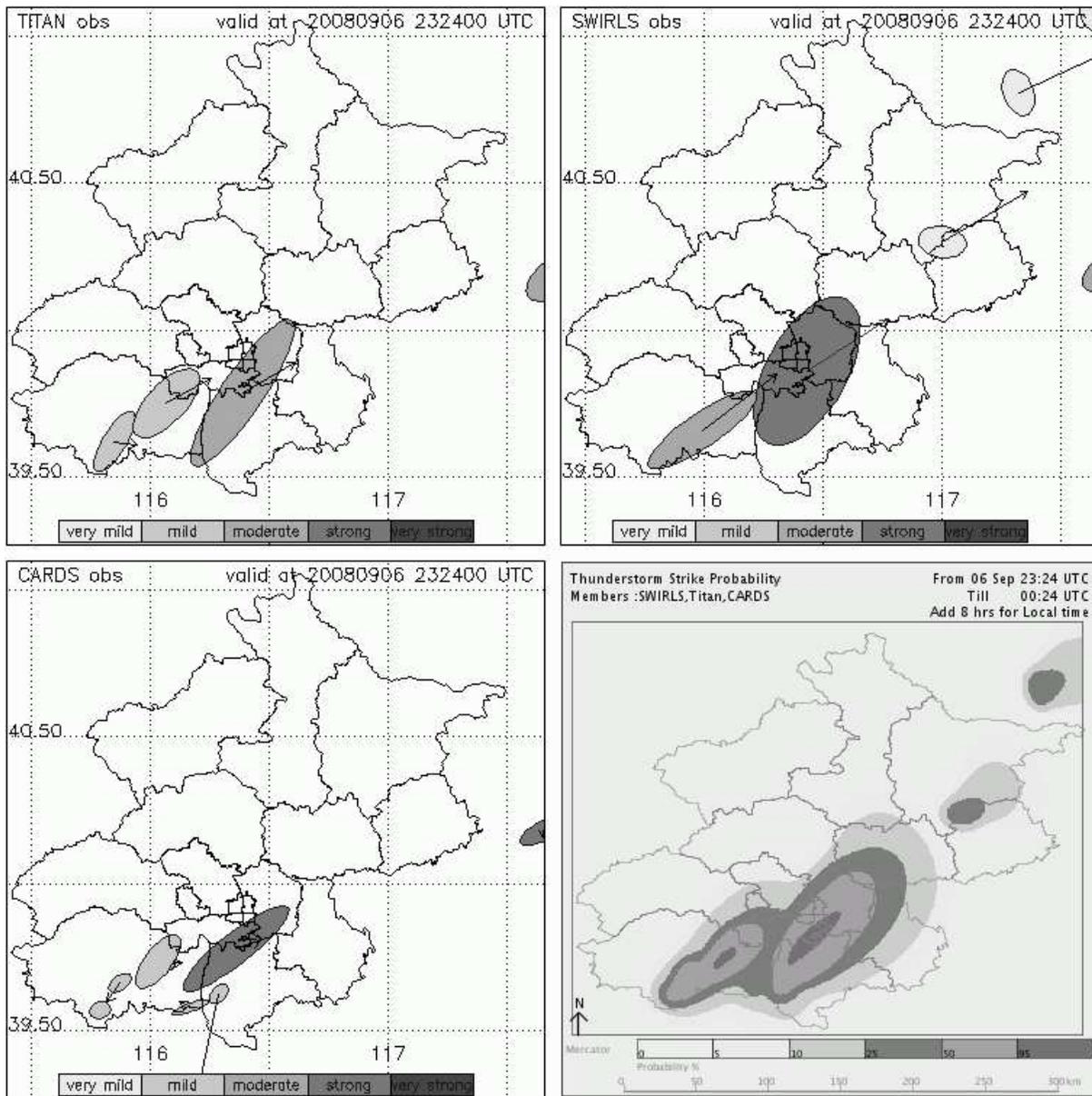


FIG. 17. Thunderstorm nowcasts from the TITAN, SWIRLS, and CARDS nowcasting systems, and consensus strike probability (lower right panel) for the hour beginning at 23:24 UTC on 6 September 2008. The arrows show the predicted movement of the thunderstorms during the next 60 minutes.