ASSESSMENT OF TROPICAL RAINFALL POTENTIAL (TRaP) FORECASTS DURING THE 2003-04 AUSTRALIAN TROPICAL CYCLONE SEASON

Elizabeth Ebert

Bureau of Meteorology Research Centre
GPO Box 1289K, Melbourne, Vic., Australia 3001

ABSTRACT

Tropical Rainfall Potential (TRaP) forecasts provide estimates of 24 h rainfall accumulation in landfalling tropical cyclones (TCs) based on the advection of a field of satellite-estimated precipitation. Validation of TRaP forecasts for five Australian tropical cyclones during the 2003-04 season showed significant skill in predicting heavy rainfall. The predictions of maximum rain at landfall compared well with gauge observations in most cases. The AMSU-based TRaPs performed noticeably better than the SSM/I-based TRaPs in predicting spatial rain coverage and amount, giving higher correlations with the observations, more accurate estimates of rain area and conditional rain rate, and lower root mean squared errors. The TRaPs performed neither better nor worse than mesoscale numerical weather prediction models. A decomposition of the TRaP error for regions of heavy rain suggests that only a small portion was related to errors in the track forecasts. Pattern errors, which relate to the shape, size, and fine scale structure of the forecast entity, accounted for about half of the total error, while rain volume error was about one third of the total error. These relate to errors in the satellite rain rate retrieval as well as the assumption of a time-invariant rain pattern. An ensemble of TRaP forecasts could account for some of these uncertainties, leading to more useful objective guidance.

1. INTRODUCTION

To help provide objective guidance for flood forecasts associated with heavy rain in landfalling tropical cyclones, the NESDIS Satellite Services Division issues short-term space-based rain forecasts called Tropical Rainfall Potential (TRaP). TRaP forecasts provide estimates of rainfall accumulation based on the simple advection of a field of satellite-estimated precipitation, essentially a satellite precipitation "nowcast". It is computed as

$$TRaP(x,y) = \int_t R(x,y,t)dt \qquad (1)$$

where R is the satellite-derived rain rate, x and y denote the spatial location and t is the time. The integration period is normally 24h. The rain rate is assumed constant in a Lagrangian sense, while the storm motion is obtained from official track forecasts from operational centers. Areal TRaPs for cyclones that are 24-36 hours or less from landfall are now available for most tropical storms and cyclones around the globe (Kidder et al., 2004; see also http://www.ssd.noaa.gov/PS/TROP/trap-img.html). A subset of vetted (checked and approved) TRaP products are available in near real time on the web as (gif) images at 4 km horizontal resolution. The digital forecasts can also be downloaded from SSD as McIDAS areas or text files.

Although the primary goal is to estimate the maximum rainfall amount and location, the spatial distribution of rainfall predicted by the areal TRaPs are potentially very useful. To make best use of this product it is important to know how well it predicts the maximum rain amount, the location of maximum rainfall, the spatial extent of the rainfall, and the total volume of rainfall. Operational forecasters also need to know whether the TRaP forecasts are more accurate than those available from numerical models. Algorithm developers and more sophisticated users of TRaP may wish to understand how the rainfall errors are related to errors in track forecasts, errors in satellite-estimated rain rates, and the assumption of (Lagrangian) steady-state rainfall during a 24h period.

TRaPs have been validated for hurricanes in the western hemisphere during the 2001 and 2002 seasons (Ferraro et al., 2002, 2004). Since the primary objective of the TRaP is to provide an early warning of potential maximum rainfall for locations near the coast, the emphasis of Ferraro et al.'s (2002) validation was on the location and magnitude of the predicted rain maximum. Examining results for seven storms they found that the TRaP maximum rainfall was generally within ±80% of the maximum value observed at rain gauges. Inaccuracies in hurricane track forecasts led to errors in TRaP rainfall. However, there was no systematic relationship between errors in storm motion and under- or over-estimates of rain, suggesting that errors in the satellite rain rate retrievals, combined with the simplifying assumption of a steady state spatial rain distribution in the travelling storm, also contributed strongly to the total error.

A study of TRaPs for the 2002 Atlantic hurricane season focused on the validation of spatial rain forecasts (Ferraro et al., 2004). Using 24 h accumulations calculated from the National Centers for Environmental Prediction (NCEP) Stage IV hourly gauge-radar rainfall analysis as validation data, they found that the most accurate TRaP forecasts were those based on the Tropical Rain Measuring Mission (TRMM) rainfall estimates, but all TRaPs tended to underestimate the maximum rainfall. The TRaP forecasts outperformed Eta model forecasts according to most statistical measures.

In this study TRaP forecasts for five tropical cyclones in the Australian region during the 2003-04 season were verified against daily rain gauge observations and an operational gridded daily rainfall analysis. The complete set of automated areal TRaPs were available for validation. In practice, most users have access to only the vetted TRaPs, which would be expected to be of the highest quality because they are evaluated and checked by an operational satellite meteorologist. An object-oriented contiguous rain area (CRA) verification approach was used to verify storm rainfall properties and to separate the error contributions due to incorrect cyclone track, rain volume, and spatial distribution. Mesoscale model forecasts were verified in a similar manner to assess the "competition". The most important results from this study are highlighted here; details can be found in the paper by Ebert et al. (2004).

2. VALIDATION METHODOLOGY

Australia has a national network of over 5000 rain gauges that measure 24 h rain accumulation at 9 am local time each day[1]. These rain gauge observations were used to validate the maximum rainfall predicted by the TRaPs by comparing the observed maximum rainfall at the first 9 am observation time after landfall to all TRaP forecasts valid within ±12 h of that time. To investigate

[1] 9 am local time corresponds to 2300 UTC in Queensland, 2330 UTC in the Northern Territory, and 0100 UTC in Western Australia.

the impact of timing differences, a second comparison was made using only TRaPs valid within ±3 h of the rainfall observations.

This maximum rainfall validation is less than ideal, for a few reasons. Gaps in the rain gauge network and gauge "undercatch" at high wind speeds mean that the true maximum rain accumulation is unlikely to be observed. As noted, timing differences between the TRaP and the observations will lead to apparent errors in the predicted rainfall, even for a perfect TRaP. Since TRaP forecasts are derived from coarser resolution satellite rainfall estimates, they do not represent the spatial scale of the point observations from gauges and thus the comparison suffers from errors of representativeness. Nevertheless, to the extent that users are tempted to take TRaP estimates of maximum rainfall at face value, the comparison is appropriate.

The areal distribution of rainfall in the TRaP forecasts was validated against the Australian operational daily rainfall analysis. The gauge data are analyzed onto a 0.25° latitude/ longitude grid using a 3-pass variable length scale Barnes objective analysis scheme (Weymouth et al., 1999). In the absence of gauge-calibrated radar observations the gauge analysis provides the best estimate of spatial rainfall distribution. It is limited in its accuracy by gaps in the network in the tropics (spacing of roughly 1 gauge per 60 km except in the vicinity of Darwin) and therefore cannot represent the strongest spatial gradients in the cyclone rainfall.

The TRaP rainfall fields were remapped onto the same grid as the analysis for the spatial validation. To focus only on the rainfall of interest, the validation domain was limited to a (moving) 10° latitude/longitude box centered on the observed cyclone position. The observed tropical cyclone positions were provided by the responsible Tropical Cyclone Warning offices in Brisbane (Queensland), Darwin (Northern Territory), or Perth (Western Australia), based on a careful analysis of enhanced infrared imagery and synoptic observations.

The spatial validation was performed for the 85 TRaP forecasts with valid times falling within ±3 h of the observations. Of the 40 vetted forecasts that were released to the public, nine met this timing criteria. TRaPs constructed from satellite passes with incomplete coverage were not included in the validation. All spatial validation results refer to the portion of the TRaP that was over land.

As pointed out by Kidder et al. (2004) there are three main sources of uncertainty in TRaP rainfall forecasts:

- the satellite-estimated rain rates,
- the forecast storm track,
- the invariant spatial structure.

Different validation strategies may be more appropriate for assessing each source of error.

Considering first the errors in instantaneous rainfall rates diagnosed by the passive microwave instruments, previous validation studies have found typical error magnitudes of 100% or more in the tropics (Ebert et al., 1996; Smith et al., 1998). Different sensors have different measurement error characteristics. For example, the SSM/I and TRMM TMI algorithms used at NOAA are known to systematically underestimate low rain rates and overestimate heavy rain rates over land and ocean, while the AMSU algorithm tends to overestimate the area of light rain over land (R. Ferraro and S. Kusselson, unpublished results). Errors in the satellite rainfall "snapshot", compounded by the assumption of steady state rainfall, will be reflected as errors in the predicted 24 h accumulation. These will be seen mainly as errors in rain amount and extent, which can be quantified using the mean absolute error (MAE) and root mean square error (RMSE), and the ratio of forecast to observed rain amount (multiplicative bias) or area (frequency bias).

Another source of uncertainty is the predicted cyclone motion, which originates from human forecasters in the various tropical cyclone prediction centers. If the storm is forecast to move too quickly or too slowly, the rain extent and accumulation will be incorrect. If the predicted direction of motion is incorrect the TRaP will put the forecast rain in the wrong location. The correlation coefficient is a good indicator of whether the rain pattern is correct. To evaluate predicted rain occurrence the threat score (TS) measures the ratio of the number of hits (rain both predicted and observed) to the number of points with rain either predicted or observed. To help interpret the threat score it is useful to compute the probability of detection (POD), which measures how often the observed rain was correctly predicted, and the false alarm ratio (FAR), which indicates the fraction of rain predictions that were false alarms. It is common to set a rain threshold for computing the frequency bias, POD, FAR, and TS; we use values of 1 mm d^{-1} to measure success for rain / no rain prediction, and 20 mm d^{-1} to measure success for the heavier, more important, rain.

Errors due to incorrect storm motion versus those due to incorrect rainfall amount and spatial structure were separated using the object-oriented "contiguous rain area" (CRA) method of Ebert and McBride (2000). Forecast and observed rain entities were defined by a rain threshold of 20 mm d^{-1} to isolate the more important rain. The location error of a forecast entity relative to the observed entity can be estimated using automated pattern matching, or it can be specified externally. For our application we used the observed cyclone position determined by the Australian forecast office responsible for monitoring the TC to specify the position error of the forecast. The properties of the forecast entity, namely, the rain area, volume, conditional rain rate, and maximum rain rate, as well as the spatial pattern of rainfall, were verified after horizontally translating the forecast rain to the corrected location. This should be a reasonable approximation of the TRaP that would be obtained with a perfect track forecast. The CRA methodology allows the total error to be decomposed into contributions from location error, volume error, and pattern error.

3. RESULTS FOR AUSTRALIAN TROPICAL CYCLONES

TRaP forecasts were available for five tropical cyclones in the Australian region during the 2003-04 season. The observed cyclone tracks are shown in Figure 1. Three of the TCs (Debbie, Fritz, and Evan) affected the northernmost tropical latitudes of Australia, while two (Monty and Fay) produced heavy rainfall in the subtropical northwestern part of the continent. TRaP forecasts were generated using rainfall estimates from three different sensors, AMSU, SSM/I and TRMM, and track forecasts from Australian operational forecast centers and the U.S. Joint Typhoon Warning Center (JTWC). This paper presents only the aggregated results for all five cyclones. Results for individual

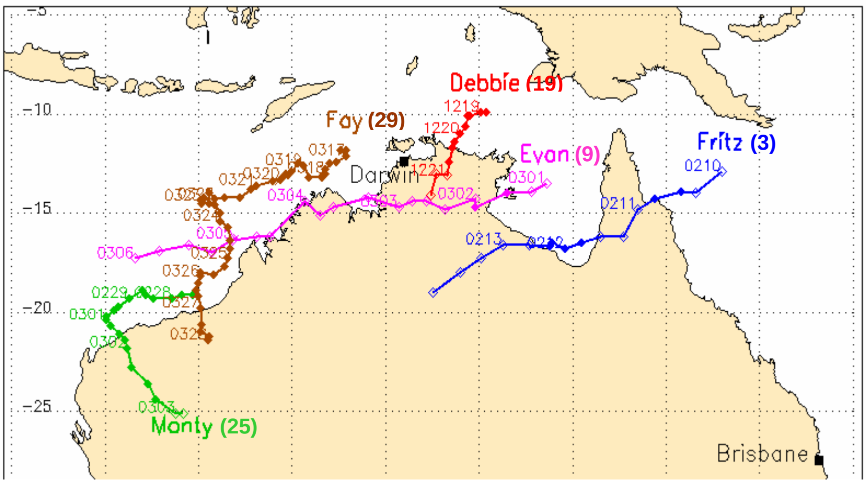

Figure 1. Observed tracks of five tropical cyclones during the 2003-04 season. The labelled dates indicate the 0000 UTC position of the storm. The open symbols denote times when the storm did not have tropical cyclone status. The number in parentheses gives the number of TRaPs validated for each storm for which spatial validations were **done.**

cyclones are described by Ebert et al. (2004).

The TRaPs usually predicted reasonably good values of maximum rainfall when compared to gauge observations (Fig. 2). Looking first at the results for TRaP forecasts valid within ±12 h of the observation time, the AMSU-based TRaP estimates appeared to be the most reliable and conservative, with a mean absolute error of 34% of the observed value. The SSM/I-based TRaPs were too high by a factor of three for TCs Debbie and Fay, but gave quite accurate estimates (MAE of 11% of the observed maximum) for TCs Fritz and Monty. TRaPs computed from TRMM data had a MAE of 56%, and seemed to suffer some of the same overestimation problems found with the SSM/I-based TRaPs. Comparisons of validation results for TRaP forecasts valid within ±3 h versus ±12 h of the observation time show no consistent improvement, but with such a small sample size it is impossible to draw any strong conclusions.

The SSM/I-based TRaPs produced unrealistically high rain maxima in two of the storms. If this was related mainly to the spatial resolution of the satellite observations then the SSM/I-based TRaPs would be expected to have lower rain maxima than the TRMM-based TRaPs, which was not the case. The errors in track forecasts for SSM/I-based TRaPs were similar to those for the other TRaPs. We therefore believe that this overestimation problem is related to errors in the satellite rain rate retrievals. In contrast, Ferraro et al. (2004) found that for U.S. hurricanes the TRaPs from all three sensors underestimated the maximum rain at landfall when compared to the Stage IV gauge-radar analysis (TRaP maxima of 65-220 mm d^{-1} compared to radar maxima of 225-475 mm d^{-1}). This difference may be related to the higher spatial density of surface observations in the US validation and the inherent differences between gauge and radar analyses, but also to the difference in atmospheric moisture between the moist southeastern United States and the relatively drier Australian tropics. A drier atmosphere would result in more rain evaporating before reaching the surface, implying that the satellite rain rate algorithms may need to be recalibrated for Australian conditions.

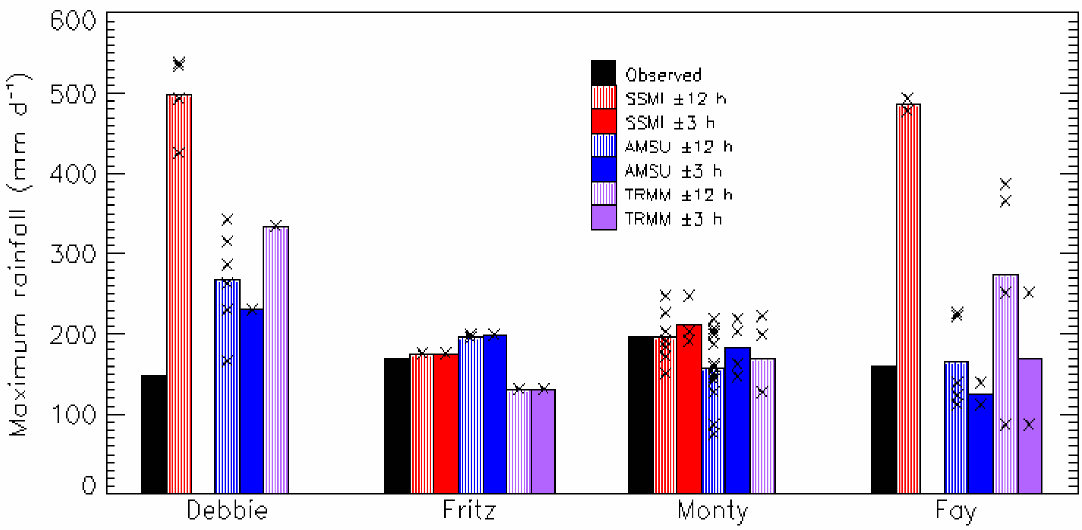

Figure 2. Maximum rainfall corresponding to 9 am local time on the first day following landfall. No TRaPs were available within 12 h of landfall for TC Evan. The black bars represent the gauge observations, while the shaded and textured bars give the mean TRaP values. The left bar (lighter shade) in each pair corresponds to TRaPs valid within ±12 h of the observation, while the right bar (darker shade) corresponds to TRaPs valid within ±3 h of the observation. The individual estimates are shown by the x's.

Box plots of the spatial validation statistics for TRaPs are given in Fig. 3. In this plot the box shows the central 50% of the distribution, with the median shown by the horizontal line, and the vertical line shows the full range. For the Australian 2003-04 tropical cyclone season, the AMSU-based TRaPs performed slightly better than the TRMM-based TRaPs and noticeably better than the SSM/I-based TRaPs, giving higher correlations with the observations, more accurate estimates of rain occurrence, and lower values of RMSE. All three products systematically underestimated both the rain area and volume.

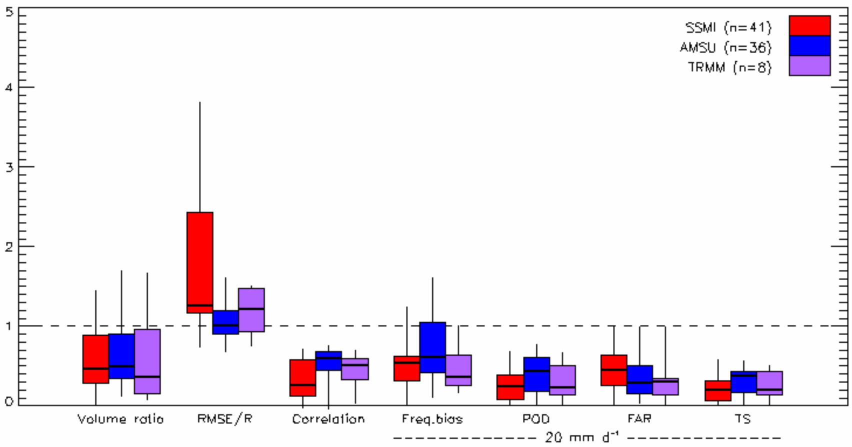

Figure 3. Validation results for TRaPs based on SSM/I, AMSU, and TRMM precipitation estimates.

The CRA validation was applied to the vetted TRaPs and focused on the regions of rain exceeding 20 mm d^{-1} in the forecast and observations (Fig 4). Compared to the results for the full 10° spatial domain the conditional rain and volume ratios were closer to 1 and the normalized RMSE was lower, implying relatively better performance of the TRaPs for the heavier rain. The location (track) errors ranged from 0 to 113 km, with a mean value of 55 km. The error decomposition suggests that on average less than 20% of the forecast error was related to track errors. Pattern errors, which relate to the shape, size, and fine scale structure of the forecast entity, accounted for about half of the total error. The rain volume error component was quite variable but averaged about one third of the total error. It appears that the errors associated with the rain rate retrieval and the assumption of steady state rain distribution in the 24h period outweigh the errors associated with incorrect track forecasts.

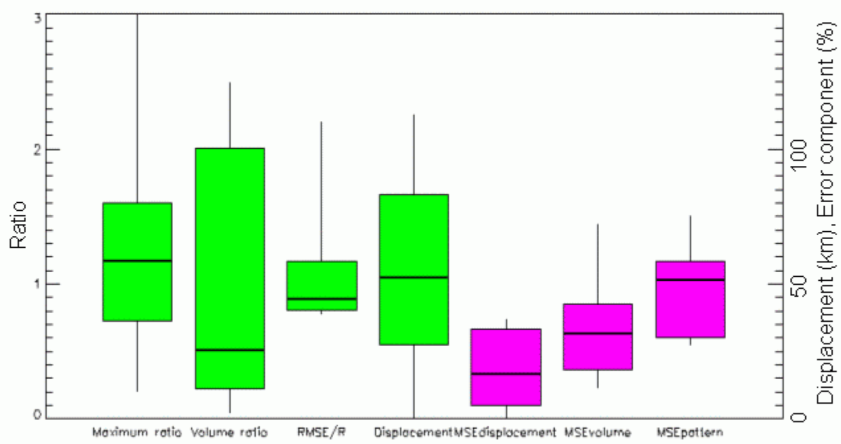

Figure 4. CRA validation results for vetted TRaPs.

Correcting the location of the TRaPs had the effect of more than doubling the average correlation coefficient for rain exceeding 20 mm d^{-1}, but was equally likely to decrease or increase the RMSE. This is in contrast with Ferraro et al.'s (2004) findings that recomputing the TRaP using a perfect (observed) track often improved, but never degraded, the performance.

To test whether the TRaPs gave better rainfall forecasts than the mesoscale NWP guidance, the validation results were compared for the eight days (spread over four cyclones) during which AMSU- and SSM/I-based TRaPs and mesoLAPS and TC-LAPS model (Davidson and Weber, 2000) 24 h rain forecasts were all available. Figure 5 shows examples of model forecasts for one day during TC Monty. The model forecasts have quite a different, arguably more realistic, appearance to the TRaPs, with a suggestion of rotation rather than streakiness. However, the excessively heavy rain in the TC-LAPS forecast, and to a lesser degree in the mesoLAPS forecast, appeared to be a common occurrence.

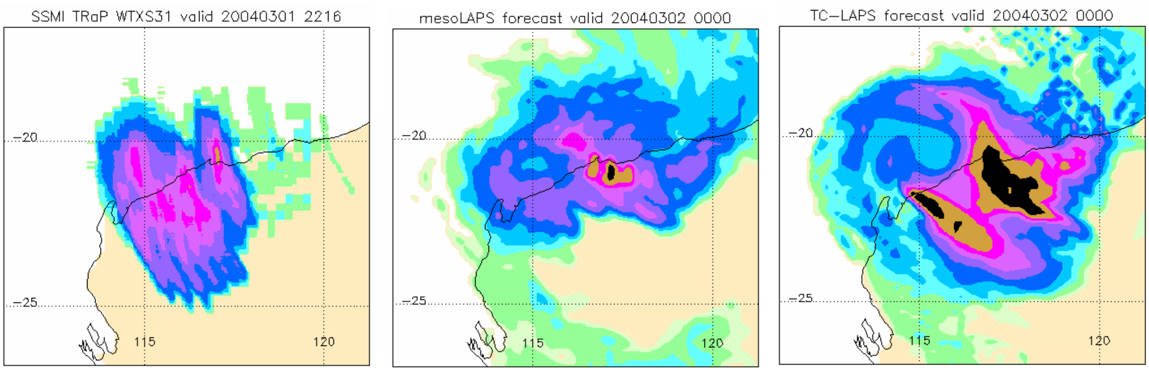

Figure 5. TRaP, mesoLAPS model, and TC-LAPS model valid at 00 UTC on 2 March 2004 **(TC**

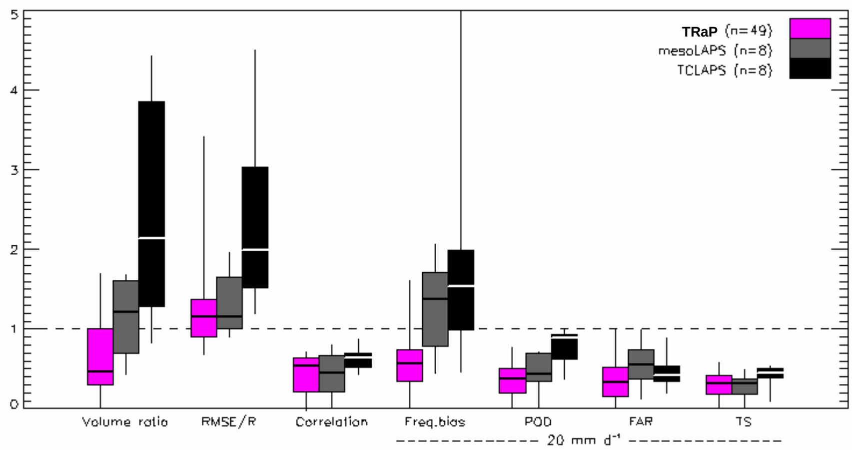

Figure 6. Comparison of TRaP to NWP models.

Figure 6 shows the distributions of selected verification statistics for the three forecast products. Looking first at rain volume, the mesoLAPS model gave the best estimates while TC-LAPS overestimated it quite severely, leading to large errors. The TRaPs had the lowest RMS errors, while the correlations were highest for the TC-LAPS model. The mesoLAPS model gave the best estimates of heavy rain area. The TC-LAPS model had the highest probabilities of detection, which contributed to it achieving the highest median threat score of the three products. Based on this comparison, there is no clear-cut "winner" – the choice of forecast would vary, depending on which quantities were deemed most vital to predict correctly.

4. DISCUSSION AND SUGGESTIONS FOR IMPROVEMENT

Both the Australian and US validation studies show that these satellite-based rainfall extrapolation products clearly have skill in predicting heavy rainfall. The AMSU-based TRaPs gave more reliable estimates of heavy rain magnitude at landfall than the SSM/I-based TRaPs. The TRaPs and models had comparable skill in predicting the location of the heavy rain, as measured by the threat score and correlation coefficient, but the mesoLAPS model gave better estimates of rain area and volume.

These results were based on a small number of cases. Validation efforts should continue so that more robust conclusions can be drawn. Australia does not currently have an hourly gauge-radar analysis similar to the Stage IV product in the US, so it must continue to rely on 24 h rain gauge observations for the bulk of the observational data. Some of the difference in the results between this study and the Ferraro et al. (2004) study are undoubtedly related to the differences in the reference data between Australia and the U.S. Differences in environmental influences on TC rainfall are also likely to be an important factor.

The areal TRaP is still being further developed and improved (Kidder et al., 2004). The CRA error decomposition suggests that most of the error is due to errors in rain volume and pattern, as opposed to incorrect track forecasts. As long as the tracks are reasonably accurate, and given that the TRaP methodology does not "grow" or "decay" rain, systematic errors in rain volume will be primarily associated with errors in the satellite rain retrieval. Further validation of satellite rain rates against observations from coastal radar, the TRMM precipitation radar, or gauge observations from atolls and island stations might help clarify some of the rain retrieval errors. The pattern errors are related to the assumption of steady state (in a Lagrangian sense) rain including the lack of rotation in the advected rain field. The streaky appearance of the TRaPs that results from advection of local rain maxima and minima is clearly unrealistic and may deter some forecasters from having confidence in the product.

Several physically based improvements are possible. Storm rotation could be incorporated into the storm motion. Atmospheric moisture retrievals from passive microwave measurements could be used to increase or decrease the TRaP rainfall based on moist or dry advection. Similarly, a shear factor, perhaps derived from NWP, could be used to increase or decrease TRaP rainfall. Including a factor for orographic enhancement of rainfall might produce more realistic land-based rain estimates. Some simple statistical methods could be applied to reduce the streakiness of the TRaPs to give a more realistic looking, and more accurate, TRaP forecast. The simplest approach is to apply a spatial smoother to the TRaP output, but this has the undesirable effect of reducing the maximum rain. To overcome this problem probability matching can be used to transform the smoothed rain rates back to the original rain frequency distribution.

Rain probabilities can also be estimated by sampling the rain distribution in the local spatial and/or temporal neighbourhood of each pixel. For example, all rain estimates from pixels within 100 km radius of the pixel of interest, or within 12 hours, could be considered equally likely, so that the probability of precipitation exceeding a given threshold would be simply the fraction of pixels with rain exceeding that threshold. Such an approach was successfully demonstrated by Theis et al. (2004) using mesoscale NWP model output; the principle would be equally applicable to high resolution TRaP forecasts. In the example shown in Figure 7 all TRaPs valid within ±12 h of 0000 UTC 2 March 2004 were used to generate rain amount and probability forecasts.

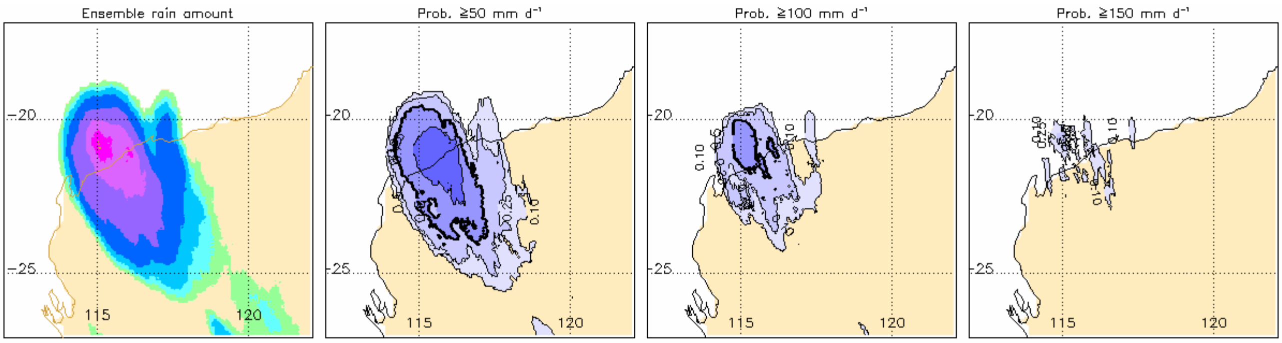

Fig 7. Ensemble TRaP based on 27 AMSU, SSM/I, and TRMM TRaPs within ±12 h of 0000 UTC, 2 March 2004 (TC Monty).

A better strategy would be to generate an ensemble of TRaPs by adding realistic perturbations to the forecast speed and direction of the cyclone. The parameters of the microwave rain rate retrieval could also be varied to provide several initial fields. The ensemble strategy acknowledges that there are many uncertainties in the forecast, and explicitly takes them into account. The ensemble mean TRaP, corrected using probability matching as above, would almost certainly be a more skilful forecast than a single realization (i.e., the current TRaP product). The great advantage of the ensemble approach is the ability to easily produce probability forecasts for critical rain thresholds. This would add enormous value to the TRaP product in decision-critical situations like tropical cyclones approaching landfall.

5. REFERENCES

Davidson, N.E. and H.C. Weber, 2000: The BMRC high-resolution tropical cyclone prediction system: TC-LAPS. *Mon. Wea. Rev.*, **128**, 1245-1265.

Ebert, E.E., S. Kusselson and M. Turk, 2004: Validation of NESDIS operational Tropical Rainfall Potential (TRaP) forecasts for Australian tropical cyclones. *Aust. Met. Mag.*, submitted.

Ebert, E.E. and J.L. McBride, 2000: Verification of precipitation in weather systems: Determination of systematic errors. *J. Hydrology*, **239**, 179-202.

Ebert, E.E., M.J. Manton, P.A. Arkin, R.J. Allam, G.E. Holpin, and A.J Gruber, 1996: Results from the GPCP Algorithm Intercomparison Programme. *Bull. Amer. Met. Soc*., **77**, 2875-2887.

Ferraro, R., P. Pellegrino, S. Kusselson, M. Turk, and S. Kidder, 2002: Validation of SSM/I and AMSU derived Tropical Rainfall Potential (TRaP) during the 2001 Atlantic hurricane season. *NOAA Tech. Rep. NESDIS 105*, 43 pp. [Available from the NOAA/NESDIS Office of Research and Applications, 5200 Auth Rd., Camp Springs, MD 20746-4304.]

Ferraro, R., P. Pellegrino, M. Turk, W. Chen, S. Qiu, R. Kuligowski, S. Kusselson, A. Irving, S. Kidder and J. Knaff, 2004: The Tropical Rainfall Potential (TRaP) technique. Part 2: Validation. *Wea. Forecasting*, submitted.

Kidder et al., 2004: The Tropical Rainfall Potential (TRaP) technique. Part 1: Description and examples. *Wea. Forecasting*, submitted.

Smith, E. A., Lamm, J. E., Adler, R., Alishouse, J., Aonashi, K., Barrett, E., Bauer, P., Berg, W., Chang, A., Ferraro, R., Ferriday, J., Goodman, S., Grody, N., Kidd, C., Kniveton, D., Kummerow, C., Liu, G., Marzano, F., Mugnai, A., Olson, W., Petty, G., Shibata, A., Spencer, R., Wentz, F., Wilheit, T., Zipser, E. 1998: Results of WetNet PIP-2 Project. *J. Atmos. Sci.*, **55**, 1483–1536.

Theis, S., A. Hense, and U. Damrath, 2004: Probabilistic precipitation forecasts from a deterministic model: a pragmatic approach. *Meteorol. Applications*, submitted.

Weymouth, G., G.A. Mills, D. Jones, E.E. Ebert, and M.J. Manton, 1999: A continental-scale daily rainfall analysis system. *Aust. Met. Mag.*, **48**, 169-179.