The Australian Operational Daily Rain Gauge Analysis

Beth Ebert and Gary Weymouth
Bureau of Meteorology Research Centre, Melbourne, Australia
e.ebert@bom.gov.au

Daily rainfall data and analysis procedure

The Australian Bureau of Meteorology produces objective analyses of daily rainfall that can be used for hydrological applications, validation of model forecasts and satellite rainfall estimates, and climate monitoring. The analyses give accumulated daily rainfall for 00-24 UTC on a 0.25° grid over the land area of Australia. The latitude range of 10ºS to 45ºS allows tropical, sub-tropical, and mid-latitude regimes to be investigated.

Two rainfall analyses are produced, one in near-real time, and the other a few months later. The operational near-real time analysis is based on 9 a.m. gauge observations of 24-hr accumulated rainfall at up to 2000 synoptic and telegraphic stations, and is available a few hours after the observations are made. Figure 1 shows the location of rain gauges in Australia. The populated regions of eastern and southwestern Australia contain high gauge densities, 5-20 per 1° latitude/longitude box, while some desert regions in western and central Australia have few or no gauge observations.

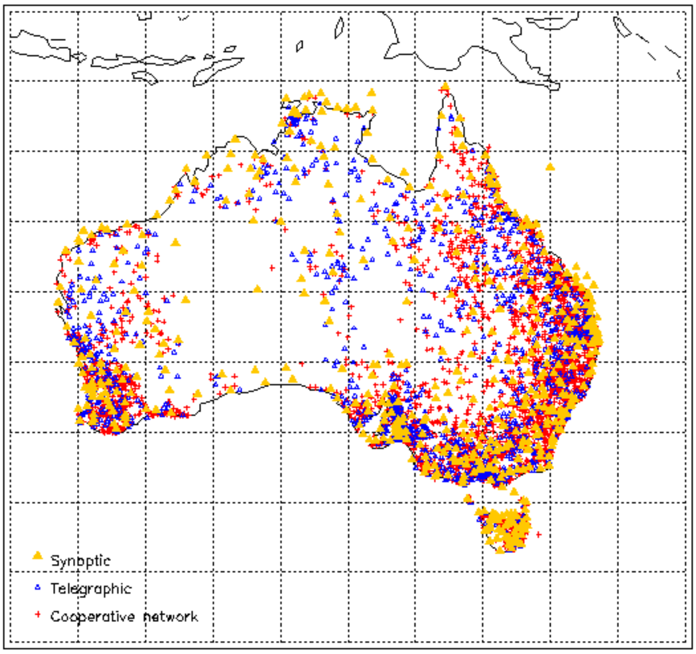

Figure 1. Australian rain gauge locations.

The 9 a.m. LST observation time corresponds to between 22 and 01 UTC, depending on season and time zone. These readings are assumed to approximate the rainfall accumulation between 00 and 24 UTC. Analysis errors associated with the timing mismatch are usually negligible, and tend to be very much smaller than those associated with the satellite estimates or NWP forecasts that they may be used to validate.

After automatic range and buddy checking to eliminate erroneous data, a three-pass Barnes successive corrections scheme is used to map the gauge data onto a 0.25° grid over the land area of Australia (Weymouth *et al.*, 1999). Analysis is not attempted in regions with no gauge coverage. The Barnes scheme estimates the rainfall at a gridpoint as a weighted average of surrounding observations, where the weights are inversely related to the distance from the gridpoint using a Gaussian weighting function. The first pass uses a length scale that is large compared with the correlation scale of the data, resulting in a relatively smooth rainfall field. Inner passes, which yield incremental refinements to the initial rainfall field, use shorter length scales so that relatively greater weight is assigned to observations close to an analysis grid point. The inner passes control the level of detail provided in the analysis, effectively acting as a filter for the data.

The operational daily rainfall analyses produced using the 3-pass Barnes scheme have been compared against analyses made using more sophisticated schemes such as Statistical Interpolation (SI) and Indicator Kriging (a scheme that uses IR satellite data to detect rain-free gridpoints before performing a kriging analysis over the remaining area; Sun *et al.*, 2002). Results showed that their bias and RMS errors differed by no more than a few percent. For daily rainfall over Australia (at least), the errors associated with incomplete spatial sampling appear to be much more important than the choice of analysis scheme. These errors are investigated further in the next section.

The more complete "climate" daily rainfall analysis incorporates more than 4000 additional gauge observations from the Cooperative Network (see Fig. 1), and is produced a few months after the event. The gauge density is improved in most regions, increasing to 15-50 per 1° latitude/longitude box in populated areas. The "climate" analysis uses variable, as opposed to fixed, length scales to take best advantage of the improved sampling in data-dense regions. In addition to being more accurate, this analysis product has reduced bias compared to the near-real time product because the observations of zero rainfall from telegraphic stations (normally non-reporting in real time) are included (Weymouth *et al.*, 1999).

Analysis errors

The usefulness of the rainfall analyses is limited without some knowledge of their expected errors. This is important for using these rainfall estimates in data assimilation and merging schemes, estimating uncertainties in hydrological or other products that may use rainfall data as input, and properly accounting for uncertainty in "ground truth" when validating other precipitation estimates (e.g., Krajewski *et al.*, 2000). In this section we present some results from an investigation on the accuracy of daily rainfall analyses produced using an experimental SI scheme. As previously mentioned, the analysis errors differed little from those of the 3-pass Barnes scheme, so the following results can be taken as representative of the operational products.

The magnitude of the analysis errors depends to a large extent on how well the observational network samples the natural variability of the rainfall. In Australia the correlation length scale (where $\rho(x)$ drops to $1/e\ \rho(x{=}0)$) is about 350 km in winter, when most Australian rainfall is associated with mid-latitude synoptic scale systems, and decreases to around 110 km during tropical summer, when convection is the primary regime. This means that in order to adequately sample daily rainfall, stations should be closer together in the tropics than elsewhere (unfortunately not the case, see Fig. 1). As a result analysis errors are expected to be larger there than in other regions and seasons.

Cross-validation techniques are commonly used to compute statistics of analysis errors at the point scale. Unfortunately, direct calculation of *grid-scale* analysis errors requires large numbers of independent observations. However, by making some reasonable assumptions about the independence of analysis errors, it is possible to use an error separation method similar to that described in Krajewski et al. (2000) to estimate their magnitudes.

Suppose two independent analyses X_1 and X_2 are produced by randomly dividing the spatial observations in half and analysing each set separately. The variance of their difference is

$$\mathrm{var}(X_1 - X_2) = \mathrm{var}[(X_1 - X_t) - (X_2 - X_t)]$$

$$= \mathrm{var}(X_1 - X_t) + \mathrm{var}(X_2 - X_t) - \mathrm{cov}[(X_1 - X_t), (X_2 - X_t)]$$

where X_t is the true value and $(X_1 - X_t)$ and $(X_2 - X_t)$ are the (unknown) analysis errors. We can expect the analysis errors to be independent of each other, since they were generated from different observations. This means we can neglect the covariance term. Given the parallel method of producing the two analyses, we can also expect that their error variances are approximately equal, i.e., $\mathrm{var}(X_1 - X_t) = \mathrm{var}(X_2 - X_t)$, and, for a sufficiently large observational dataset, a reasonable approximation to $\mathrm{var}(X - X_t)$, the error variance of the analysis based on all data. We thus arrive at a simple approximation for the analysis error variance,

$$\mathrm{var}(X - X_t) = \mathrm{var}(X_1 - X_2) / 2 .$$

The standard error is the square root of the variance, normalized by the mean value.

This approach was used to estimate analysis errors for the Australian rainfall data. Figure 2 shows the analysis standard error σ/R decreasing with increasing station density as spatial sampling is improved. It also decreases with increasing rainfall accumulation. Similar behaviour was demonstrated by Huffman (1997) and others for satellite precipitation analysis errors.

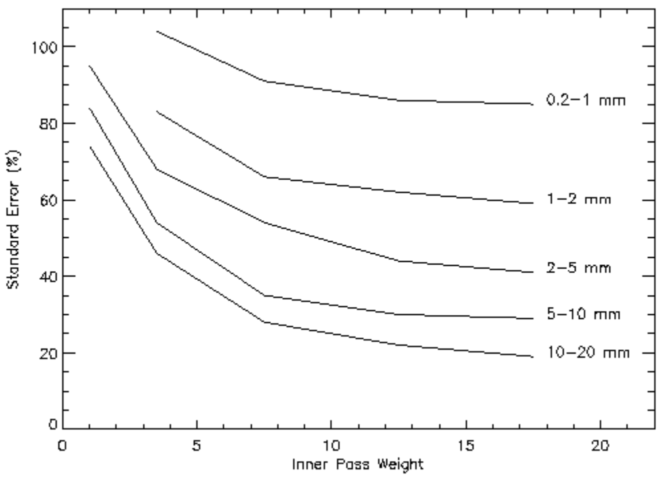

Figure 2. Standard error (%) in 0.25° grid boxes as a function of summed inner pass weight from the Barnes analysis (related to station density), for five ranges of daily rainfall accumulation.

The percent standard error in Fig. 2 can be empirically fit by a function of the form,

$$\frac{\sigma}{R}(\%) = a + b \ln W + cR^{-1/2}$$

where W is the summed inner pass weight, R is the analysed rain accumulation, and the coefficients have the values a=65, b=-21, and c=50. This expression can be used to generate error estimates to accompany the rainfall analyses.

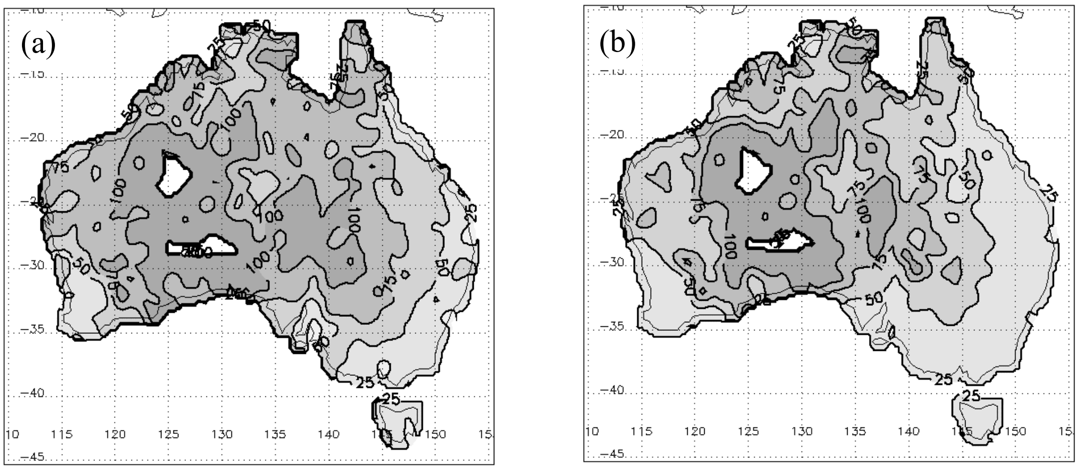

Figure 3. Annual mean analysis error standard deviation, expressed as a percentage of annual mean rain during January-December 2000, for (a) near-real time daily analyses, and (b) "climate" daily analyses.

The spatial distribution of analysis errors (Figure 3) shows that the standard errors average between 25% and 50% in eastern and far southwestern Australia, but can exceed 100% in the sparsely sampled regions of central Australia. Tropical values range between 50% and 100%, with lower errors in the "climate" daily analyses than in the near-real time analyses. Note that errors on any given day are likely to be much lower, or even zero, in areas with no observed rainfall, or much higher in the case of regions with convective rainfall.

The analysis errors can be reduced somewhat by spatial averaging. Because the rain field is spatially correlated, the error reduction is less than would be predicted for independent samples. For the Australian analyses, averaging from a 0.25° grid scale to a 1° grid scale decreases the analysis errors by only about 25%.

Future work

BMRC is working toward producing a combined rain gauge-radar analysis over Australia on hourly and daily time scales. The greatest improvements in the daily rainfall analyses are expected to be in the coastal regions of western and tropical Australia that have radar coverage.

Recommendations

1. Make use of Australian gauge data and gridded analyses for hydrological applications, validation of NWP forecasts and satellite precipitation estimates, and climate monitoring.

2. Provide error estimates with the analyses, to assist users in making the most sensible use of the data.

3. Encourage the development and use of validation methodologies that account for the uncertainties in the "ground truth" data.

4. Make use of the full gauge "climate" Australian dataset, not just the values put out on GTS, in producing merged rainfall analyses.

Obtaining Australian daily gauge data and gridded analyses

The near-real time daily rain gauge data and gridded analyses can be downloaded via FTP from the Bureau of Meteorology's National Climate Centre at the URL,

ftp://ftp.bom.gov.au/anon/home/ncc/www/rainfall/totals/daily.

The more complete "climate" rain gauge data and analyses, based on up to 6000 observations per day, can be obtained on request from the first author.

References

Huffman, G.J., 1997: Estimates of root-mean-square random error for finite samples of estimated precipitation. *J. Appl. Meteor.*, **36**, 1191-1201.

Krajewski, W.F., G.J. Ciach, J.R. McCollum, and C. Bacotiu, 2000: Initial validation of the Global Precipitation Climatology Project monthly rainfall over the United States. *J. Appl. Meteor.*, **39**, 1071-1086.

Sun, X., M.J. Manton, and E.E. Ebert, 2003: Regional rainfall estimation using double-kriging of raingage and satellite observations. *BMRC Research Report No. 93*, 41 pp.

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