15.5 **THE AUSTRALIAN NATIONAL THUNDERSTORM FORECAST GUIDANCE SYSTEM: CURRENT DESIGN, VERIFICATION AND FUTURE PLANS**

Harald Richter*
Centre for Australian Weather and Climate Research, Melbourne, Australia

1. INTRODUCTION

The prediction of the likelihood and severity of thunderstorms in Australia involves the examination of observations and numerical model predictions with the aim of identifying the basic convective ingredients – buoyancy (moisture and instability), lift and shear (Johns and Doswell 1992). With tight time constraints imposed on the operational delivery of day 1 and day 2 thunderstorm forecasts, it has long been desirable to employ a supporting system that skillfully suggests where to focus the human effort in the thunderstorm forecast process.

The production of thunderstorm forecasts (or outlooks) in Australia is the responsibility of the seven individual forecast offices (Fig. 1). Individual staff in these offices can have relatively limited exposure to thunderstorm forecasting, in contrast to staff in the Storm Prediction Center in the U.S. which has national responsibilities in a convectively very active region of the planet. A system that would alert a forecaster to potential thunderstorm threats that are easy to miss is therefore beneficial.

Fig. 1: Australian forecast offices including the associated states (or territories) that mark the area of forecast and warning responsibility for that office. Melbourne is the location of the Bureau's Head Office which is co-located with the Victorian forecast office.

Corresponding author address: Harald Richter, Centre for Australian Weather and Climate Research, PO Box 1289K, Melbourne VIC 3001, Australia; e-mail: h.richter@bom.gov.au.

This paper describes Australia's National Forecast Guidance System (NTFGS) which aims to address the two goals outlined above (Hanstrum 2004). Section 2 will focus on the current system design, section 3 will present some preliminary verification results and section 4 will outline plans for the future NTFGS development.

2. DESCRIPTION OF THE CURRENT NTFGS

The National Thunderstorm Forecast Guidance System (NTFGS) is a numerical model post-processing algorithm for diagnosing environments favourable for deep moist convection. It ingests operational NWP model output four times a day (06z, 12z, 18z and 00z). The model that currently feeds the NTFGS is ACCESS-A, which stands for "Australian Community Climate and Earth System Simulator - Australian domain" (Puri et al. 2012). At present (2012), ACCESS-A is based on the UK Met Office's Unified (global) Model version 6.4 and runs on the regional (Australian) domain shown in Fig. 2. The non-hydrostatic model is formulated on a 0.11 degree (~12 km) lat-lon grid with 50 vertical levels that utilize a hybrid vertical coordinate which is terrain-following near the ground and approaches a constant height surface in the lower stratosphere. The model runs four times a day and out to 48 hours. The model physics consist of

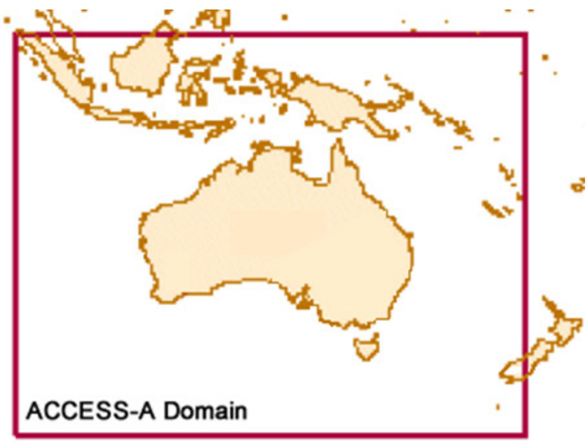

Fig. 2: Spatial domain of the ACCESS-A numerical model which produces the input to the NTFGS algorithm.

the Smith (1990) diagnostic cloud scheme, a modified Edwards and Slingo (1996) radiation scheme, a boundary layer scheme following Lock (1998, 2000, 2001), Wilson and Ballard (1999) microphysics, a convective scheme based on Gregory and Rowntree (1990) and Grant and Brown (1999), and the Moses-II land surface scheme (Essery et al., 2003). Data are assimilated using a 4D-Var scheme that runs every 6 hours. More details on the NTFGS parent model can be found in the operations bulletin http://www.bom.gov.au/nwp/doc/bulletins/apob83.pdf.

The algorithm currently provides operational forecasters at the Australian Bureau of Meteorology with a set of thunderstorm-related diagnostics across all of Australia out to two days in lead time. An example of the NTFGS hazard summary graphic for 17 October 2012 is shown in Figure 3:

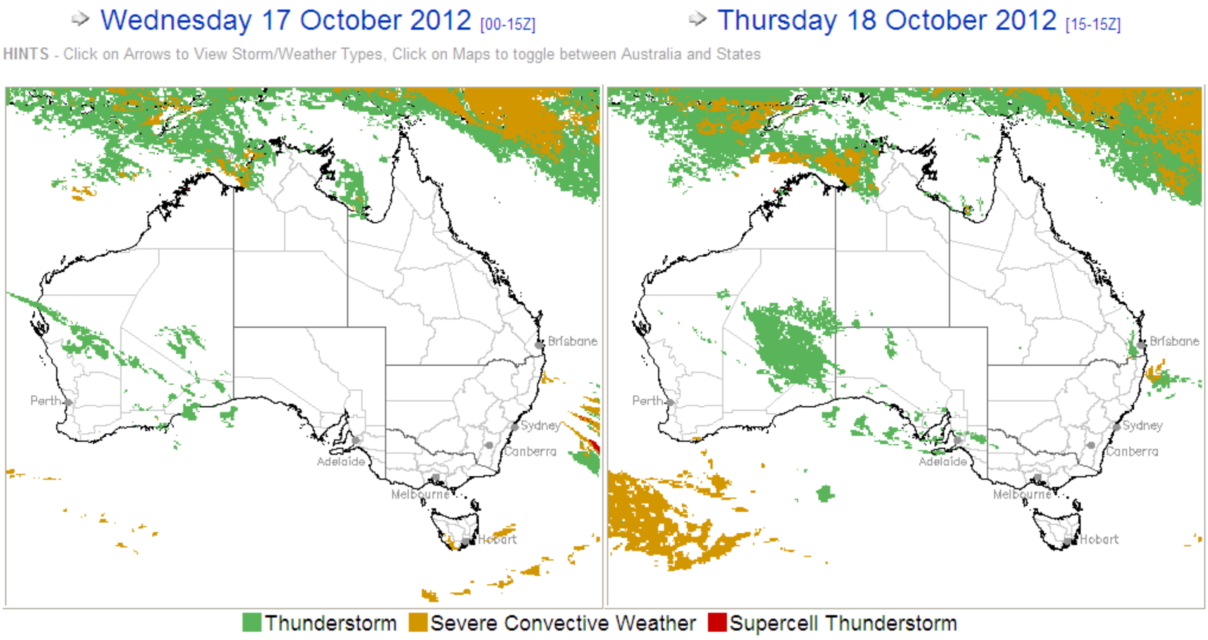

Figure 3: Summary of the NTFGS diagnostics for outlook day 1 (left) and outlook day 2 (right) for the randomly selected 17 October 2012 case.

The individual diagnostics within the NTFGS are produced by testing a defined group of ingredients for exceedance of ingredient-specific thresholds. The diagnostic switches from "not favourable" to "favourable" only when each contributing ingredient simultaneously exceeds its specific threshold. A subset of those diagnostics have three probability categories - not favourable, favourable and very favourable, where the distinguishing factor between the 'favourable' versus 'very favourable' categories lies in the ingredients satisfying higher threshold levels. An example of the graphical collation of all NTFGS diagnostic categories for 17 October 2012 is given in Figure 4.

Note that, at present, the categorical probabilities of occurrence (not favourable, favourable and very favourable) are not calibrated against any numerical probabilities. It is therefore likely that the occurrence probability "favourable" for tornadoes equates to quite a different actual probability of occurrence compared to "favourable" for surface-based thunderstorms. For most convective hazards in Australia, any such calibration effort quickly runs into a well known but very significant under-reporting problem away from the major cities due predominantly to the sparse population across most of the continent.

In the diagnostic summary in Fig. 3, thunderstorms are the aggregate of surface-based and elevated thunderstorms in Fig. 4, whereas severe convective weather includes damaging/destructive winds, large hail, tornadoes and heavy convective rain. Supercells carry over unchanged.

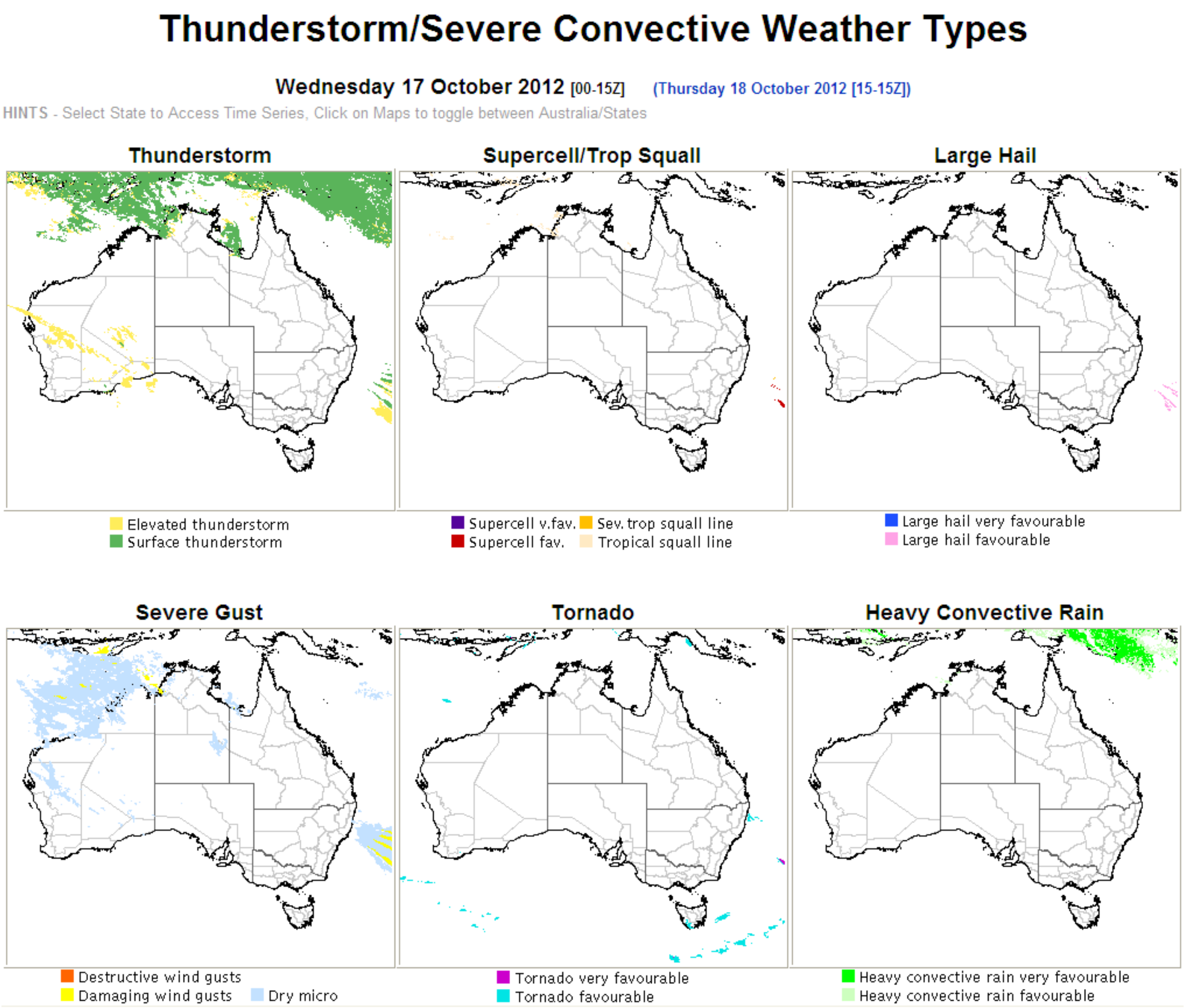

Fig. 4: Display of all available NTFGS diagnostics with their respective categorical probabilities of occurrence. The six panels show a disaggregation of the left panel in Fig. 3.

At this stage of the NTFGS output inspection the forecaster is equipped with location pointers that highlight areas in which specific storm types or specific convective hazards have an elevated probability of occurrence compared to the background. A crucial next step is the forecaster's interrogation of the contributing ingredients to the algorithm's diagnosis. As an example, Table 1 lists the five ingredients and corresponding thresholds for the surface-based thunderstorm diagnostic. Note that the evaluation of the lifted index (LI) for a 50 hPa mixed surface parcel distinguishes between a cool season and a warm season scenario, separated by the model temperature at 850 hPa exceeding 12 degrees Celsius.

The choice of ingredients is guided by an ingredients-based thinking: the conditional instability required for deep moist convection is incorporated through the lifted index, cloud electrification is parameterized through a cold

Ingredient	Threshold
LI(500) warm season T_{850}>=12C LI(700) cool season T_{850}<12C	<= -1
Equilibrium Level	<= -20 C
Max(omega) in sigma layer [0.9988, 0.8500]	>= 10 hPa/hr
|CIN|	<= 25 J/kg
Cold Cloud Depth	>= 3 km

Table 1: Ingredients that contribute to the NTFGS not favourable/favourable decision for the surface-based thunderstorm diagnostic. LI is the lifted index of a surface parcel mixed through the lowest 50 hPa, T_{850} is the air temperature at 850 hPa, omega is the model's vertical velocity in pressure coordinates (evaluated between two levels of the standard vertical sigma coordinate), and CIN is convective inhibition.

parcel equilibrium level combined with a requirement of a minimum cloud depth below zero degrees Celsius (or, cold cloud depth), and potential for convective initiation is expressed through criteria for low convective inhibition in conjunction with adequate boundary layer lift.

At this stage of the NTFGS development, the choice of ingredient thresholds has been guided by a limited number of local case studies (Hanstrum et al. 2002; Treloar and Hanstrum 2002) in conjunction with subjective tuning over a larger range of cases.

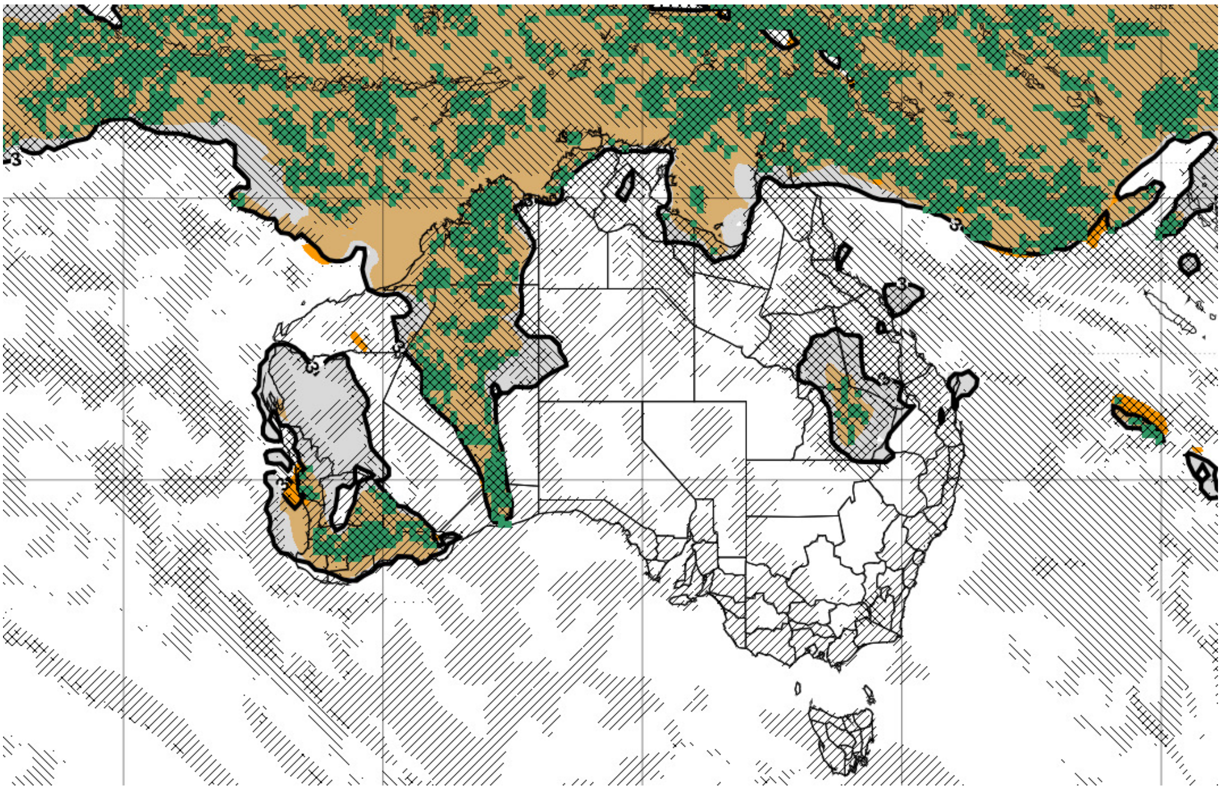

Fig. 5: NTFGS surface-based thunderstorm diagnostic output (green pixels) at 06 UTC, 31 December 2011. Shown are all five surface-based thunderstorm ingredients, plotted only where they exceed their respective thresholds as indicated in Table 1. Areas of lifted index < -1°C are shaded in orange, the semi-transparent grey shading marks parcel equilibrium levels colder than -20°C, the thick black contour outlines areas of cold cloud depth exceeding 3 km, northwest-to-southeast diagonal stripes mark areas of |CIN| < 25 J/kg, and southwest-to-northeast diagonal stripes show regions where omega exceeds -10 hPa per hour of upmotion.

Figure 5 shows the individual ingredients for surface-based thunderstorms on 31 December 2011 (06 UTC). For this example the cloud electrification criteria impose no constraint on the surface-based thunderstorm area (green pixels) beyond the conditional instability condition (LI < -1°C). Within extensive areas of |CIN| < 25 J kg^{-1}, the model upward vertical velocity for this case provided a second substantial constraint on the extent of the thunderstorm area. Therefore, the areal extent of surface-based thunderstorms for this case is dominated by a combination of sufficient conditional instability in regions of decent upward vertical velocity in the model. The quality of this diagnosis will be addressed in the next section.

3. PRELIMINARY VERIFICATION RESULTS

Thirty-one NTFGS surface-based and elevated thunderstorm (hereafter: thunderstorm) diagnoses based on 12-hour forecasts of the ACCESS-R model were compared against lightning data. All forecasts were valid at 06 UTC (local afternoon) during 1-31 December 2011. ACCESS-R is essentially identical to ACCESS-A described above, but is defined on a larger domain (65°S to 17.125°N; 65°E to 184.625°E) and a coarser grid (0.375° or ~37.5 km).

The lightning data used are provided by the Global Position and Tracking Systems Pty. Ltd. or GPATS. GPATS is a time-of-arrival system that detects signals in the very low and low frequency band (1.5 kHz - 400 kHz). It therefore has its highest detection efficiency for cloud-to-ground lightning, rather than cloud-to-cloud which emits more strongly in the VHF range. The sensor network has become adequately dense over Australia in recent years, but still leaves some sparsely covered areas in the western half of the continent (Fig. 6). Overall, the interpretation of any verification results involving GPATS lightning needs to qualitatively take into account a low bias for the reasons stated above.

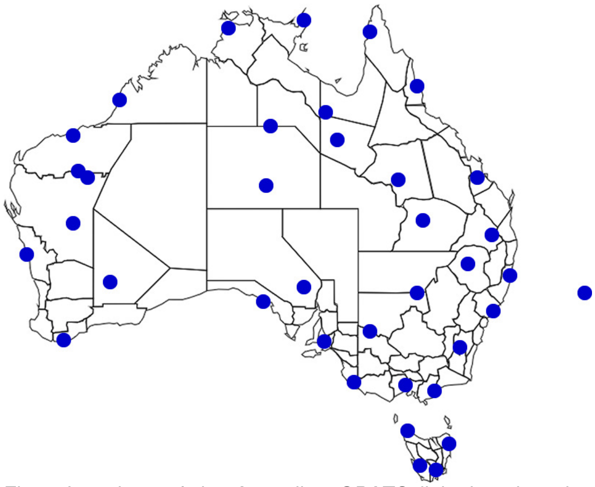

Fig.6: Locations of the Australian GPATS lightning detection sensors as of October 2012.

A typical comparison between the NTFGS 06 UTC thunderstorm diagnostic and the corresponding lightning flash density is given in Fig. 7. The lightning data are aggregated over a 3-hour period centered on 06 UTC 31 December 2011, which matches a gapless output time-centered time interval for the 3-hourly ACCESS-R or NTFGS output.

Subjective evaluation shows for all thirty-one December cases an overprediction of thunderstorms, in particular over the tropical ocean areas. Over the Australian continent most thunderstorms are detected by the NTFGS, with the majority of the missed events showing low ($N < 50$) flash densities. While not as dramatic as over the tropical oceans, the NTFGS also overpredicts maritime thunderstorms in the temperate zones, but allowances need to be made for the decrease in GPATS detection efficiency with increasing distance from the coastline.

Noting the differences between the large high bias over tropical maritime regions and a reduced high bias over land, a selection of neighborhood verification strategies summarized in Ebert (2008; 2009) was employed separately in those two areas

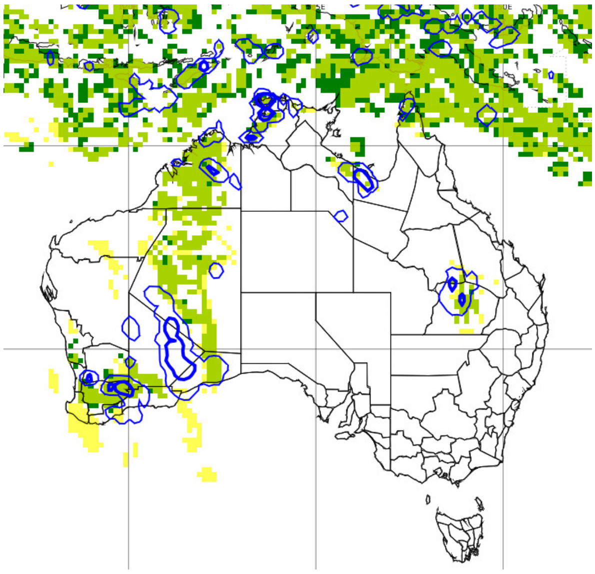

Fig.7: Comparison of the 06 UTC 31 December 2011 surface and elevated thunderstorm diagnostics with the lightning flash density between 0430 and 0730 UTC. Dark green pixels represent surface-based thunderstorms, yellow pixels show elevated thunderstorms, and light green pixels indicate an overlap of both. The flash density is a count of detected strikes N over 3 hours inside an ACCESS-R grid box (approximately 37.5 km x 37.5 km) with the thin blue contour showing N=1, the bold blue contour denoting N=50 and a second thin contour (apparent near Darwin) showing N=200.

(see Fig. 8). A neighbourhood approach rather than the more traditional point verification approach was chosen to avoid penalizing the NTFGS output for small spatial placement errors. Rather, any forecast thunderstorm event that is close to observed storms in either space or time (or both) is rewarded for “almost” producing a totally accurate event prediction. Fig. 9 shows the verification results from four different neighbourhood verification methods applied to 31 cases over the tropical maritime domain.

The multi-event contingency table method (Atger 2001) in Fig. 9a (top left panel) centers progressively larger neighborhoods on the location of an observed event and tests for the presence of an event forecast. The near-zero Hanssen and Kuipers score (HK; the difference between the probability of detection, POD, and the false alarm rate, F) indicates that the NTFGS shows practically no skill over the tropical oceans north of Australia. While the algorithm pretty much captures every storm event (POD ~ 1), the vast number of false alarms (F ~ 1) negates those previous benefits.

Low values of the Fractions Skill Score (FSS) indicate that the fractional coverage of the NTFGS far outweighs that of the observed thunderstorms which underscores the over-forecasting tendency across all tested neighborhood scales (Fig.9b, top right panel). The pragmatic method (Fig.9c, lower left panel), which compares the forecast event probability inside a neighborhood with the observations in the center (Theis et al. 2005), shows a decidedly negative Brier Skill Score (BSS) indicative of a very high forecast fractional thunderstorm coverage. The NTFGS results are also far from practically perfect (Fig. 9d, panel 4), which means that they deviate substantially from the thunderstorm area that would have been drawn by a forecaster with hindsight access to all the thunderstorm observations.

Over the Australian continent all scores show substantial improvements (Fig.10). Given the fractional coverage of the domain by observed thunderstorms, f_{TS}, is relatively small, the NTFGS assessment for continental December 2011 thunderstorms has useful skill (FSS > $0.5+f_{TS}/2$ ~ 0.5) on most neighborhood scales (Roberts and Lean 2008). The NTFGS diagnoses are most useful on the 1-2 degree scale where HK is maximized while FSS > 0.5.

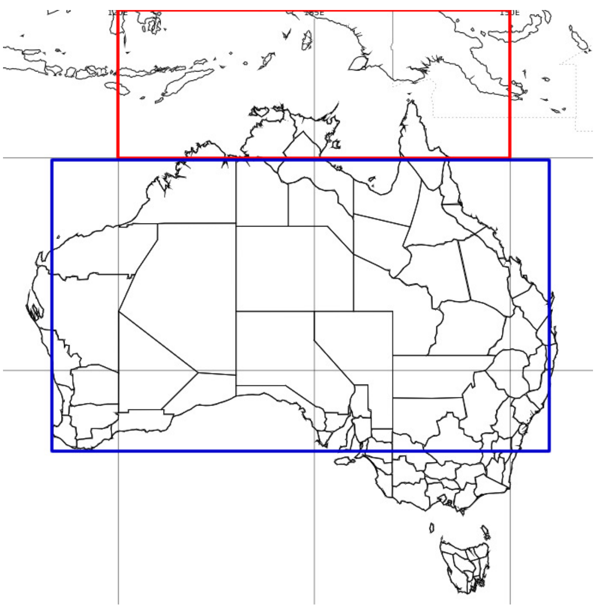

Fig.8: Neighborhood verification regions employed in the preliminary verification approach. The tropical maritime region (15°S to 4°S, 120°E to 150°E) is shown as a red box, the continental region (35°S to 15°S, 115°E to 153°E) is shown as a blue box.

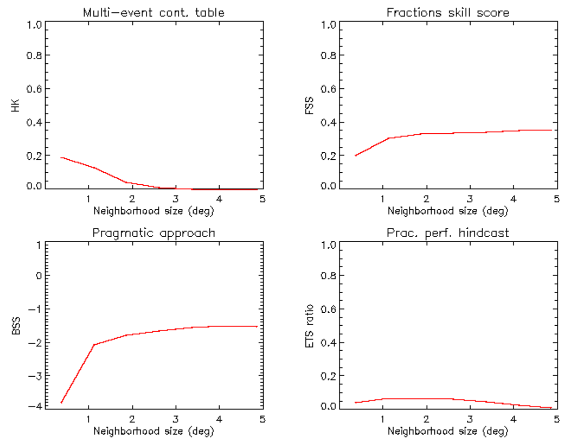

Fig.9: Neighborhood verification scores for the tropical maritime domain in Fig 8. Shown, as a function of the (~square-shaped) neighborhood size in degrees are the Hanssen and Kuipers discriminant (HK; top left) associated with the multi-event contingency table method from Atger (2001), the Fractions Skill Score (FSS; top right) based on Roberts and Lean (2008), the Brier Skill Score (BSS; bottom left) associated with the pragmatic method (Theis et al. 2005; Ebert 2008), and the Equitable Threat Score ratio (ETS ratio; bottom right). The ETS ratio compares the ETS score for the NTFGS thunderstorm forecast to the ETS of an equivalent "perfect" hindcast based on the practically perfect hindcast approach described in Brooks et al. (1998).

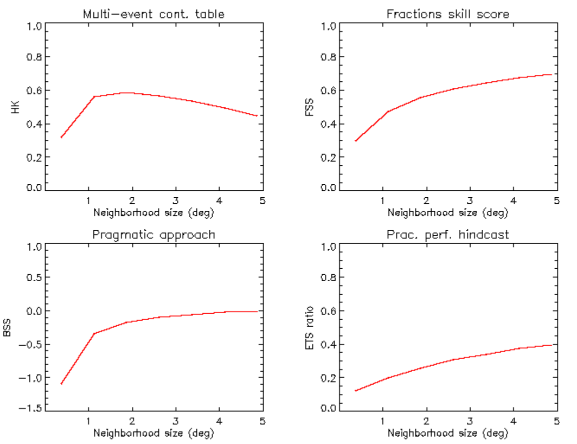

Fig.10: As in Fig.9, but for the continental domain shown in Fig. 8.

4. PLANS FOR NTFGS DEVELOPMENT

An extensive consultation process was conducted during the first half of 2012 which received input from practically all Australian forecast offices and the Training, Services Policy and Research branches of the Bureau. The received suggestions cluster around the following points:

- The dependence of the NTFGS on a single deterministic model run needs to be removed in favor of either a lag ensemble using ACCESS, or a true ACCESS ensemble (AGREPS-R).
- The allocation of qualitative, categorical probabilities based on the exceedance of hard single thresholds needs to give way ‘softer’ thresholds, either through employing multiple hard thresholds or “fuzzification” of individual thresholds.
- Verification of the two types of thunderstorm assessments (surface-based and elevated) is already under way using GPATS lightning data; this should allow for calibration and the transition of categorical into numerical probabilities.
- Verification of the convective hazard diagnostics in Australia suffers from significant under-reporting away from the major population centers; a blended approach of using more rigorously verified ingredients from the U.S. combined with an ongoing verification process restricted to the major Australian population centers is desirable.
- The display of the NTFGS diagnostics and ingredients needs to transition from its current less flexible web display to the operational integrated data viewer (Visual Weather) at the Bureau. Figs 5 and 7 show samples of a prototype display in Visual Weather.
- The quality of model-assessed convective initiation (CI) likelihood needs addressing; at present threshold exceedance for mixed parcel CIN and low-level omega are used as the CI proxy. Should convective hazard potential (excluding CI) be shown separately to overall convective hazard probability (potential including likelihood of CI)?
- A range of individual ingredients (superfluous or missing) and ingredient threshold changes have been identified. Examples are the addition of a lifted condensation level < X criterion (X is a low threshold height above ground level) to the tornado diagnostic, or the consideration of 0-6 km shear instead of the current ~0-4 km layer for supercells.
- The oceanic high bias needs to be addressed.

5. ACKNOWLEDGEMENTS

Special thanks is due to Paul Hettrick for his advice on the GPATS network attributes, and to Beth Ebert for her support and input on the neighborhood verification aspects of this paper. I would also like thank my three internal reviewers, Alain Protat, Jeff Kepert and Tony Bannister.

REFERENCES

Atger F., 2001: Verification of intense precipitation forecasts from single models and ensemble prediction systems. *Nonlinear Processes in Geophysics,* **8,** 401–417.

Brooks H. E., Kay M., Hart J. A., 1998: Objective limits on forecasting skill of rare events. Extended Abstracts, 19th Conf. on Severe Local Storms, Minneapolis, MN, Amer. Meteor. Soc., 552-555.

Ebert, E. E., 2008: Fuzzy verification of high resolution gridded forecasts: A review and proposed framework. *Meteor. Appl.,* **15,** 51-64.

Ebert, E. E., 2009: Neighborhood verification: A strategy for rewarding close forecasts. *Wea. And Forecasting,* **24,** 1498-1510.

Edwards, J. M., and Slingo, A., 1996: Studies with a flexible new radiation code. I: Choosing a configuration for a large-scale model. *Q. J. R. Meteorol. Soc.*, **122,** 689–719

Essery, R. L. H., M. J. Best, R. A. Betts, P. M. Cox, and C. M. Taylor, 2003: Explicit Representation of Subgrid Heterogeneity in a GCM Land Surface Scheme. *J. Hydrol.,* **4,** 530-543.

Grant, A. L. M., and Brown, A. R., 1999: A similarity hypothesis for shallow-cumulus transports, *Quart. J. Roy. Meteor. Soc.,* ***125,*** 1913–1936.

Gregory, D., and P. R. Rowntree, 1990: A mass flux convection scheme with representation of cloud ensemble characteristics and stability-dependent closure, *Mon. Weather Rev.,* ***118,*** 1483–1506.

Hanstrum, B. N., 2004: A National NWP based thunderstorm and severe thunderstorm forecasting guidance system. *Preprints, International Conference on Storms, AMOS 11th National Conference 5-9 July 2004*, Brisbane, Australia, 31-36.

Hanstrum, B. N., Mills, G. A., Watson, A., Monteverdi, J.P., Doswell, C.A., 2002: The cool-season tornadoes of California and Southern Australia, *Wea. and Forcasting* **17,** 705-722.

Johns, R. H., and C. A. Doswell III, 1992: Severe local storms forecasting. *Wea. Forecasting*, **7**, 588–612.

Lock A. P., 1998: The Parameterization of entrainment in cloudy boundary layers, *Quart. J. Roy. Meteor. Soc.,* ***124,*** 2729-2753.

Lock A. P., Brown, A. R., Bush, M. R., Martin, G. M. and Smith, R. N. B., 2000: A new boundary layer mixing scheme. Part I: Scheme description and single-column Model, *Mon. Wea. Rev.,* ***128,*** 3187-3199.

Lock, A. P., 2001: The numerical representation of entrainment in parameterizations of boundary layer turbulent mixing. *Mon. Wea. Rev.*, **129**, 1148–1163.

Puri, K., G. Dietachmayer, P. Steinle, M. Dix, L. Rikus, L. Logan, M. Naughton, C. Tingwell, Y Xiao, V. Barras, I. Bermous, R. Bowen, L. Deschamps, C. Franklin, J. Fraser, T. Glowacki, B. Harris, J. Lee, T. Le, G. Roff, A. Sulaiman, H.

Sims, X. Sun, Z. Sun, H. Zhu, M. Chattopadhyay, C. Engel, 2012: Implementation of the initial ACCESS Numerical Weather Prediction system, *Aust. Met. Oc. Journal,* submitted.

Roberts N. M., Lean H. W., 2008: Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Wea. Rev.,* **136,** 78–97.

Smith, R. N. B. 1990: A scheme for predicting layer clouds and their water contents in a general circulation model, *Quart. J. Roy. Meteor. Soc.,***116,** 435-460.

Theis S. E., Hense A., and Damrath U., 2005: Probabilistic precipitation forecasts from a deterministic model: A pragmatic approach. *Meteor. App.,* **12,** 257–268.

Treloar, A. B., and B. N. Hanstrum, 2002: A Study of Australian Warm Season Thunderstorms. *Extended Abstracts, Weather Services Scientific Conference 3-7 June 2002*, Melbourne, Australia, 99-100.

Wilson, D. R., and Ballard, S. P., 1999: A microphysically based precipitation scheme for the UK Meteorological Office Unified Model, *Quart. J. Roy. Meteor. Soc.,***125,** 1607-1636.