Aust. Met. Mag. 53 (2004) 269-277

Verification of operational cool-season tornado threat-area forecasts from mesoscale NWP and a probabilistic forecast product

Graham A. Mills
Bureau of Meteorology Research Centre, Melbourne, Australia

(Manuscript received January 2004; revised May 2004)

Cool-season tornado threat-area forecasts based on numerical weather prediction (NWP) model forecasts of vertical shear, thermodynamic instability, and convergence have been issued twice daily in Australia since the southern winter of 1999. This paper summarises a tiered verification of these operational forecasts for the four cool seasons from 1999-2002 inclusive. These verifications included 27 case studies, detailed model threat-area forecast climatologies, and ingredients verification. It is shown that in most cases useful forecast guidance was provided, that the model climatology has similar spatial and temporal structure to that observed, and that an assessment of the ingredients forecasts by means of contingency tables provides a benchmark objective verification that can be used to monitor product and numerical model forecast performance in predicting these rare, extreme events. A probabilistic threat-area product, based on mesoscale numerical weather prediction model output and verifying radiosonde data for the 2000-2002 seasons, to identify areas where environmental shear and instability are such that there is potential for cool-season tornado occurrence is described. These probabilities are conditioned on the model climatology, and so show a continuous field of probability rather than the earlier categorical threat forecast maps. Testing the system using independent data for the cool season of 2003 shows useful skill.

Introduction

Hanstrum et al. (2002) show from proximity soundings that cool-season tornadoes in southern Australia occur in environments characterised by relatively weak, but still positive, buoyancy in the lower troposphere, and strong wind shear from the surface to 1 – 1.5 km. Threshold criteria based on these proximity soundings (hereafter termed ‘proximity thresholds’) showed that a surface to 700 hPa lifted index* (SLI7) less than –0.5° C, and a 10 m – 850 hPa shear greater than 10^{-2}

Corresponding author address: Dr G.A. Mills, Bureau of Meteorology Research Centre, GPO Box 1289K, Melbourne, Vic. 3001, Australia.
Email: g.mills@bom.gov.au

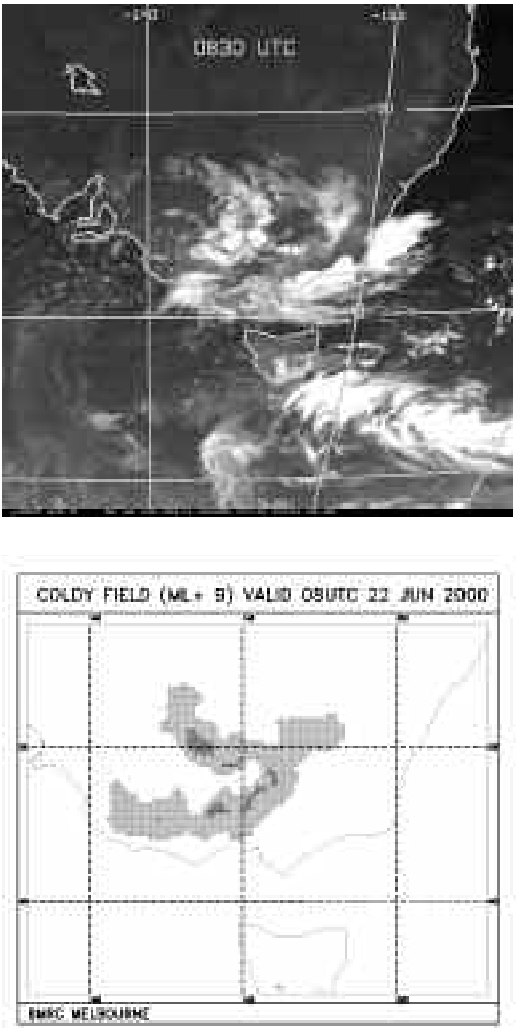
* The surface lifted index (SLI) to a given level is the difference between the temperature at that level and the temperature of a surface parcel (with a known temperature and dew-point) that is lifted dry-adiabatically to saturation, and then saturated adiabatically to the given level.

sec^{-1} were necessary environmental ingredients before such storms were observed. The values of these criteria are consistent with proximity soundings for similar tornado events in California (Lipari and Monteverdi 2000) and over the Great Plains of the USA in the northern hemisphere cool season (Johns et al. 1993).

The Australian Bureau of Meteorology operational mesoscale NWP system, LAPS (Puri et al. 1998), was shown to have some skill in forecasting areas where shear and buoyancy proximity thresholds were observed to be exceeded so long as the model thresholds were adjusted to account for model bias. Using these 'forecast thresholds' and with the addition of a convergence criterion, an operational threat-area forecast based on the Bureau's meso-LAPS model has been issued twice per day by the National Meteorological and Oceanographic Centre of the Bureau of Meteorology since the winter of 2000. A typical product is shown in Fig. 1, for 22 June 2000, the evening of the Kyabram (approx 36.4°S 145.4°E) tornado in north-central Victoria. Two sets of model thresholds are set, one less stringent than the other, to provide some allowance for model error and for case-to-case variation. As detailed by Hanstrum et al. (2002) the less stringent thresholds were selected to maximise the probability of detection for the events in the development dataset while maintaining the physical basis implied by the observational studies. The second set of thresholds was designed to more closely match the proximity sounding thresholds and reduce the size of the outlined areas, reducing the false alarm ratio at the expense of potentially missing some events. The areas outlined by these two sets of thresholds will be referred to as weak and strong threat areas in the remainder of this paper.

Verification of forecasts of rare events can be problematic. In verifying these forecasts a three-tier approach is used. First, the performance of the system in 27 known cases of cool-season tornadoes during the four cool-seasons of 1999 to 2002 is assessed. Second, the climatology of the forecasts is presented to assess whether over-forecasting influenced the case-study assessment. Third, the individual forecast ingredients, shear and SLI7, are directly verified against observations at radiosonde locations. It will be shown that, in spite of the difficulties inherent in verifying rare events, the forecast product usefully discriminates environments in which cool season tornadoes could occur. There are, however, issues associated with the categorical thresholds used in the assessment, and the verification dataset used in the ingredients verification is used to provide a continuous probability function based on the meso-LAPS model climatology that can replace the two-level threat-area product shown in Fig. 1.

Fig. 1 **GMS infrared satellite imagery at 0830 UTC 22 June 2000 (top), and fields showing forecast areas of cool-season tornado potential at the same time from operational 0.125° meso-LAPS forecasts. Light shading indicates the weak thresholds have been exceeded, dark shading indicates the strong thresholds have been exceeded.**



Verification of threat-area forecasts for the 1999-2002 seasons

It is difficult to devise completely satisfactory verification strategies for the forecasts of the environments in which small-scale, rare events such as cool-season tornadoes occur. First, the reporting of these very small-scale, short-lived events is moot away from population centres. Second, it is not the tornadoes themselves that are being predicted, but rather the ingredients (shear, buoyancy, and convergence) that define an environment in which tornadoes may develop. Accordingly, a three-tier verification strategy was developed based first on a subjective assessment of model performance in cases of reported cool-season tornadoes, second on the relation of the model's climatology of threat-area forecasts to observed occurrence of cool-season tornadoes, and third to a verification of the model's forecasts of shear and SLI7 against radiosonde data. The verification period was from 1999-2002 for the LAPS 0.375° system, and 2000-2002 for the meso-LAPS 0.125° system. It was decided not to undertake the model verification earlier than 1999 for LAPS, and 2000 for meso-LAPS in order that the model data had a common vertical resolution and boundary-layer parametrisation. It was, however, considered worthwhile to use the extra year of data available from the coarser grid to broaden the verification database.

The primary requirement for a system that forecasts cool-season tornado potential environments is that the system correctly identifies such an environment at the time and place when such an event occurs. Using listings from the web sites maintained by the Severe Weather Section of the Western Australian Regional Office, and the Severe Weather Summaries issued monthly by the Bureau of Meteorology Services Policy Branch (now Weather and Oceanographic Services Policy Branch), 27 occurrences of tornadoes in Australia between 15 May and 15 September for 1999, 2000, 2001 and 2002 were noted. These are listed in Table 1, together with the lead time of the forecast, the State in which the event occurred, a comment on the synoptic situation and a subjective assessment of the usefulness of the forecast. In order to keep the assessment tractable, it was decided to only assess one forecast for each event. A lead time of 9 to 24 hours was considered appropriate, since it is intended that this forecast of threat areas acts as an alert to the forecast office, not as a nowcast tool, and so there needs to be sufficient lead time for the model run to complete and the guidance to be available in the forecast office some hours prior to the event. Equally, it is an extreme, small-scale event being forecast, and so shorter lead times are more attuned to operational practice.

Table 1. List of events cool-season tornado events for the years 1999-2002. See text for details.

Case No.	*Date*	*Lead time (h)*	*State*	*Synoptic character*	*Model success*
1	17/6/1999	12	WA	Active frontal cloudband/cold-pool convection	Good
2	22/6/1999	18	WA	Active frontal cloudband	Good
3	13/7/1999	18	WA	Strong cold-pool convection	Good
4	14/7/1999	18	WA	Cold-pool convection	Fair
5	18/7/1999	21	SA	Convective bands in strong cyclonic flow	Good
6	25/8/1999	24	WA	Convective band in strong cyclonic flow	Fair
7	2/9/1999	18	WA	Convective band in active westerly trough	Good
8	20/6/2000	15	WA	Convective band in active westerly trough	Fair
9	21/6/2000	9	Vic	Strong northwesterly flow – negatively tilted squall-line	Good
10	22/6/2000	9	Vic	Convection – trough in strong westerly flow	Excellent
11	30/6/2000	9	WA	Active convective development in frontogenetic region	Fair
12	13/7/2000	12	WA	Active convective line and later cold-pool convection	Good
13	20/7/2000	24	SA	Convective line in strong cyclonic westerly flow	Good
14	24/7/2000	9/18	SA	Open-cell cumulus in large cold-air field	Good
15	2/8/2000	21	WA	Comma-cloud development in cold-air field	Fair
16	8/8/2000	21	Vic	Convective development in surface trough development	Good
17	10/8/2000	15	SA	Trough to west of mature low	Poor *
18	9/9/2000	21	Vic	Negatively tilted convective band	Good
19	25/7/2001	12	NSW	Easterly dip on NSW coast	Poor *
20	7/9/2001	18	SA	Convective line in surface trough development	Good
21	8/9/2001	21	SA	Trough to west of mature low	Good
22	18/5/2002	9/12	SA	Frontogenetic cloudband in active trough	Good
23	13/6/2002	21	WA	Convection in frontogenetic region	Good
24	15/6/2002	15	SA	Convective cluster ahead of occluding deep low	Excellent
25	24/7/2002	15	WA	Developing frontal cloudband	Excellent
26	3/8/2002	18	Vic	Convection developed to rear of upper low	Good
27	12/8/2002	9	Vic	Trough to west of mature low	Good

The quality of the forecast guidance was assessed as excellent if the forecast fields showed a strong correspondence to the evolving intensity of the convection as interpreted from satellite imagery, and also to be accurately located in time and space. Forecasts were assessed as good if there was little phase error, but perhaps some error in evolution – that is, the size of the convective areas varied compared to the size of the forecast areas. A fair forecast was one that did indicate that the weather system was one in which cool-season tornado environments were possible, but the forecast showed phase, scale, or evolution differences. Finally, a poor forecast was one in which there was no threat-area forecast, but a tornado was reported. Of the 27 forecasts assessed, three were rated excellent, 17 good, five fair, and two poor. (It must be acknowledged that some of the shorter lead-time forecasts would have been rated at least one category lower if the 12-hour earlier (longer lead-time) forecast had been used.) In only two of the assessed forecasts was there no indication of a cool-season tornado environment, so in 25 out of 27 cases there was a forecast that the significant regional weather system which did result in a tornado was likely to produce areas where the ingredients in that environment would support tornadoes, and 20 of the 27 forecasts produced good to excellent guidance. The two cases in which no signal was identified were convectively weak cases as assessed by the GMS-5 infrared satellite imagery, but inspection of the ingredients involved in the decision making show a band of enhanced shear in the correct area at the correct time, but with lesser intensity than the forecast threshold used. Mills (2004) presents forecast threat areas, forecast fields of ingredients, and satellite imagery near the time of occurrence of the tornado, and a short commentary, for each of the 27 cases.

While the case-study assessment above shows a satisfactorily high probability of detection, it is necessary to show that this hit rate has not been achieved simply by over-forecasting. However, due to sampling and reporting issues, it is difficult to devise a totally satisfactory null case assessment, and so indirect methods are needed to address this issue. Hanstrum et al. (2002) showed that the distribution of observed cool-season tornadoes favoured areas over the far southwest of Western Australia and over coastal southeastern South Australia, and that the climatological distribution of observed simultaneous shear and buoyancy proximity threshold exceedances at radiosonde observation sites closely matched this distribution (see their Figs 2 and 15). If the number of forecasts of weak or strong threat at each forecast grid-point per season is calculated, then this model climatology can be compared with Hanstrum et al.'s distributions. These were calculated for each model base time (0000 and 1200 UTC), forecast interval, and model resolution (LAPS/meso-LAPS), and year. As an example, summary plots showing the mean fields for the 12-hour and 24-hour LAPS forecasts valid at 0000 UTC and at 1200 UTC for 1999-2002 are presented in Fig. 2. These show maxima of the same order and the same geographic distribution as seen in the analysis of the climatological radiosonde data record and of observed cool-season tornadoes presented in Hanstrum et al. (2002). This indicates that the diagnosed threat areas are not forecast excessively frequently at any individual grid-point, and thus provides some evidence that the high rate of detection seen in the case study assessment above has not been achieved purely by over-forecasting. Mills (2004) also shows that subtle year-to-year variations in these patterns can be related to interannual regional circulation differences and also bear some relation to the numbers of reported tornadoes.

The above two assessments, while satisfactory in themselves, are not strictly objective against a directly comparable variable. One objective and reproducible verification can be developed if it is remembered that it is the ingredients – shear, SLI7, and vertical motion – that are being forecast, rather than the occurrence of tornadoes. The first two of these ingredients can be verified directly against radiosonde observations, although this does have the limitation that it can only be achieved at radiosonde locations and at one or two times per day (and these times might not coincide with the peaks in tornado occurrence). Table 2 shows contingency tables for all 12-hour and 24-hour meso-LAPS forecasts valid 0000 or 1200 UTC at the 10 radiosonde stations shown in Fig. 3, and satisfying the weak and strong joint model thresholds, compared with the radiosonde data satisfying the joint proximity thresholds of shear > 15 m s^{-1} and SLI7 < -0.5° C. A model forecast was counted if any of the 3x3 array of grid-points over the station indicated a threat.

The first conclusion is that the overwhelming majority of cases are correct null forecasts, as would be expected for rare events. Second, there is a tendency to over-forecast with the weak thresholds, which is expected (and was intended, see Hanstrum et al. 2002) and to underforecast with the strong thresholds. Overall, the Probability of Detetection (POD)/False Alarm Ratio (FAR) (see, for example, Mason (2003) and Ebert and McBride (1997) for definitions and discussion) for weak threat are 0.50/0.89, and for strong threat 0.15/0.75. These verifications perhaps present a somewhat more pessimistic interpretation of the threat-area forecasts than did the case-study assessment. However, the contingency table verifications of the joint shear/lifted index threshold exceedances have

Fig. 2 **Climatology of LAPS grid-point forecasts for weak and strong threat for the years 1999-2002. Forecasts valid 0000 UTC are all 12 and 24-hour forecasts valid at 0000 UTC, forecasts valid 1200 UTC are all 12 and 24-hour forecasts valid at 1200 UTC. Contour interval for weak hits is 8, with a lowest contour of 2, for strong hits 1 with a contour interval of 1.**

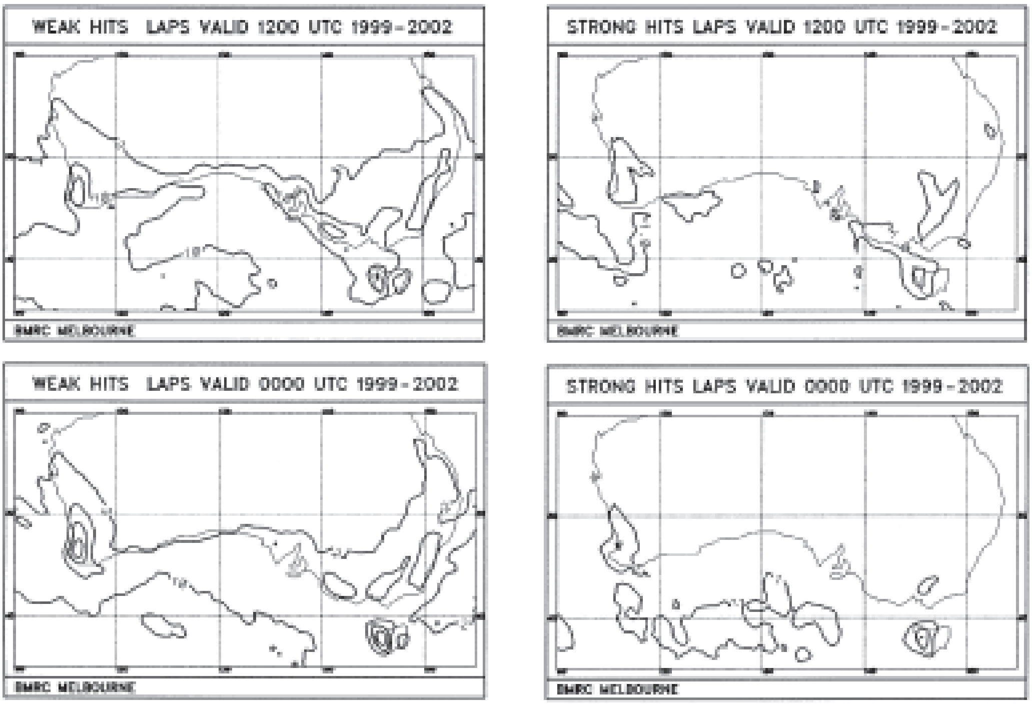


Table 2. **Contingency tables for forecasts from the meso-LAPS model for the seasons of 2000-2002 for forecasts when joint forecast thresholds are exceeded verified against radiosonde data using proximity thresholds for the 10 stations shown in Fig. 3.**

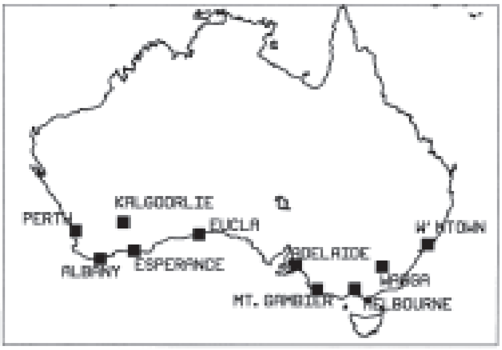
	Weak thresholds		*Strong thresholds*	
	Model No	*Model Yes*	*Model No*	*Model Yes*
Obs No	13 309	727	13 949	87
Obs Yes	91	90	153	28

similar levels of performance to those in the very limited published performance of tornado watch forecasts (see the discussion in Hanstrum et al. (2002)). The great advantage of such a verification strategy is that it is objective and reproducible, and so can be used to monitor and document system performance and potentially model improvement. A limitation is that the thresholds used are 'hard' – there is no reward for a narrow miss, either in time, space, or magnitude. In addition, there is considerable variability in model performance when measured against the different radiosonde locations, and one must acknowledge that there may be errors of representivity here. These factors may contribute to the perception that the pattern-recognition approaches of the case-study verification appear rather more positive than the ingredients verifications presented in this section. Mills (2004) presents contingency tables for individual stations, and also direct verification plots of the individual ingredients.

Probability threat-area forecasts

It is evident from the case studies and composite fields shown in Hanstrum et al. (2002) (see their Figs 16, 17, 18 and 19) and the 27 case studies in Mills (2004) that the shear and instability maxima are usually not co-located – there is some offset. Thus the areas where the thresholds are jointly exceeded (that is, the threat areas) are often in the overlap zone between the instability maximum and the shear max-

Fig. 3 **Locations of radiosonde stations used in ingredients verification.**



imum rather than coinciding with the individual centres, and these are frequently areas of marked horizontal gradient in the individual ingredients. This characteristic of the guidance makes the precise areas marked as threats very sensitive to the choice of thresholds, and also to model error in the forecasts of the individual ingredients, although the specification of two levels of threat does alleviate this issue somewhat. Further, the exact POD/FAR values from the third verification described above are sensitive to the values of these categorical thresholds. It is desirable to produce a more continuous function that avoids these issues.

One approach that has been used over many years to produce a continuous function indicating severe weather likelihood is to produce an 'index'. However, specifying an 'index' such as the Supercell Composite Parameter, or the Significant Tornado Parameter (Thompson et al. 2002) requires that an appropriate normalisation of each parameter be applied, and on a physical basis it is difficult to determine the appropriate bias and multiplier used in this normalisation, and also the appropriate weight to be given to each normalised parameter in calculating the index.

The matched sets of forecast and observed shear and SLI7 used to produce the ingredients verification do provide an alternative approach. The desired forecast quantity is whether, at a grid-point, the proximity shear and lifted index thresholds will both be exceeded. Using the three-year verification dataset, then for all forecasts with values of shear and lifted index that exceed specified values, the frequency with which the proximity thresholds of 15 m s^{-1} and –0.5°C are exceeded can be calculated. As an example, for Perth over the three cool seasons of 2000-2002, a shear greater than 10 m s^{-1} together with a SLI7 less than -0.5° C was forecast 187 times of the possible 1471 12-hour and 24-hour meso-LAPS forecasts based at either 0000 or 1200 UTC. Of those 187 forecasts, the proximity thresholds were exceeded in the verifying radiosonde data on 41 occasions. If the shear thresholds and SLI7 thresholds are incremented in small steps, then a two-dimensional array is obtained, with each element of the array representing the model's conditional climatology of the probability that for a given shear and lifted index threshold, the proximity thresholds will be exceeded. These can then be applied to the forecast fields to obtain a contoured probability field.

This process was applied to the 12-hour and 24-hour forecasts based at both 1100 and at 2300 UTC for Perth and Adelaide (the stations with the greatest numbers of observed events) and probabilities in bins of 2 m s^{-1} shear and 0.5°C lifted index were calculated (see Table 3). These probability values are approximately monotonically increasing both with increasing shear and with increasing instability, as would be intuitively expected, apart from a few bins containing low numbers of events (particularly at the higher shear/lower SLI7 ranges) where there were some anomalies. It is found that the weak and strong thresholds used in the operational product correspond to probabilities of approximately 25 and 50 per cent respectively.

After applying a simple interpolation procedure to force a monotonic increase in probability through the array, and (arbitrarily) imposing lower/upper limits of 1 and 85 per cent, the model forecast fields can be used to generate a contoured probability field, rather than the two-level shaded threat areas as shown in Fig. 1. An example is shown in Fig. 4, the case of the Kyabram tornado of 22 June 2000. These forecasts have the great benefit of producing a continuous field indicating threat areas, and the waxing and waning of the zone of highest probability can be clearly seen, peaking near Kyabram late on the afternoon of 22 June (the third panel can be compared with Fig. 1). The vertical motion (the convergence ingredient necessary for convection) has not been included in the probability function, but in Fig. 4 the areas of low-level ascent are shaded to allow some visual incorporation of this information. An additional benefit is that areas where there is a probability below 25 per cent are still identified, providing useful forecast insight in some of the (forecast) environmentally weaker cases shown in Mills (2004). Indeed, the two cases (2 May 2000 near Albany, WA, and 24 July 2001 on the central coast of NSW) where no threat areas were forecast, but tornadoes were reported, each show probabilities greater than 15 per cent in coherent zones over the reported tornado locations.

Table 3. Probability (percent) that if a meso_LAPS forecast of shear (m s^{-1}) is greater than the listed values, together with a SLI7 (°C) less than a given value, that the proximity thresholds of 15 m $^{-1}$ and –0.5°C will be exceeded. Data is for all 12 and 24-hour meso-LAPS forecasts based at 0000 and 1200 UTC for the cool seasons of 2000, 2001 and 2002. Values are bold if there were less than 20 forecasts in a category. The asterisk indicates no forecasts in this range.

Shear *SLI7*	*0.0*	*2.0*	*4.0*	*6.0*	*8.0*	*10.0*	*12.0*	*14.0*	*16.0*
4.0	4.1	4.2	4.8	6.0	9.1	14.3	17.1	25.9	**16.7**
3.5	4.3	4.5	5.0	6.2	9.4	14.5	17.3	27.3	**16.7**
3.0	4.7	4.8	5.4	6.6	9.9	15.5	18.7	31.3	**20.0**
2.5	5.0	5.2	5.8	7.0	10.3	16.1	19.5	31.9	**21.4**
2.0	5.5	5.6	6.2	7.4	11.0	17.6	22.5	39.5	**33.3**
1.5	6.0	6.1	6.7	8.1	11.8	18.7	24.2	45.5	**37.5**
1.0	6.7	6.8	7.3	8.8	12.7	20.5	27.8	51.9	**42.9**
0.5	7.3	7.4	8.0	9.5	13.5	21.9	30.4	60.9	**60.0**
0.0	8.4	8.4	9.0	10.4	14.6	23.2	32.8	**60.9**	**60.0**
-0.5	8.6	8.6	9.0	10.2	14.7	23.0	35.4	**70.6**	**66.7**
-1.0	8.8	8.8	9.2	10.0	14.7	23.1	35.0	**73.3**	**66.7**
-1.5	10.8	10.8	11.3	12.4	18.4	27.9	46.4	**83.3**	**100.0**
-2.0	13.3	13.3	14.1	15.6	22.2	33.3	**58.3**	**80.0**	**

It is most desirable that forecast products are objectively verifiable, and the probability forecasts can be verified by means of the Relative Operating Characteristic (ROC) as described by Mason (2003) and Mason and Graham (1999). The ROC graph is generated by issuing a forecast of an event from a probabilistic forecast if a pre-defined threshold probability is exceeded. A series of hit and false alarm rates can be calculated for different probability thresholds, and plotting these hit rates against false alarm rates generates the ROC curve. In general, for a skillful forecast the hit rates are larger than the false alarm rates, and the ROC curve bends towards the top-left corner of the graph. As the diagonal line on the graph indicates no skill, the further the ROC curve deviates above the diagonal, the more skillful the forecast, and the area under the ROC curve has been used as an indicator of forecast skill. Forecast values of SLI7 and of shear were matched with observed data from the 10 radiosonde stations shown in Fig. 3 from 15 May to 15 September 2003 to provide an independent dataset with which to verify the probability forecasts described above, and Hit Rate and False Alarm Rate were calculated for a range of probability thresholds from five to 85 per cent in five per cent steps. These results are shown in Fig. 5 and the ROC curve is well above the diagonal 'no skill' line. The area under the ROC curve is 0.879, indicating a very useful degree of skill in these forecasts.

Discussion and conclusions

An assessment of the performance of the cool-season tornado threat-area forecasts issued by the Australian Bureau of Meteorology has been summarised for the four cool seasons from 1999-2002. Subjective assessment indicates the guidance has a useful level of skill, with most events being relatively well forecast if small timing or location errors are allowed. Climatological analysis indicates that the climatology of the forecasts is comparable to the observed climatology, and verification of the ingredients against radiosonde observations, while subject to some errors of representivity, does provide an objective, reproducible verification with which the system performance can be monitored, and the influence of NWP model improvements can be assessed.

A probabilistic form of threat-area forecast for cool-season tornado potential has been described in this paper. The probability forecasts are conditioned on the model climatology, as is the extreme forecast index of Lalaurette (2003) using the ECMWF ensemble forecast system, although those were univariate forecasts rather than the bivariate forecasts described in this paper. The probabilities thus do require that the model climatology does not change dramatically when model improvements are made, and so these should also be monitored during parallel testing to ensure smooth transition to later model versions. How restrictive this is will need to be determined.

Fig. 4 **Contoured fields of forecast probability of joint threshold exceedance for 3, 6, 9, and 12-hour forecasts for the Kyabram tornado event of 22 June 2000. Contours are at 15, 25, 40, 60 and 80% values. Shaded areas indicate ascent at 900 hPa. The X marks the position of Kyabram. These forecasts are based on the operational 0.125° grid spacing meso-LAPS NWP model.**

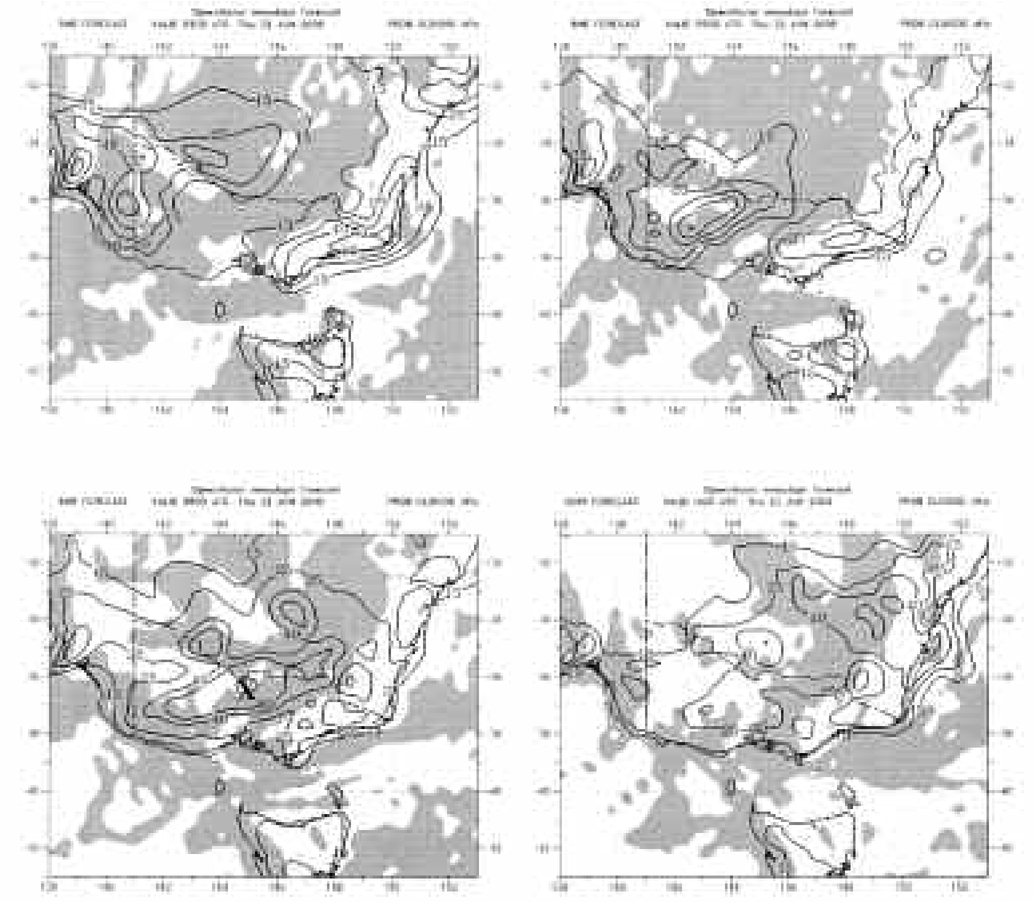
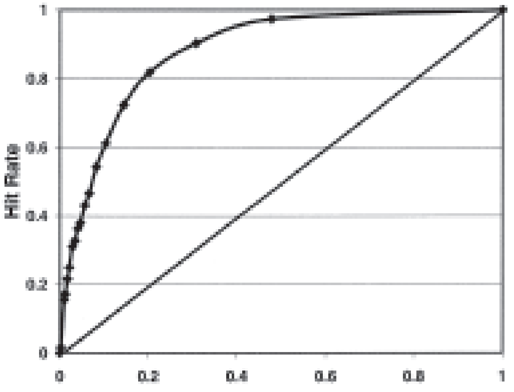


Fig. 5 **Relative Operating Characteristics (ROC) diagram for probability forecasts of cool-season tornado threat at 10 radiosonde locations during the cool season of 2003. The diamonds in the body of the figure mark probability thresholds from 5% to 85% at 5% intervals.**



The probability framework for indicating threat areas has several benefits over the previous product based on ‘hard’ thresholds. It is a continuous function and so does not suffer from the arbitrary yes/no decisions inherent in the earlier product, avoiding the issue of narrow misses or large areas of weak threat. It thus lends itself to verification methods such as the use of the ROC diagrams, enabling NWP system skill at forecasting rare and extreme weather phenomena to be objectively monitored without the limitations imposed by categorical thresholds. In addition, it is perhaps more clearly related to the environmental ingredients that determine the conditions under which cool season tornadoes can occur than to the occurrence of the tornadoes themselves, thus lessening the possibility that a forecast threat-area may be seen as a failed forecast if a tornado is not observed in the area. This approach could be extended to other model forecast guidance systems where threshold values are used, such as the thunderstorm decision tree described by Mills and Colquhoun (1998), or the Bureau of Meteorology's new National Thunderstorm Forecast Guidance System (Hanstrum 2003).

References

Ebert, E.E. and McBride, J.L. 1997. Methods for verifying quantitative precipitation forecasts: application to the BMRC LAPS model 24-hour precipitation forecasts. *BMRC Techniques Development Report No.2*. Available from BMRC, GPO Box 1289K, Melbourne, Vic. 3001, Australia. 87pp.

Hanstrum, B.N., Mills, G.A., Watson, A., Monteverdi, J.P. and Doswell, C.A. 2002. The cool-season tornadoes of California and Southern Australia. *Weath. forecasting, 17*, 705-22.

Hanstrum, B.N. 2003. A national NWP-based thunderstorm and severe thunderstorm forecasting guidance system. In "Current issues in the parameterisation of convection": extended abstracts of presentations at the fifteenth annual BMRC Modelling Workshop 13-16 October 2003. P.J. Meighen and A.J. Hollis, Eds. Pp7-11. *BMRC Research Report No. 93*. Available from BMRC, GPO Box 1289K, Melbourne, Vic. 3001, Australia. 124pp.

Johns, R.H., Davies, J.M. and Leftwich, P.W. 1993. Some wind and instability parameters associated with strong and violent tornadoes. 2: Variations in the combinations of wind instability parameters. The Tornado: Its Structure, Dynamics, Prediction, and Hazards, C. Church, D. Burgess, C. Doswell, and R. Davies-Jones (eds). *Geophys. Monogr. 79*, Amer. Geophys. Union, 583-90.

Lalaurette, F. 2003. Early detection of abnormal weather conditions using a probabilistic extreme forecast index. *Q. Jl R. Met. Soc., 129*, 3037-57.

Lipari, G.S. and Monteverdi, J.P. 2000. Convective and shear parameters associated with northern and central California tornadoes during the period 1990-94. *Preprints, 20th Conf. on Severe Local Storms*, Orlando, FL, Amer. Meteor. Soc., 518-521.

Mason, I.B. 2003. *Binary Events. Forecast Verification: A Practitioner's Guide in Atmospheric Science*. I.T. Jolliffe and D.B. Stevenson (eds), Wiley.

Mason, S.J. and Graham, N.E. 1999. Conditional probabilities, relative operating characteristics, and relative operating levels. *Weath. forecasting, 14,* 713-25.

Mills, G.A. 2004. Verification of operational NWP model cool-season tornado threat area forecasts in Australia. *BMRC Research Report No. 103*. Available from BMRC, GPO Box 1289K, Melbourne, Vic. 3001, Australia. 114pp.

Mills, G.A. and Colquhoun, J.R. 1998. Objective prediction of severe thunderstorm environments: linking a decision tree with an operational regional NWP model. *Weath. forecasting, 13*, 1078-92.

Puri, K., Dietachmeyer, G.D., Mills, G.A., Davidson, N.E., Bowen, R.A. and Logan, L.W. 1998. The new BMRC Limited Area Prediction System, LAPS. *Aust. Met. Mag., 47,* 203-23.

Thompson, R.L., Edwards, R. and Hart, J.A. 2002. Evaluation and interpretation of the Supercell Composite and the Significant Tornado Parameters at the Storm Prediction Centre. *Preprints, 19th Conference on Weather analysis and Forecasting and 15th Conference on Numerical Weather Prediction*. Amer. Met. Soc. August 2002, J11-J14.