18th World IMACS / MODSIM Congress, Cairns, Australia 13-17 July 2009
http://mssanz.org.au/modsim09

A comparison of the BAWAP and SILO spatially interpolated daily rainfall datasets

Beesley, C. A. [1], **A. J. Frost** [1] **and J. Zajaczkowski** [2]

[1] *Water Division, Australian Bureau of Meteorology*
[2] *Queensland Climate Change Centre for Excellence, Department of Environment and Resource Management*
Email: c.beesley@bom.gov.au

Abstract: High-quality gridded rainfall datasets are required for modelling of hydrological systems and other environmental applications. For example, Chiew et al. (2008) use daily 5 km gridded interpolated historical rainfall, scaled according to projected changes in rainfall distributions provided by scenarios, for assessing the local impact of climate change over the Murray Darling Basin. Biases and errors in the interpolated surface can, therefore, adversely affect the quality of such calibrations and the resulting estimates of current/future water availability. Two products containing archived Australia-wide gridded (0.05° by 0.05°) gauge-based daily rainfall are currently publicly available: the widely used SILO product (www.longpaddock.qld.gov.au/silo) produced by the Department of Environment and Resource Management; and the new freely available product produced by the Bureau of Meteorology as part of the Australian Water Availability Project (www.bom.gov.au/jsp/awap), herein denoted BAWAP.

This study presents the first direct comparison of the two surfaces using a range of daily rainfall cross-validation statistics, including mean error, mean absolute error and root mean squared error. Cross-validation datasets have been produced by the respective organizations responsible for SILO (leave-one-out cross validation) and BAWAP (1% cross validation) for data spanning 2001-2007. As described in the paper, differing subsets of the total available sites were used in production of the cross-validation surfaces, however, using records common to both, summary error statistics have been calculated for each of the datasets.

The results indicate that the error statistics are similar for both methods, with the SILO method producing slightly lower error statistics overall for the 2001 to 2007 period. This result is clouded by the fact that only a subset of observations (stations with complete months) were used in the SILO cross-validation analysis which is not representative of the current operational process. The BAWAP and SILO methods both contain small positive bias of rainfall amount on dry days (observed rainfall equal to 0mm) and a negative bias of amount on wet days (rainfall greater than 0 mm) increasing with magnitude of the rainfall observation (indicated in Figure 1). Both methods overestimate the number of wet days in the analysis period due to smoothing effects of the interpolation methods at the edges of rainfall events, however this is more pronounced in the BAWAP dataset. A geographical comparison of BAWAP and SILO error statistics suggest that there are only small differences in the errors across most of the continent. However, there are higher BAWAP errors along the east coast of Australia, particularly in high gauge density regions. Also, root mean square error statistics generally increase the closer the sites are to the equator. The positive bias for dry day amounts and negative bias for larger rainfall amounts result in a negative/positive bias for higher/lower rainfall areas.

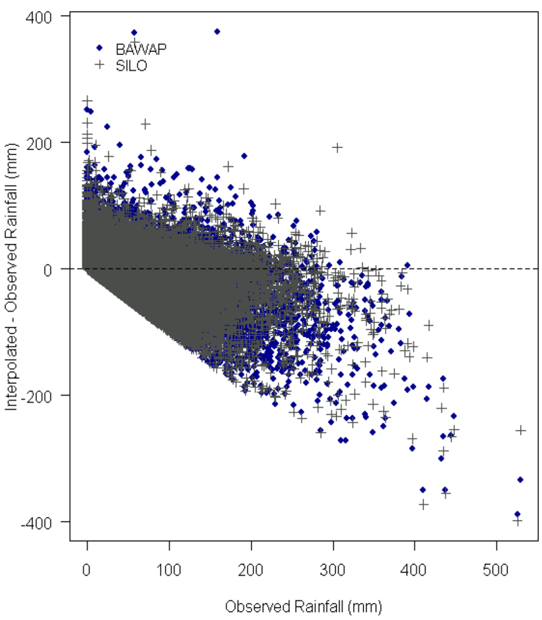

Figure 1. Residual (Interpolated-Observed) versus observed daily rainfall.

Keywords: *Rainfall, SILO, AWAP, cross validation*

1. INTRODUCTION

Rainfall is highly variable over area and time and, therefore, is difficult to reliably interpolate from surrounding station observations (Jeffrey et al., 2001; Hutchinson, 1998a,b). Some factors that contribute to the complexity of daily rainfall analyses include various sources of observation error, irregular geographical distribution of rainfall observations, under-representation of high elevation areas which often tend to have higher rainfall and the influence of topography.

Two archived Australia-wide gridded gauge-based interpolated daily rainfall products are currently publicly available: the SILO and BAWAP datasets. Both products provide surfaces of spatially-interpolated observed daily rainfall at a spatial resolution of 0.05° x 0.05° and attempt in different ways to incorporate information about the influence of topography on rainfall in order to more reliably interpolate daily rainfall.

The SILO method (documented in Jeffrey et al. (2001)) is produced by the Queensland Climate Change Centre of Excellence (QCCCE) within the Queensland Department of Environment and Resource Management (formerly within the Environmental Protection Agency) and is available commercially as patched point records (with missing data at rain-gauge locations being 'patched' through interpolation) and as a synthetic dataset generated over a set of evenly spaced grid locations, referred to as the 'drilled data'. This product has been used extensively in many Australian hydrological studies (e.g. Chiew et al., 2008). As discussed in Jeffrey et al. (2001) and Jeffrey (2006), the current SILO operational product is based on a three step process: a) a thin plate smoothing spline (Hutchinson, 1993, 1998a,b) is used for interpolation of monthly rainfall climatology normalisation parameters; b) ordinary kriging (a geostatistical technique – see Cressie (1993)) is then applied to the normalised monthly rainfall; and c) the monthly values are disaggregated to daily values using the relative temporal distribution generated by ordinary kriging of the observed daily values within each month.

The BAWAP method was developed by the Bureau of Meteorology (BoM) National Climate Centre (NCC) for the Australian Water Availability Project (http://www.csiro.au/awap/). This product also provides patched and interpolated datasets from 1900 onwards (Jones et al., 2007). This method is based on the method used to derive operational daily rainfall grids produced by the NCC from real-time data records (http://www.bom.gov.au/climate/austmaps), and hence has been used extensively to produce the daily operational maps output by the BoM. The BAWAP method uses a two-step process: a) a thin plate smoothed spline is used for interpolation of monthly rainfall climatology; and b) interpolation of anomalies of daily rainfall (expressed as a percentage of the climatological rainfall) using Barnes' successive correction method (Mills et al., 1997; Jones and Weymouth, 1997).

Previous error analyses for the methods described by Jeffrey et al. (2001) and Jones et al. (2007) are incomparable as they were undertaken for differing underlying observed data. For example, the errors were derived for different analysis periods, a different method was used to calculate the summary statistics, some of the rainfall station coordinates had been updated since the SILO study and more general quality of BoM data has been undertaken between those studies. Revised error statistics have since been published based on the analysis period 2001-2007 (Jones et al., 2009; Zajaczkowski, 2008). The present study details an independent comparison of the two datasets using common cross-validation statistics.

2. METHOD

2.1. Cross Validation

Cross validation is a method that is commonly used to assess the error associated with interpolated climate data (or more generally where estimation techniques are being used for prediction). One or more observations at a time are omitted from the analysis and a value is interpolated at the location of the omitted station. The difference between the interpolated value and the actual observed value at that location is used to assess the accuracy of the interpolation. Typically, this is repeated for each observation in turn. Cross-validation analysis relies on the calculation of several metrics to measure the performance of the interpolated field compared with the observed site values. These metrics are defined here as they are used within both past analyses, and calculations in this study. The three typical measures are the mean error (ME - also referred to as bias), mean absolute error (MAE) and root mean square error (RMSE). These metrics are calculated according to the following equations:

$$ME = \frac{1}{n}\sum_{i=1}^{n}\left(E_i - O_i\right) \qquad (1)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|\left(E_i - O_i\right)\right| \tag{2}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(E_i - O_i\right)^2} \tag{3}$$

where E_i is the interpolated (estimated) value at a station on a particular day, O_i is the observed value at the same station on the same day and n is the total number of station-days included in the analysis.

2.2. Data

Given the incompatibility of previous studies (and underlying cross-validation datasets), QCCCE and BoM NCC were approached to provide revised data. The following data were supplied for this analysis:

- 'leave-one-out' cross validation for a method following the SILO operational method (daily disaggregation of interpolated monthly rainfall) for each day for the period 2001 to 2007;
- 'leave-one-out' cross validation for direct ordinary kriging of daily rainfall data (SILO operational method for days in the most recent month) for each day for the period 2001 to 2007; and
- daily cross validation for the BAWAP method leaving 1% of stations out at a time (100-fold cross validation) for the period 2001 to 2007.

The slight difference in terms of cross-validation techniques (100-fold BAWAP cross validation versus leave-one-out SILO cross validation) is not expected to affect model results significantly, as sites were selected randomly within the 100-fold cross validation.

Importantly, the SILO dataset provided for the monthly disaggregation method varies from the method applied operationally in two ways: a) the pre-processing quality control procedure currently undertaken as part of the operational method (removal of outliers according to various rules) was not applied prior to final interpolation; and b) only stations with a complete set of daily values for any given month (i.e. no missing/accumulated rainfall days for a given month) and which satisfy other gross error quality criteria were used in the analyses. The latter means that the SILO monthly disaggregation cross-validation dataset supplied typically contains only approximately 4000 sites on any one day. This contrasts with the SILO operational methodology (and within the BAWAP supplied method) where there are typically 6000 sites included on any given day.

Due to these reasons, the cross-validation statistics will not entirely reflect the current operational method used within SILO and the results must thus be interpreted in light of this. This difference from the method undertaken operationally was introduced as cross validation data was not able to be output using the existing operational methodology. The methodology was recoded from scratch for the purposes of this study. However, it was found that incorporation of the complex coding required to undertake monthly disaggregation for months containing missing/accumulated data was not possible within the time constraints of this project. This inadequacy may be addressed in a future study. A second comparison (reported in Beesley et al. (2009) but not reported here due to space limitations) was undertaken using the SILO direct ordinary kriging data (which includes all available sites) and the BAWAP methodology, thus providing a reasonable basis for comparison for the operational methodologies (as direct ordinary kriging is used as a component of the current SILO operational method). That comparison showed that the SILO method produces slightly greater error statistics in that case. It is also noted that not all of the available sites were used in the BAWAP interpolation (those sites which have been opened recently).

2.3. Error Analysis

With the supplied data, a cross-validation comparison of methods for the period spanning 2001 to 2007 has been undertaken. All analyses were undertaken using the freeware statistical software package R (R Development Core Team, 2008). Only records existing in both the BAWAP and SILO datasets were included in the error calculations to make the results as comparable as possible. However, additional stations may have been used in the generation of either the BAWAP or SILO surfaces that were not used in the other and, therefore, any areas with additional good quality data points might be expected to have a better overall accuracy.

3. RESULTS

Summary error statistics for the BAWAP and SILO cross-validation datasets provided for the analysis period 2001 to 2007 are presented in Table 1. The most recently published error statistics for both the BAWAP and SILO methods (based on the same data) are also provided in the table. It is noted that these previously published statistics use differing methods of calculation of validation statistics. Error statistics for the direct kriging of daily rainfall are provided in Table 1. However, unless otherwise specified, subsequent figures and tables refer to the SILO monthly disaggregation method (as this is the method closest to that used operationally for historical data).

The figures in Table 1 suggest that the overall errors associated with the cross-validation datasets are very similar. For the statistics calculated in this study, both the SILO datasets appear to have a slightly higher overall accuracy when compared with the BAWAP error statistics. The SILO direct daily rainfall interpolation (using approximately 4000 sites) and the monthly disaggregation methods (similar to that used operationally but only using 2/3 of sites) have very similar accuracies.

Table 1. Summary error statistics for the BAWAP and SILO cross-validation daily rainfall datasets for the 2001 to 2007 analysis period.

Dataset	Method	Number of observations	ME (mm)	MAE (mm)	RMSE (mm)
All Stations - All days	BAWAP	12049752	0.01	0.85	3.43
	SILO – monthly disaggregation		-0.02	0.71	3.06
	SILO – direct daily kriging		0.004	0.70	3.02
	BAWAP (statistics from Jones *et al.*, 2009)	NA	0.0	0.9	3.1
	SILO - monthly disaggregation (statistics from Zajaczkowski, 2008)	NA		0.71	2.6
All Stations - Wet days (observed rain > 0mm)	BAWAP	2705631	-0.61	3.16	6.98
	SILO – monthly disaggregation		-0.54	2.70	6.18
	SILO – direct daily kriging		-0.46	2.64	6.07

The ME results in Table 1 indicate underestimation of rainfall for wet days (days with greater than 0mm observed rainfall) for both the SILO and BAWAP methods. This is generally a feature of weighted average interpolation techniques, as the interpolated value can never be greater than the surrounding observed values. When all observations are included, the bias is reduced and reversed. It is suggested that this over/under-dispersion at low/high rainfall values is due to the fact that neither method accounts for the zero bounded nature of rainfall, nor the increasing variability of rainfall with increasing intensity – see Beesley et al (2009) for further discussion.

To investigate the apparent wet day amounts bias in the interpolated fields, the observed versus BAWAP and SILO interpolated rainfall estimates for all data are plotted in Figure 1. Due to the weighted average interpolation methods, it is generally not expected that the interpolated rainfall will significantly exceed the observed rainfall, unless there is an observational error at the target station, an error in the station coordinates or the surrounding stations experience higher rainfalls. It is observed that for both methods the absolute errors tend to increase with increasing observed rainfall amount.

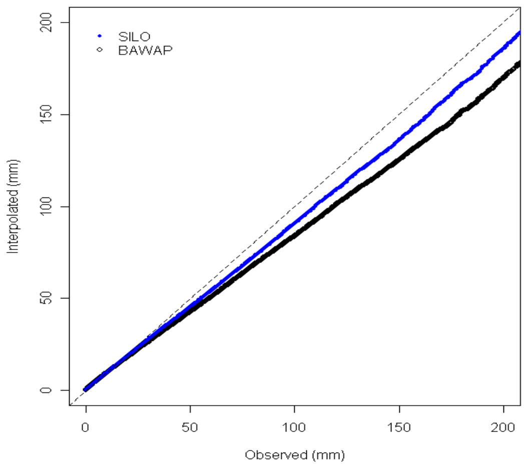

Figure 2. Quantile-quantile plot of the observed versus interpolated daily rainfall.

Table 2. Frequency of interpolated values as a ratio of the frequency of observed values.

Rainfall amount (mm)	**0**	**0-2**	**2-10**	**10-100**	**>100**
SILO	0.90	1.7	1.12	1.00	0.76
BAWAP	0.80	2.4	1.25	1.00	0.60

Figure 2 displays the differences between the BAWAP and SILO rainfall distributions (in the form of a quantile-quantile plot) and the observed rainfall distribution based on the pooled recorded rainfall values over all sites Australia-wide. Quantile-quantile plots show the overall distribution rather than the observed versus interpolated rainfall at a station for a particular date. This plot suggests there is a tendency for both methods to underestimate the higher observed rainfalls and that this is more pronounced in the BAWAP interpolations. Table 2 presents the frequency of rainfall values for each method compared to the frequency of observed values. Both methodologies underestimate the number of observed zero values, whilst tending to overestimate the values up to 2 mm, caused by the smoothing effect of the weighted average interpolation techniques here. This effect is also highlighted in Figure 3, which shows an increase in the number of wet days across Australia for both interpolation methods, particularly along the southern coast. These results are important in terms of the interpretation of any analysis on wet and dry days (eg. trends in dry/wet spell length) based on this data. It is noted that such studies usually apply a threshold of approximately 1mm to define wet days, so as to reduce the impact of inconsistently recorded trace amounts. It is also noted that the impact of the underestimation of the number of dry days on rainfall and rainfall-runoff modeling studies is somewhat offset by the overestimation of the number of small rainfall values observed (0 mm – 2 mm), as these values typically do not contribute significantly to annual rainfall or runoff totals.

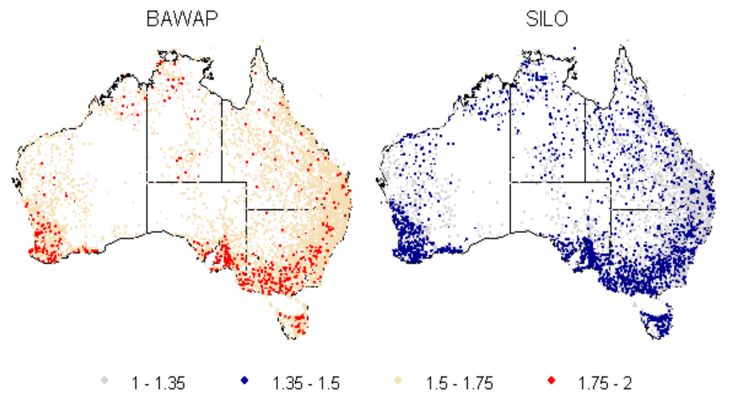

Figure 3. Ratio of the number of wet days (rain > 0mm) interpolated versus observed for stations with more than 300 records.

In order to investigate trends in spatial performance of each method of the continent, a geographic representation of the error statistics is presented. Figure 4 shows the ME and RMSE statistics for the two methods at each site Australia-wide for wet days only, and Figure 5 shows the differences between the RMSE for the two methods. Due to lack of records at some locations during the analysis period, these figures display only stations with a minimum of 100 observations.

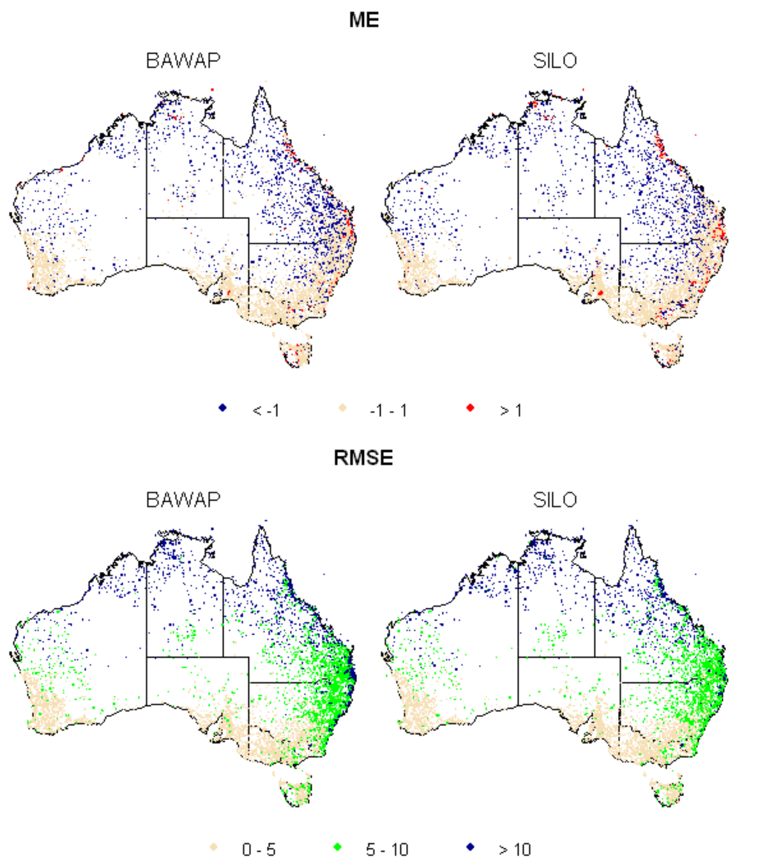

Figure 4. Geographical representation of ME and RMSE error statistics for observed wet days only (i.e. where the observed rainfall is greater than 0 mm) for individual stations (with a minimum of 100 records).

When only wet days are considered, the bias of both interpolation methods towards underestimating rainfall is apparent, particularly in the higher rainfall regions around the eastern and northern coasts of Australia. The bias appears to increase and fluctuate between positive and negative in mountainous areas, such as the north Queensland coast near Cairns, along the Great Dividing Range in southeast Australia and in southwest Tasmania, which may be indicative of the complex influence of topography on rainfall events. In southern Australia, there is an increased tendency to overestimate the rainfall. This may be the due to the dominant occurrence of more widespread rainfall patterns associated with rain bands in this area compared with the smaller scale, high rainfall convective storms that occur more frequently in the north of Australia during the summer months. These relatively low rainfall larger area events may exhibit lower variability and may, therefore, be expected to have less error and an increased frequency of overestimation as the rainfall is spread over a larger area.

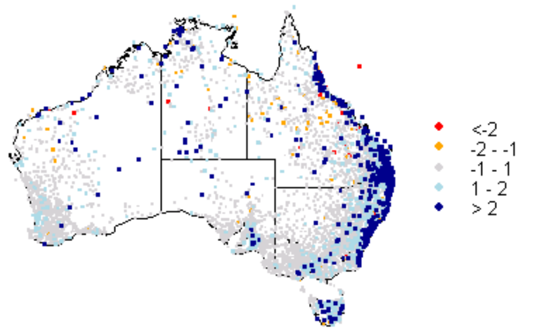

Figure 5. Differences between the BAWAP and SILO station RMSE for observed wet days (for stations with a minimum of 100 records).

The RMSE maps (Figure 4) display similar patterns for both the SILO and BAWAP station errors. Errors are lowest in southern Australia, with an increasing gradient towards the north, east and west coasts. The largest errors occur in the northern region of Australia. When only wet days are considered, the pattern is similar but the error statistics increase in magnitude.

In the RMSE difference plots shown in Figure 5, there are small differences in the errors across most of Australia when all records are included in the analysis. Stations with lower BAWAP errors appear to be relatively evenly distributed with lower SILO error stations. However, there are higher BAWAP errors along the east coast, particularly in high observation density areas. A possible reason for this may be that in high rainfall areas, the variable range parameter used in the SILO method allows higher resolution of smaller scale rainfall features in high gauge density areas and also that this reduces the impact of edge effects associated with having no data out to sea. Higher quality rainfall records associated with city areas could also improve the accuracy of the SILO results due to the absence of an observational error term (nugget) in the SILO method. Similar patterns, but with increased magnitude of errors, are apparent when only observed wet days are considered. The SILO dataset has lower error for several stations in the data-sparse tropical areas, which may indicate that the station density-dependent range or length scale associated with the SILO method may improve the prediction in those locations. These biases (and errors in general) are discussed in greater detail and further investigated through an annual rainfall analysis, showing significant underestimation by both methods for the (runoff producing) higher altitude coastal fringes of Australia in Beesley et al.(2009).

4. DISCUSSION AND CONCLUSIONS

Error analyses have been undertaken previously for the SILO and BAWAP daily rainfall interpolation methods. However, the ability to compare the results of these analyses is hindered by the use of different datasets and calculation methods. In an effort to achieve a more direct comparison, the providers of these interpolation methods were approached to provide updated cross-validation statistics for comparison purposes.

Based on the analyses of the available datasets, the results indicate that the SILO method has produced slightly better error statistics overall for the 2001 to 2007 comparison data. However, the results are somewhat clouded by the use of a subset of the data available for each method in the analysis. In particular, for the data provided for the method closest to that used operationally in SILO, only 2/3 of data were used. Both the BAWAP and SILO methods have been shown to contain biases for differing levels of rainfall (generally small positive bias of amounts on dry days and negative bias on wet days).

Using the sites common to both analyses for the data provided, there are higher BAWAP errors particularly along the east coast of Australia in high gauge density regions. The following reasons are suggested for this performance:

- The SILO method does not contain a measurement error term (nugget) and, therefore, may have a higher accuracy when good quality high density data is used (i.e. in cities);
- The range or length scale factor used in the SILO method varies with observation density and may, therefore, capture the smaller scale rainfall variability better in high density, high rainfall areas;
- Given the distance weighted nature of the BAWAP estimate, if anywhere within the given weighting radius there is a positive rainfall observed, there will be a positive rainfall interpolated at the target site. This is not the case for kriging methods which can allow for a greater degree in sparsity of rainfall; and

- The BAWAP analysis uses 1% cross validation compared to leave-one-out cross validation for SILO. This is expected to slightly degrade the relative results of BAWAP compared to SILO. This is somewhat offset, however, by the fact that much fewer sites are used in the SILO monthly disaggregation method.

Overall the SILO and BAWAP operational methods performed very similarly within the cross validation comparison. However, significant errors in terms of RMSE and bias showed spatial coherence Australia-wide, with northern areas typically under-representing the variability present, as indicated by higher RMSE values. It is suggested that this is caused by the underestimation of convective rainfall variability (especially in summer), which typically has a shorter spatial correlation length than elsewhere, which further causes underestimation if an Australia-wide correlation length that fits all types of events is used. It is noted that the analysis period used here is relatively short (seven years only) and corresponds with a relatively dry period of the Australian rainfall, particularly in the south and eastern part of the continent. This is likely to mean that the statistics presented in this study are potentially low compared with the errors associated with the BAWAP and SILO datasets over a longer, wetter period. Furthermore, the pooled (spatially and temporally) error statistics presented tend to be dominated by stations with complete records and higher rainfall mean/variability (as highlighted by the spatial distribution of errors presented). The comparison presented here is presented in greater detail in report form in Beesley et al. (2009).

ACKNOWLEDGMENTS

We express our gratitude to William Wang and David Jones from the National Climate Centre, Bureau of Meteorology for their data provision and advice regarding the BAWAP daily rainfall data product. We also greatly appreciate the efforts of Stephen Jeffrey and Ken Day from QCCCE in organizing the input of Juliusz Zajaczkowski for coding and calculation of the SILO cross validation datasets. Thanks also to Alan Seed, Doerte Jacob, Beth Ebert, Sri Srikanthan, Eloise Nation from the BoM and two anonymous reviewers for providing helpful input and reviews of this work.

REFERENCES

Beesley, C. A., Frost, A. J., Zajaczkowski, J. (2009) In preparation, A comparison of the BAWAP and SILO spatially interpolated daily rainfall datasets, Water Division, Bureau of Meteorology.

Chiew, F. H. S., Vaze, J., Viney, N. R., Jordan, P. W., Perraud, J-M., Zhang, L., Teng, J., Young, W. J., Penaarancibia, J., Morden, R. A., Freebairn, A., Austin, J., Hill, P. I., Wiesenfeld, C. R. and Murphy, R. (2008), Rainfall-runoff modelling across the Murray-Darling Basin, A report to the Australian Government from the CSIRO Murray-Darling Basin Sustainable Yields Project, Water for a Healthy Country Flagship. CSIRO. 70 pp.

Cressie, N. (1993), Statistics for Spatial Data, John Wiley and Sons, 900pp.

Hutchinson, M. F., 1993. On thin plate splines and kriging, In: Tarter ME and Lock MD.(eds), Computing and Science in Statistics 25, University of California, Berkeley: Interface Foundation of North America, 55-62.

Hutchinson, M. F. (1998a), Interpolation of rainfall data with thin plate smoothing splines: I two dimensional smoothing of data with short range correlation, *Journal of Geographic Information and Decision Analysis*, 2, 152-167.

Hutchinson, M. F. (1998b), Interpolation of rainfall data with thin plate smoothing splines: II analysis of topographic dependence, *Journal of Geographic Information and Decision Analysis*, 2, 168-185.

Jeffrey, S. J., Carter, J. O., Moodie, K. B. and Beswick, A. R. (2001), Using spatial interpolation to construct a comprehensive archive of Australian climate data, *Environmental Modelling & Software*, 16, 309-330.

Jeffrey, S.J. (2006), Error analysis for the interpolation of monthly rainfall used in the generation of SILO rainfall datasets.

Jones, D. A. and Weymouth, G. T. (1997), An Australian monthly dataset, Tech Report 70, Bureau of Meteorology.

Jones, D. A., Wang, W. and Fawcett, R. (2007), Climate data for the Australian Water Availability Project, final milestone report. Bureau of Meteorology.

Jones, D. A., Wang, W. and Fawcett, R. (2009), High-quality spatial climate analyses for Australia, *Submitted to* Australian Meteorological Magazine.

Mills, G. A., Weymouth, G., Lorkin, J., Manton, M., Ebert, E., Kelly, J., Jones, D. and deHoedt, G. (1997), A national objective daily rainfall analysis system, BMRC Research Report, Bureau of Meteorology.

R Development Core Team, (2008), R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. ISBN 3-900051-07-0. URL http://www.R-project.org.

Zajaczkowski, J. (2008), Cross-validation of rainfall data, Queensland Climate Change Centre of Excellence, Queensland Environmental Protection Agency.