Journal of Hydrology 493 (2013) 105–114

Contents lists available at SciVerse ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Evaluating geostatistical methods of blending satellite and gauge data to estimate near real-time daily rainfall for Australia

Adrian Chappell [a,1,*], Luigi J. Renzullo [a], Tim H. Raupach [a], Malcolm Haylock [b]

[a] CSIRO Land and Water, GPO Box 1666, Canberra, ACT 2601, Australia
[b] Partner Reinsurance Company, Zurich, Switzerland

ARTICLE INFO

Article history:
Received 30 August 2012
Received in revised form 14 January 2013
Accepted 13 April 2013
Available online 24 April 2013
This manuscript was handled by Konstantine P. Georgakakos, Editor-in-Chief, with the assistance of Hervé Andrieu, Associate Editor

Keywords:
Gauge
Satellite
Precipitation
Geostatistics
Cokriging
Rainfall intermittency

SUMMARY

Rain gauges provide valuable information about the amount and frequency of rainfall. In Australia, the majority of rain gauges are located in populated, wet coastal regions. Approximately 2000 gauges reporting within 24 h of a target day were used to make near real-time (NRT) estimates of daily precipitation. The remaining ≈4000 gauges for the same target day were used to evaluate bias and estimation performance using several traditional statistics. There is considerable potential to improve the estimation of rainfall in Australia using related ancillary data, particularly in sparsely gauged areas. The Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis (TMPA-RT) near real-time product (3B42RT) provided images (0.25° resolution) of precipitation across Australia. Daily precipitation was estimated in 2009/10 approximately every 5 km across Australia. This study evaluated selected geostatistical methods for estimating daily rainfall maps across Australia. It tackled the change of support problem and spatial intermittency of daily rainfall data in blending satellite and gauge data. Dissension occurred amongst traditional global statistical measures of performance which were compounded by extremes of gauge density. Overall, our assessment is that blending the 3B42RT satellite and rain gauge data was not worthwhile. However, the blending considerably reduced the estimation variance indicating that uncertainty of the map estimates was a neglected property necessary to detect change and difference in patterns.

1. Introduction

Knowledge of the amount, spatial distribution and temporal variation of daily precipitation is essential for provision of public information, hydrological modelling and flood forecasting, climate monitoring and climate model validation. Rain gauges provide valuable information about the amount and frequency of rainfall. However, they are commonly located in, or near to, population centres to provide timely information for water resource management. Automated weather stations including rain gauges are too expensive to form a dense network of measurements. In Australia, the majority of rain gauges are located in populated, wetter coastal regions and relatively few stations are located in drier, inland regions. Areas of high rain gauge density provide more reliable estimates at unsampled locations than areas of low rain gauge density.

The estimation of the spatial distribution of rainfall can be improved with the inclusion of ancillary data (radar, satellite and/or topography) that is related to the rainfall. For example, Krajewski (1987) found improved estimates when blending rain gauge and radar information with bias. Notably, he found little improvement in the blending when the bias in the radar data was removed. Goovaerts (2000) augmented monthly rainfall totals for Portugal with digital elevation data. Velasco-Forero et al. (2009) showed that radar data considerably improved the estimation of hourly rainfall in Catalonia, Spain. Although weather radar is very useful for augmenting rainfall estimation, radar networks in Australia are limited to the coastal regions. Satellite-based estimates provide synoptic images of the spatial distribution of precipitation events at 0.5–3 h intervals, at between 0.07° and 0.25° resolution (Joyce et al., 2004; Kubota et al., 2007; Huffman et al., 2007). While the absolute accuracy of satellite rainfall products is questionable (Tian and Peters-Lidard, 2011) and remains the subject of on-going worldwide assessment (Ebert et al., 2007; IPWG, 2011), they nevertheless provide unique information on the spatial extent of the rainfall, particularly for those regions of the Australian continent where the network of rain gauges is sparse. It is only relatively recently that the statistical blending of satellite-derived precipitation products and rain gauge measurements has been explored to generate fine-resolution rainfall estimates at continental scales (Vila et al., 2009; Li and Shao, 2010; Clarke et al., 2011).

* Corresponding author. Tel.: +61 2 6246 5925; fax: +61 2 6246 5965.
E-mail address: adrian.chappell@csiro.au (A. Chappell).
[1] Present address: Environmental Remote Sensing Laboratory (LTE), École Polytechnique Fédérale de Lausanne (EPFL), GR C2 564, 1015 Lausanne, Switzerland.

0022-1694/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jhydrol.2013.04.024

There remain two key issues with the blending of daily rain gauge measurements and satellite data: (i) the change of support problem (COSP) and (ii) spatial intermittency of daily rainfall data. There is a considerable difference in the resolution of information with satellite data providing either radiance measurements or rainfall estimates over a large pixel ($0.25^\circ \times 0.25^\circ$) whilst rain gauge observations represent such a small area they may reasonably be considered to be a point. Any blending of these different sources of data should adequately account for these differences. The support of a sample has come to mean simply the volume of a datum. The COSP is how the spatial variation in one variable associated with a given support relates to that of another variable with a different support (e.g., Gotway and Young, 2002; Kyriakidis, 2004). Spatial intermittency is related to the delivery of daily rainfall in discrete patches in space and time. This causes a discontinuous surface of rainfall i.e., areas of zero rainfall between areas of non-zero rainfall (Creutin and Obled, 1982). Spatial intermittency undermines the common assumption of continuity in the geostatistical treatment of the data. However, Barancourt et al. (1992) showed that a straightforward method to tackling this mixed distribution is to threshold the rainfall distribution with an indicator transform, map the presence and absence of rainfall, and combine it with a wet area map.

This study is an evaluation of geostatistical techniques for blending satellite and gauge data to estimate daily rainfall maps across Australia. Only geostatistical techniques will be considered based on the premise that rainfall estimates of these techniques give more accurate results than methods of interpolation (Grimes and Pardo-Igúzquiza, 2010; Hofstra et al., 2008). Many studies (e.g., Creutin and Obled, 1982; Lebel et al., 1987; Goovaerts, 2000) have used geostatistics to estimate rainfall. Kriging is the central tool for geostatistics and our experiments will be conducted using only kriging-based techniques. The aim of this paper is to demonstrate the impact of upscaling and rainfall intermittency on the blending of satellite data with rain gauge observations for the spatial estimation of daily rainfall. Point estimates of daily rainfall are made at 5 km intervals across Australia every day for an intensive observation period (IOP) between 31 May 2009 and 31 May 2010. To achieve the aim, we have two objectives and employ several estimation techniques. The first objective is to explicitly tackle the COSP defined above. We use block-kriging to make rainfall estimates with the same support as the satellite data and develop an appropriate cross-variogram before using cokriging and the gauge data to make the maps. The second objective is to consider the impact of (inappropriately) assuming continuity in daily rainfall which may be intermittent over space. The two-stage process of Barancourt et al. (1992) is used with the techniques described previously to account for the COSP. We compare these estimation techniques with ordinary point kriging by employing several validation metrics to isolate the benefits of the satellite data and to determine how best to blend the satellite data. The local estimation variance maps demonstrate the impact of the number and configuration of rain gauge information and consequently provide a rudimentary measure of a technique's relative uncertainty.

2. Precipitation data

2.1. Australian rain gauge networks

Daily gauge data used in our investigations covering the period 31 May 2009 – 31 May 2010 were obtained from the data holdings of the Australian Bureau of Meteorology (BoM). No additional quality checks were performed on these data. The observations are interpreted as the 24-h accumulated rainfall to 9 am (local time) on the day of interest. The data include the measurements from the network of gauges reporting in near real-time (NRT) i.e., within 24 h of 9 am on the day of interest, and the measurements from a larger post real-time (PRT) monitoring network reporting daily rainfall for the day of interest some days to months later (Fig. 1).

It is common for there to be twice as many PRT gauges as NRT gauges. The significance to this study of the differences in which observations of rainfall are recorded by the BoM is the opportunity to use the PRT network of rain gauge observations to test the performance of the mapping techniques that use the NRT set of observations. Approximately 2000 NRT rain gauges were available daily and were used to make estimates at unsampled locations across Australia. A further approximately 4000 PRT rain gauges available for each day were set aside and used to evaluate the performance of the mapping techniques.

2.2. Satellite-based precipitation products

Satellite-based precipitation estimation began in the 1970s with the advent of weather satellites. Multi-channel radiometers aboard these geosynchronous platforms provided visible and thermal infrared (IR) imagery of the Earth's surface at unparalleled spatial and temporal resolutions (Kummerow, et al., 2000). In the 1980s passive microwave (PM) sensors aboard polar-orbiting satellites provided stronger relations with precipitation, and estimates derived from these data are generally considered superior to those derived from IR observations. A number of algorithms have been developed that combine precipitation observations from both types of satellite observations, thus exploiting the high spatial and temporal coverage of the geostationary IR estimates with the more accurate PM-based estimates (Huffman et al., 1997; Joyce et al., 2004; Kubota et al., 2007).

The Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis (TMPA) system generates a range of quasi-global precipitation products with data ranging from 1 January 1998 to present (Huffman et al., 2007). Of interest in this work is the near real-time product (referred to as 3B42RT) that combine IR and PM estimates within a 3 h window centred on a synoptic time (0000, 0300, 0600, ..., 2100 UTC). Huffman et al. (2007) provides details on this product, including the calibration of the multiple PM and geostationary IR observations. These precipitation estimates were geographically subset for the Australian continent in this study. The spatial resolution and temporal sampling of the data are the same as the original data, the geographic extent, however, is limited to continental Australia. To facilitate comparison with surfaces of interpolated gauge data, it was necessary to construct daily TMPA-RT-based estimates of precipitation for Australia using a consistent 24-h accumulation period. That is, the interpretation of satellite-derived daily precipitation is the accumulation of the rain falling in the 24 h to 9am local time for a given day, taking into account the different time zones. An illustration of the process is given in Fig. 2 for 13 April 2003.

For a single pixel in the eastern side of the country, a time series is constructed using the precipitation rates from 2100 UTC on 11 April to 0000 UTC on 13 April 2003. Integrating under the curve between 9 am Eastern Standard Time (EST) on 12 April to 9 am EST on 13 April 2003 gives the daily precipitation (i.e., 50.4 mm) for that pixel. The process is repeated for each time zone (noting that a different set of UTC times may be required), which for standard time in Australia, are UTC + 8, UTC + 9.5 and UTC + 10 h.

3. Blending methods for rain gauge and satellite data

3.1. Data transformation

There are two considerations that we made in our treatment of daily rainfall data across Australia irrespective of and prior to, the

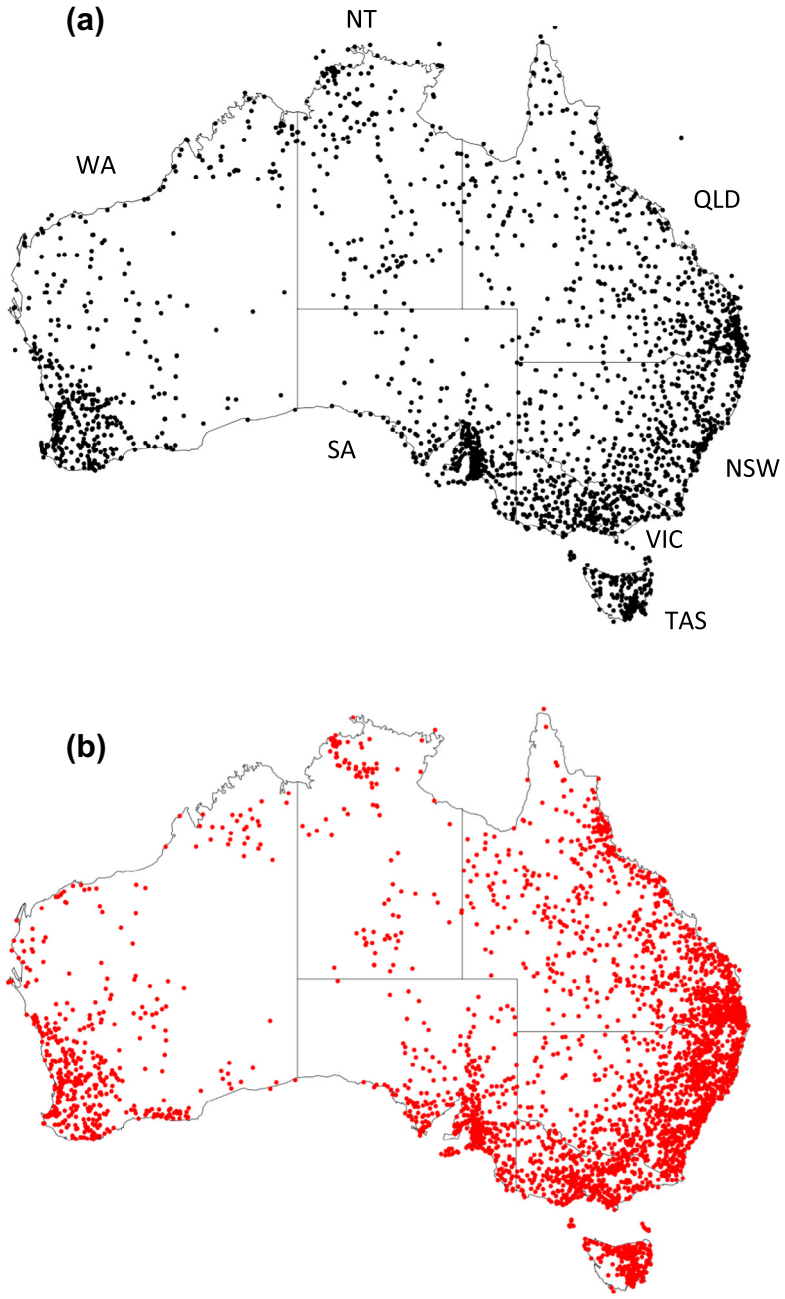

Fig. 1. Typical distribution of rain gauges reporting daily rainfall. Displayed are the gauges that reported rainfall for 1 September 2007 in (a) near real-time (NRT) and (b) in the days to months following, referred to here as post real-time (PRT).

use of a particular method. Daily rainfall comes in discrete blocks in space and time which leads to a large probability mass at zero, no fixed upper bound and a tendency for the probability density functions (PDFs) to be positively skewed (Grimes and Pardo-Igúzquiza, 2010). Highly skewed data can cause problems in the specification of the variogram, so it was transformed to a Normal distribution using the Normal score transform (Deutsch and Journel, 1998). Across continental Australia a global coordinate system such as latitude and longitude is not equal in area. This means that a 1° width near the equator (northern Australia) is considerably larger than the same 1 degree width in the south of Australia. To overcome any geographically dependent preferential inclusion of data in the geostatistical analysis, the coordinate system of the gauge and satellite data was converted (in *R* using spTransform with PROJ.4 projection arguments) to an equal area projection using Universal Transverse Mercator with the Geodetic Datum of Australia 1994 (GDA94).

3.2. Accounting for change of support

The variogram of a set of spatial data is used to estimate the expected squared difference between the locations (x) of two sample values and their respective separation distance. The semi-variance is computed as:

$$\hat{\gamma}(h) = \frac{1}{2m(h)} \sum_{i=1}^{m(h)} \{z(x_i) - z(x_i + h)\}^2, \quad (1)$$

where the $z(x)$ and $z(x + h)$ represent actual values of a property, Z, at places separated by the lag vector h for a set of data $z(x_i)$, $i = 1$,

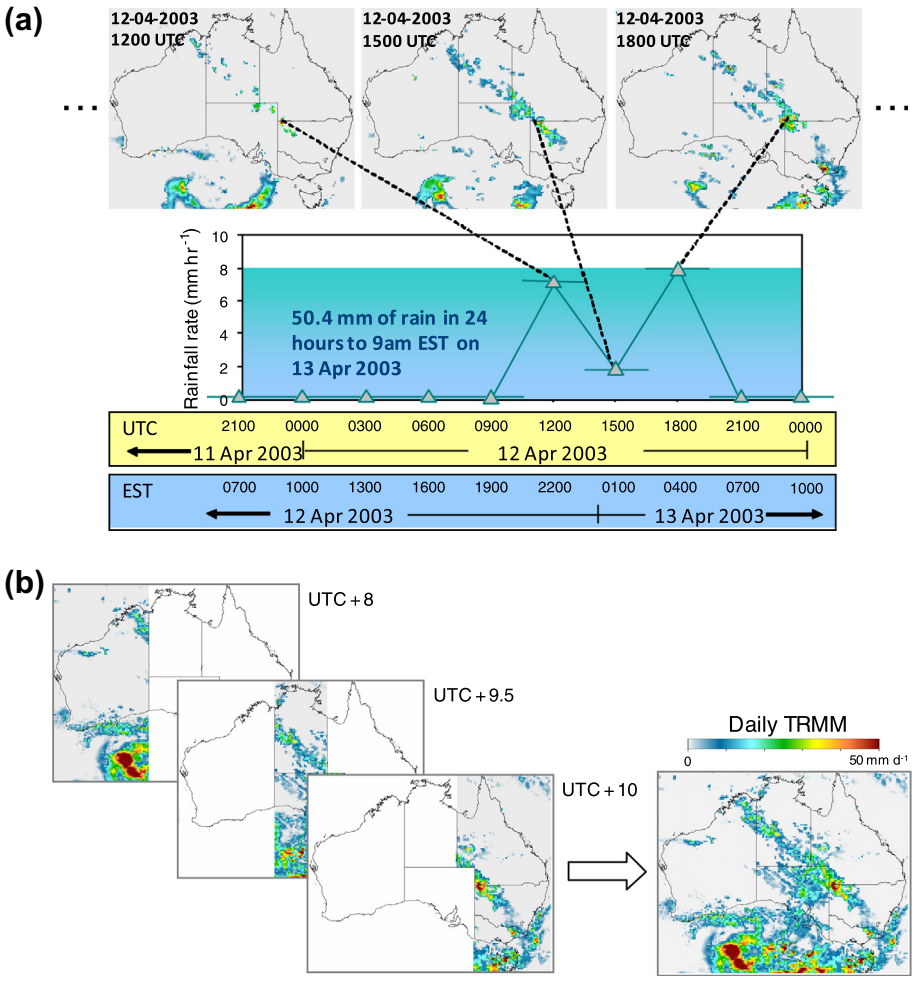

Fig. 2. Graphical depiction of the approach used to construct daily satellite-based precipitation estimates for Australia for 13 April 2003: (a) for a single pixel, using all the precipitation rates in the 24 h to 9 am Eastern Standard Time (EST), and (b) for the whole continent, taking into account the three different time zones.

2, . . ., $m(h)$, where $m(h)$ is the number of pairs of data points. The variogram summarises the spatial variation in a property and describes how that variation changes with increasing separation distance between samples. Once the variogram has been calculated it is fitted with an "authorised" model selected from several families of models (e.g., Gaussian, spherical, exponential) and a nugget term, if appropriate, which ensure validity of the model when making estimates at unsampled locations.

Remote sensing data provides a valuable secondary variable for blending with rain gauge (point) observations. However, the information available from 3B42RT is from a very large support (0.25°, ~25 km square grid) whilst rain gauge observations represent such a small support that they may be considered points. Any blending of these different sources of data should take account of these differences in support. The variogram on one support can be related to that on another. For distances h which are very large in comparison with the dimension of support v, the regularized variogram is obtained from the point variogram by subtracting a constant term $\hat{\gamma}(v, v)$ which is related to the dimensions and geometry of the support v of the regularization (Journel and Huijbregts, 1978, p. 78) i.e., the dispersion variance of the support (Webster and Oliver, 2001, p. 64). In this way, the variogram of the rain gauges can be regularized to the support of the satellite data. Since we also need to calculate the cross-variogram, we took a practical alternative to this approach that has been adopted by others (e.g., Krajewski, 1987; Chappell, 1998). Namely, we identified the NRT rain gauges that fell within satellite pixels and used block-kriging to estimate rainfall over selected blocks that coincided with those satellite pixels. This approach was used to tackle the COSP and these 0.25° square blocks (~25 km square) represented the upscaled NRT gauge data for use with cokriging. For each day of gauge and satellite data, variograms were calculated and fitted with a range of models. The best model was selected using the sum of squares difference between predictions and measurements. This automated process ensured that a bespoke spatial model was available for each day. Selected experimental variograms showed no evidence that spatial variation was dependent on direction (anisotropic). Consequently, we assumed that the variogram was isotropic and fitted a model on that basis for every day of gauge and satellite data.

3.3. Kriging and cokriging

Kriging is the method of geostatistical estimation at unsampled locations (points) or areas (blocks). It is one of the most reliable two-dimensional spatial estimators (Laslett et al., 1987; Laslett and McBratney, 1990; Laslett, 1994) and often produces more reliable estimates than methods of interpolation (Webster and Oliver, 2001). Using ordinary point kriging (PK) we estimated daily rainfall at unsampled locations on a 5 km square grid across Australia. The map provided a baseline output against which to compare the other methods. The 5 km grid spacing used here was set for consistency and subsequent comparison with two extant national gauge-only rainfall products (Jeffrey et al., 2001; Jones et al., 2009).

Table 1
Geostatistical estimation* methods used in this study and their relative conceptual advantages and disadvantages.

Labels	Technique description[a]	Advantages	Disadvantages
PK	Point kriging	Estimates made at discrete points	Points cover small area and map of points prone to erratic (small scale) spatial pattern Does not account for related secondary information to improve estimates Assumes daily rainfall are continuous in space
PcoKup	Upscaled point cokriging	As above for PK Improves estimates using related secondary information Accounts directly for spatial variability of scales of support	Assumes daily rainfall are continuous in space
PKI	Point kriging intermittency	As above for PK Directly accounts for intermittency of rainfall in space	As above for PK (excluding continuity issue)
PcoKIup	Upscaled point cokriging intermittency	As above for PKI As above for PcoKup	

[a] All kriging techniques use ordinary kriging as opposed to simple kriging because of the underlying assumption made that the rainfall process is stationary only within local neighbourhoods and hence the population mean is unknown.

Ancillary data can be used to improve the estimates of the property of interest at unsampled locations using cokriging. It has been used in many applications to improve the accuracy of expensive-to-measure properties (sparsely sampled) using one or more spatially interdependent, cheaper-to-measure properties. This is because the benefits of cokriging are maximised when the secondary data is related to the primary data and the amount of secondary data is considerably larger than the primary data. In this respect, remotely-sensed data provide cheap, intensively sampled spatial information for use with this technique (e.g., Atkinson et al., 1994). In addition to the (auto-) variogram of each variable, cokriging requires a joint model or cross-variogram, defined by analogy with the direct variogram as:

$$\hat{\gamma}_{zs}(h) = \frac{1}{2m(h)} \sum_{i=1}^{m(h)} \{[z(x_i) - z(x_i + h)][s(x_i) - s(x_i + h)]\}^2 \quad (2)$$

,where $s(x)$ is the secondary variable.

Multiple secondary variables attract considerable additional modelling and computational effort which is often not justified by improvements in the estimation variance. The linear model of co-regionalization (LMC) provides a framework for modelling the auto- and cross-variograms of two or more variables so that the variance of any possible linear combination of these variables is always positive (Journel and Huijbregts, 1978). In this case, the model for each of the sample variograms may consist of one or more authorised models. However, the same basic model (with the same range values) must appear in each auto- and cross-variogram (Isaaks and Srivastava, 1989). We used (in *R*) the LMC model for point cokriging the NRT rain gauge and 3B42RT data (PcoKup) so that for each day a bespoke spatial model was available.

Simplified 'cokriging' estimators (e.g., collocated cokriging and kriging with external drift) could not be used here because they would provide estimates at the resolution of the auxiliary data (approx. 25 km square) and the requirement for this work was to make estimates at a finer grid resolution (approx. 5 km square).

3.4. Intermittent rainfall fields

Daily rainfall data, unlike monthly or annual rainfall patterns, is usually intermittent. In other words, there are distinct boundaries between wet and dry areas caused by the localised nature of the rainfall. Barancourt et al. (1992) considered this intermittency of the rainfall fields and proposed a treatment which was recently summarised by Grimes and Pardo-Igúzquiza (2010):

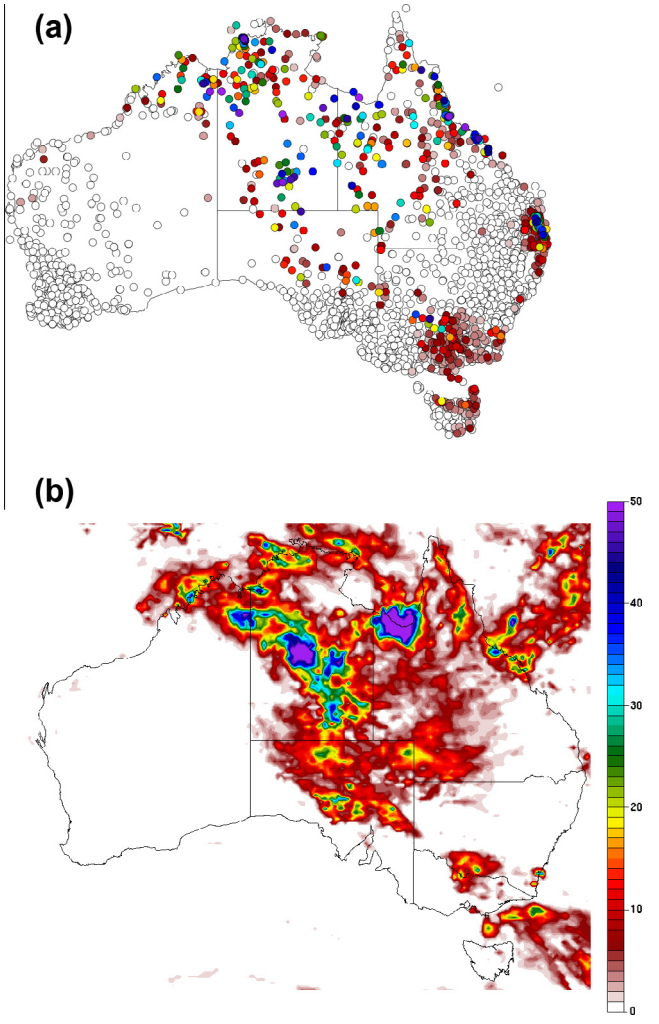

Fig. 3. Maps for 28 February 2010 showing the location and magnitude of daily rainfall measured at (a) rain gauges and (b) estimated by the TMPA 3B42RT-derived daily rainfall product for Australia.

(1) Transform the rainfall values to an indicator (binary) variable where the rainfall value is considered to be zero (0) or nonzero (1).

(2) Calculate the experimental variogram of the indicator data and fit a model in the manner described above (Section 3.1).
(3) Estimate the soft indicator field $I(x)$ on a regular grid by (indicator) (co)kriging values between 0 and 1 which represent the probability of rainfall occurrence.
(4) Produce a hard indicator field $I*(x)$ by selecting a probability threshold T such that $I*(x) = 1$ for $I(x) > T$ and $I*(x) = 0$ for $I(x) \leqslant T$ where T is set to ensure that the proportion of the target domain designated rainy is the same as the numerical proportion of gauges registering nonzero rainfall amounts.
(5) Calculate and fit a model to the variogram of the nonzero rainfall values and estimate the rainfall amount field $F(x)$ using ordinary (co)kriging.
(6) Estimate the final rainfall field $Z(x) = I*(x)F(x)$.

We coded this algorithm (in *R*) and developed a workflow and repeated it for each day of data and therefore had bespoke spatial models to implement this intermittent approach for ordinary point kriging (PKI) for use as a baseline comparison. We also implemented this approach for point cokriging by accounting for the change of support (PcoKIup). The techniques used in this study are summarised in Table 1.

3.5. Evaluation statistics

In addition to making estimates on a regular grid we made estimates using the same kriging techniques at the locations of the post real-time (PRT) rain gauges which are here used as an independent evaluation dataset (Section 2.1). The specific point estimates for each day between 31 May 2009 and 31 May 2010 were then compared with those PRT rain gauge values that were set aside. Several nation-wide error statistics were computed to assess the performance of the methods. The square root of the mean squared error between the estimate and the gauge value (RMSE) provided an average measure of absolute difference. Small RMSE values indicated smaller differences and hence better performance than larger values. The mean absolute error (MAE) is similar but gives less weight to the more extreme differences. Small values indicated the best performance. The bias (Bias) is the average difference amongst the estimates and the gauge values. An estimator is said to be unbiased if its bias is equal to zero for all values. The Pearson Product Moment correlation coefficient (R) describes the relationship between the pairs of estimates and gauge values. The correlation represents the scatter and direction of a linear relationship, but not the slope of that relation, nor many aspects of nonlinear relations.

The satellite rainfall and numerical weather prediction communities compute, in addition to the error statistics above, a suite of categorical statistics to measure the performance of interpolation or estimation techniques to predict rain or no rain. We have used here several of those statistics to assess the performance of the estimation methods. Following Ebert et al. (2007) every 5 km grid point can be classified as a hit (H, observed rain correctly detected), miss (M, observed rain not detected), false alarm (F, rain detected but not observed), or null (no rain observed nor detected) event. The probability of detection, $POD = H/(H + M)$, gives the fraction of rain occurrences that were correctly detected (perfect score is 1), while the false alarm ratio, $FAR = F/(H + F)$, measures the fraction of rain detections that were actually false alarms (perfect score is 0). The equitable threat score (ETS) gives the fraction of observed and/or detected rain but was correctly detected, adjusted for the

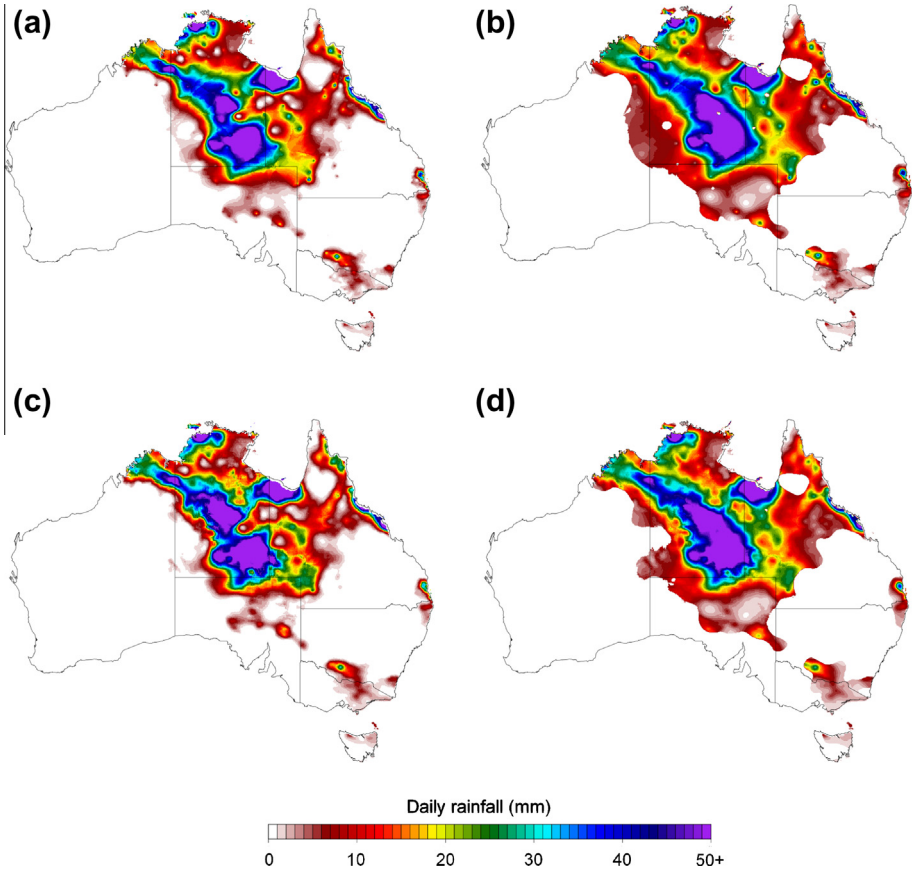

Fig. 4. Maps of daily rainfall (mm) for 28 February 2010 at points on a ~5 km regular grid across Australia assuming rainfall continuity by the point kriging estimators (a) PK with upscaling (c) PcoKup and expecting intermittency by the point kriging estimators (b) IPK with upscaling (d) IPcoKup. Table 1 describes the estimators used.

number of hits *He* that could be expected due purely to random chance, where $He = (H + M)(H + F)/N$ and N is the total number of estimates (perfect score is 1). The ETS is commonly used as an overall skill measure, with the POD, and FAR providing complementary information on bias, misses, and false alarms.

Like the NRT rain gauge data, the PRT rain gauge (validation) data are not spread evenly across Australia. They are located predominantly in the coastal regions where rainfall is likely to be larger than interior regions. These standard measures of performance are *a priori* likely to be biased by the large number of data in these coastal regions and have the potential to represent the performance of the techniques mainly around the coastal region of Australia. To tackle this issue we used cell-declustering (Deutsch, 1989) of each day's gauge data to determine the optimal size of a grid in which weights were applied to calculate an unbiased mean. In each optimal-sized grid cell weights were applied to all gauges. The standard statistics described above were recalculated on these cell-declustered or preferentially weighted data. For each day, several cell sizes and origins were tried to identify the smallest declustered mean because large values of rainfall were preferentially sampled. Erratic results caused by extreme values falling into specific cells were avoided by averaging results for several different grid origins for each cell size. Ultimately, the cell size which produced the smallest declustered mean for Australia was adopted for each day and used to select the subset of PRT rain gauge values.

4. Results

We describe the results of the study in two ways, visually for one day (Section 4.1) and quantitatively using evaluation statistics (Section 4.2).

4.1. Spatial variation of rainfall

We chose judiciously one day (28 February, 2010) to make a visual comparison of the techniques. This comparison enabled a first-order difference between the techniques to be identified readily. The maps also acted as an example when considering the time series statistics presented below. The context for the blend of techniques is provided by the map of rain gauges for this particular day and the available 3B42RT data (Fig. 3).

The gauges confirm that for this particular day a large proportion of Australia is dry but that large areas of central and northern Australia have recorded rainfall and small areas of rainfall have also occurred in eastern Queensland (QLD), around the border of Victoria (VIC) and New South Wales (NSW) and in northern Tasmania (TAS). There are noteworthy differences between the locations of rainfall measured by the gauges and the estimates of rainfall in the TMPA-RT product. For example, for some parts of central Australia there are no gauges where the TMPA-RT product shows a large amount of rain. Conversely, some gauges show rainfall (in Tasmania and the eastern border of Victoria and NSW) where there is almost no rainfall estimated by the TMPA-RT product.

There appear to be broadly similar patterns of rainfall produced by each of the kriging techniques (Fig. 4). Notwithstanding those similarities, there is a difference between those techniques which assume continuity in the rainfall (Fig. 4a and c) and those expecting intermittency (Fig. 4b and d). The latter appear to have smoothly varying rainfall which extends beyond the boundary of the other maps but which is truncated at the fringes of the wet areas. This is particularly evident at the southerly extent of the main wet area (in South Australia). In contrast, the former maps (Fig. 4a and c) are more variable as evident by the size and fragmented nature of the different coloured zones. These maps also display rainfall diminishing rapidly towards the fringe of the wet areas. The maps produced using point kriging without the TMPA-RT data (Fig. 4a and b) are very similar to those maps produced using cokriging to combine the 3B42RT data (Fig. 4c and d). It appears that the 3B42RT data has made little visual difference to the rainfall maps.

The estimation variance maps associated with the continuous point kriging techniques are shown in Fig. 5 using the same colour scale. These maps provide a measure of the technique's estimation uncertainty based largely on the configuration of the gauge data and whether satellite data was included in each technique. The kriging variance is independent of the magnitude of the estimate. A better estimate of uncertainty is obtained through stochastic simulation methods (Deutsch and Journel, 1998). The influence of the rain gauge configuration is evident from the patches of small variance (Fig. 5a). The magnitude of the estimation variance is considerably reduced with the inclusion of the 3B42RT data (Fig. 5b). The estimation variance of block kriging (not shown) displayed a further reduction in the estimation variance of these techniques.

4.2. Temporal variation of the evaluation statistics

The daily time series of the evaluation statistics (RMSE, MAE, bias and R) for each of the kriging techniques are shown in Fig. 6. Each plot shows that there is very little difference amongst these techniques when their estimates are compared to the PRT rain gauge values that were not used to produce the maps.

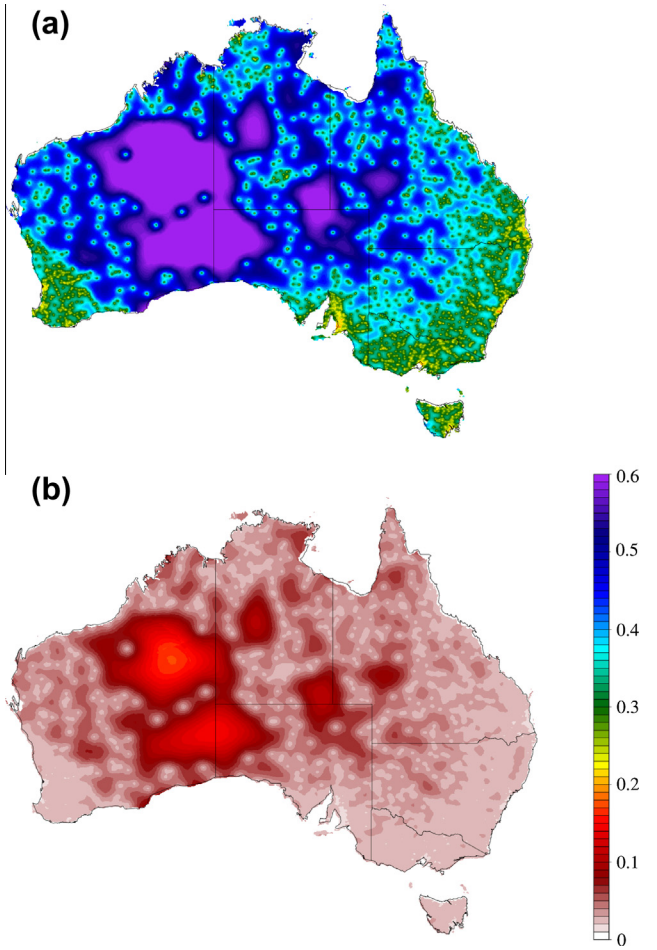

Fig. 5. Estimation variance maps of daily rainfall (mm) for 28 February 2010 produced by point kriging estimators (a) PK with upscaling (b) PcoKup at points on a ~5 km regular grid across Australia. Table 1 describes the estimators used.

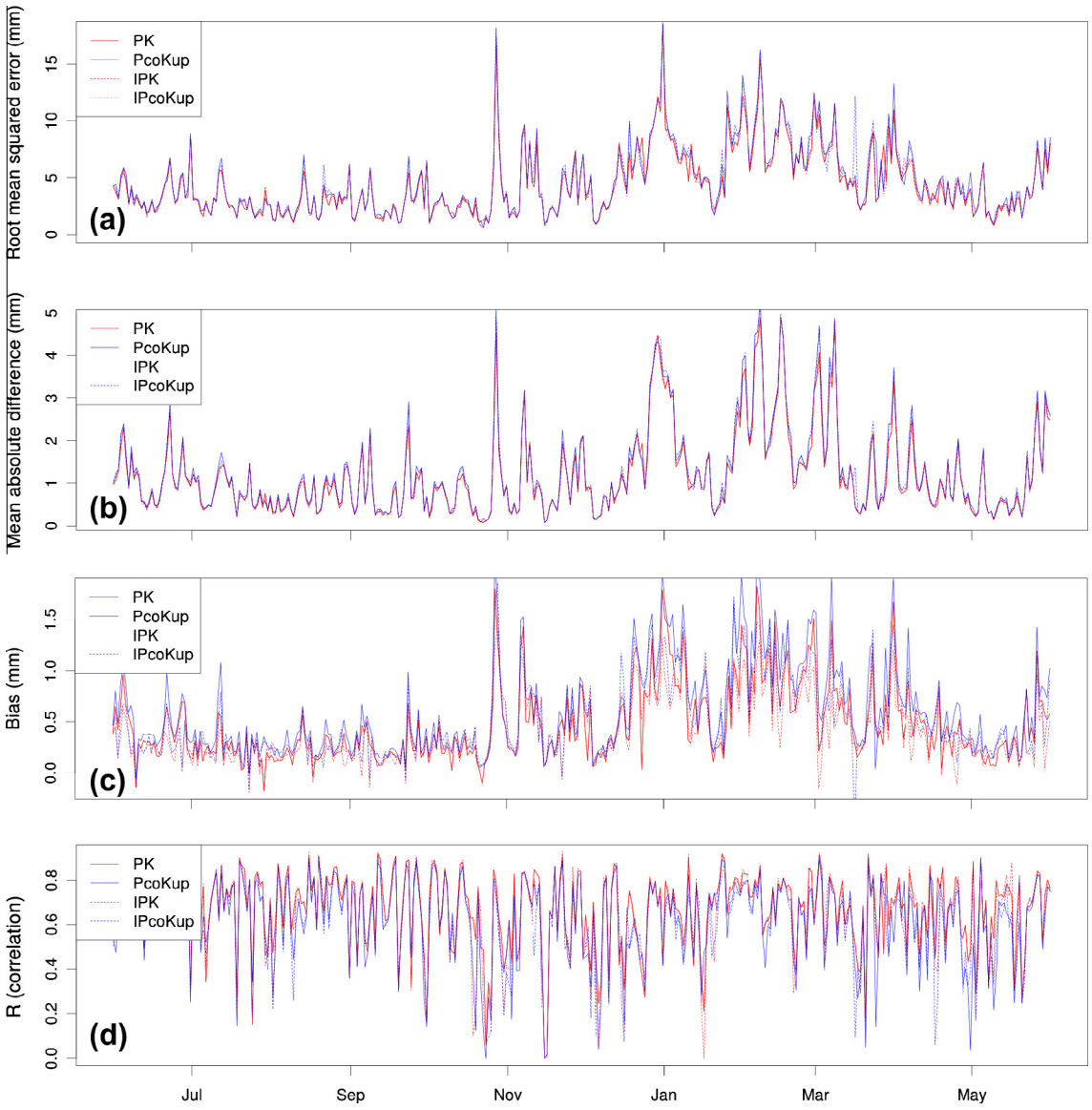

Fig. 6. Daily variation in the evaluation statistics (see Section 3.5) (a) square root of the mean squared error, (b) mean absolute difference, (c) bias and (d) Pearson product moment correlation of the kriging techniques estimates and the independent dataset of rain gauge values. Table 1 describes the estimators used.

Notably, all of the techniques show a decrease in performance (indicated by the RMSE, MAE and Bias) between January to April, the southern hemisphere summer. During this season, rainfall over a large part of the country is dominated by large magnitude and small spatially cohesive convective systems. To summarise this time series of information we have calculated the mean and the variance of the statistics for the entire time period (Table 2). The techniques of ordinary point kriging intermittency (PKI) and ordinary point kriging (PK) perform best. These results imply that the additional computational effort of cokriging and correctly handling the change of support problem (COSP) cannot be justified for the study period. Furthermore, the relative complexity of ordinary point kriging with intermittency does not appear warranted in comparison to straightforward ordinary point kriging.

The results were also cell-declustered prior to the calculation of the same evaluation statistics. The summary statistics were calculated as before. They showed no appreciable difference in the rank order of the techniques (not shown). The magnitude of the RMSE and MAE summary mean statistics was reduced slightly and the bias in the statistics also reduced.

Table 2
Annual evaluation summary statistics[a] of the kriging techniques (ranked by RMSE). Table 1 describes the spatial estimators used.

Estimators	RMSE	MAE	Bias	*R*	POD	FAR	ETS
PK	4.42	1.20	0.47	0.67	0.69	0.23	0.47
PcoKup	4.70	1.28	0.60	0.63	0.68	0.25	0.45
PKI	4.41	1.20	0.40	0.66	0.56	0.15	0.41
PcoKIup	4.66	1.27	0.50	0.64	0.54	0.13	0.39

[a] Section 3.5 contains a description of these statistics and the datasets used Shaded cells indicate the best performing method for each statistic.

The most obvious difference in performance between the two groups of techniques (i.e., the ordinary and intermittent kriging techniques) is observed in the categorical statistics (Table 2). Here, the FAR is smallest for the intermittent kriging techniques. This is not too surprising given that a first step in intermittent (co)kriging is the construction of a rain-no rain mask from the NRT gauges. Recall from Fig. 1 that most PRT gauges are distributed amongst the NRT gauges, therefore rain-no rain areas will be largely consistent between the two data sets. This results in fewer false alarms (i.e.,

estimation of rain when no rain was observed) and represented in the FAR. The rain-no rain masking of rain areas, however, is likely to result in proportionally fewer hits (i.e., estimates of rain when rain was observed) even if the satellite data are indicating rain outside the prescribed rain area. The result is smaller POD values for the intermittent kriging compared to the ordinary kriging approaches. Finally, the ETS suggests that overall, the intermittent kriging techniques perform slightly worse than the other techniques in terms of their rainfall detection performance, likely due to the masking of rain-no rain areas based on NRT gauges alone.

5. Discussion

The intermittent technique for the estimation of rainfall (Barancourt et al., 1992) offers an appropriate method of estimating daily rainfall because it explicitly tackles the occurrence of discrete patches in space and time for daily rainfall fields. Visually, there are differences between the maps produced with and without this method. However, the general patterns are the same in all of the maps indicating that the differences are at the small scale. There is also a strong basis for explicitly dealing with the change of support problem (COSP) in the blending of the satellite 3B42RT and rain gauge data. However, the inclusion of the satellite data using cokriging makes little visual impression on the rainfall map estimates. The use of the satellite data does not change the general, large scale, pattern.

For each day's NRT rain gauge data used to make the estimates across Australia, there are approximately twice as many PRT rain gauge validation data. The validation data are used in the evaluation of the performance of the techniques. Those statistics quantify the visual differences as being small and consistent over the year (Fig. 6). The results show that all techniques are more similar to each other than they are to the validation data. Despite considerable effort to identify the source of bias we were unable to reach a satisfactory conclusion. We know the PRT and NRT samples are biased in the sense that they represent the wetter parts of the continent than the drier parts of the continent. However, declustering the results prior to calculating the statistics made no difference to the interpretation.

The results indicate that there appears to be no clear benefits of including the satellite data to estimate daily rainfall across Australia, either with or without the intermittent method. Although these findings support the results of Krajewski (1987), other studies have shown improved rainfall estimates by using satellite (e.g., Grimes et al., 1999) and radar (e.g., Velasco-Forero et al., 2009) data. This apparent contradiction between our results and the literature may be due to the use of a single spatial model for the entire continent for each day. In other words, there may be mixing between different climate regimes in different parts of Australia. If this is happening then it is compounded by the accepted practice of global (continental scale) statistical assessment of performance. However, an investigation of the regional variation and the use of a pooled method of estimating the variogram within and between regions (e.g., pooled within-class variogram) was beyond the scope of this study.

The apparent contradiction led us to investigate the effect of rain gauge number and density on the blending of satellite with rain gauge data (Renzullo et al., 2011). Results showed that with more than 1000 evenly spaced gauges there is marginal benefit of including the satellite rainfall retrievals in the interpolation of gauge observations (see online supplementary material). This appears to be because the satellite 3B42RT data improves upon gauge-only analyses in (predominantly interior) regions of Australia where gauge densities are less than 4 gauges per 10,000 km^2. Bias due to the large density of gauge observations around the coastal region is likely contained partly within the satellite data and partly the validation data because of the tendency for gauges to be correlated with one another in densely sampled areas. Consequently, blending the satellite data with gauge data will, relative to the gauge-based validation, appear as bias. In sparsely sampled areas, the gauge spacing likely exceeds their correlation distance, the satellite data becomes more useful, and the bias is reduced.

There are typically more than 1000 rain gauges available for near real-time estimation of daily rainfall across Australia. Consequently, there is little benefit of using the satellite data to make estimates at unsampled locations. However, when the change of support problem (COSP) is tackled and the satellite data are combined using cokriging, we observe a considerable difference in the estimation variance. The estimation variance is reduced to approximately an order of magnitude smaller (Fig. 5) than the point kriging of rain gauges alone. Evidently, the use of satellite data makes a considerable difference to the estimation variance. These results suggest that there are considerable benefits to appropriately blending the satellite and rain gauge data. However, those benefits are not evident in the traditional performance assessment methods (Table 2). This is likely caused by those metrics indicating the accuracy of the estimation technique and not the amount of uncertainty in that estimate. In this case, uncertainty is required to demonstrate detectable difference in pattern/trend. Unfortunately, the issue of uncertainty is beyond the scope of this paper but is tackled elsewhere using conditional simulation of blended rainfall estimates (Chappell et al., 2012).

6. Conclusions

Near real-time (NRT) daily rainfall was estimated across Australia using the network of gauges reporting within 24 h of 9 am on the day of interest. Observations returned outside this timeframe were used to evaluate the performance of the estimates. Accounting for discrete patches of rainfall in space and time was a valuable technique. It performed best on average during the period of this study. However, its additional complexity hardly warranted the small improvement in performance over ordinary point kriging. Estimates of daily rainfall made by combining 3B42RT satellite and rain gauge data were on average, for the study period, worse than ordinary point kriging. This was probably because the density of the gauges, particularly in the coastal regions of Australia, was sufficient to provide accurate results without the satellite data. The satellite data will improve upon gauge-only analyses in (predominantly interior) regions of the Australia where gauge densities are less than 4 gauges per 10,000 km^2. However, the assessment of the benefits of including satellite data to the estimates of daily rainfall is confounded by the use of traditional evaluation statistics. For example, the kriging estimation variance of blended satellite and gauge data reduced by an order of magnitude compared to kriging gauge data alone. The estimation is a rudimentary estimate of uncertainty since it is largely dependent on the configuration of the sample data. Detection of difference is dependent on several additional sources of error which when combined would provide a more complete assessment of uncertainty. Nevertheless, the reduced estimation variance indicates an improved ability to detect change in daily rainfall with potential for trend detection over longer time periods. The more holistic assessment of (spatial) uncertainty is beyond the scope of this paper but is considered elsewhere (Chappell and Agnew, 2008; Chappell et al., 2012). This study shows that the evaluation of performance based on local estimates relative to an independent dataset is only part of the assessment. The estimation variance indicates the role of uncertainty in providing a more realistic assessment of a technique's performance within the context of detecting change.

Acknowledgements

This work was supported under the Water Information Research and Development Alliance (WIRADA) between the Bureau of Meteorology and CSIRO Water for a Healthy Country flagship. We are grateful for the support of Dr. Andrew Frost in arranging access to the daily gauge data from the Bureau of Meteorology. We are grateful to the anonymous reviewers whose comments did much to improve the manuscript. Any omissions or errors in the manuscript remain the responsibility of the authors.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jhydrol.2013.04.024.

References

Atkinson, P.M., Webster, R., Curran, P.J., 1994. Cokriging with airborne MSS imagery. Remote Sens. Environ. 50, 335–345.
Barancourt, C., Creutin, J.D., Rivoirard, J., 1992. A method for delineating and estimating rainfall fields. Water Resour. Res. 28 (4), 1133–1144.
Chappell, A., 1998. Using remote sensing and geostatistics to map ^{137}Cs-derived net soil flux in south-west Niger. J. Arid Environ. 39, 441–455, http://dx.doi.org/10.1006/jare.1997.0365.
Chappell, A., Agnew, C.T., 2008. How certain is desiccation in west African Sahel rainfall (1930–1990)? J. Geophys. Res. 113, D07111, http://dx.doi.org/10.1029/2007JD009233.
Chappell, A., Haylock, M., Renzullo, L., 2012. Spatial uncertainty to determine reliable daily precipitation maps.. J. Geophys. Res. 117, D17115, http://dx.doi.org/10.1029/2012JD017718.
Clarke, R.T., Buarque, D.C., De Paiva, R.C.D., Collischonn, W., 2011. Issues of spatial correlation arising from the use of TRMM rainfall estimates in the Brazilian Amazon. Water Resour. Res. 47 (5), W05539.
Creutin, J.C., Obled, C., 1982. Objective analysis and mapping techniques for rainfall fields: an objective comparison. Water Resour. Res. 18 (2), 413–431.
Deutsch, C.V., 1989. DECLUS: a fortran 77 program for determining optimum spatial declustering weights. Comput. Geosci. 15 (3), 325–332.
Deutsch, C.V., Journel, A.G., 1998. GSLIB: Geostatistical Software Library and User's Guide. Oxford University Press, Oxford, pp. 369.
Ebert, E.E., Janowiak, J.E., Kidd, C., 2007. Comparison of near-real-time precipitation estimates from satellite observations and numerical models. Bull. Am. Meteorol. Soc. 88, 47–64. http://dx.doi.org/10.1175/BAMS-88-1-47.
Goovaerts, P., 2000. Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. J. Hydrol. 228, 113–129.
Gotway, C.A., Young, L.J., 2002. Combining incompatible spatial data. J. Am. Stat. Assoc. 97 (458), 632–648.
Grimes, D.I.F., Pardo-Igúzquiza, E., 2010. Geostatistical analysis of rainfall. Geograph. Anal. 42, 136–160.
Grimes, D.I.F., Pardo-Igúzquiza, E., Bonifacio, R., 1999. Optimal areal rainfall estimation using raingauges and satellite date. J. Hydrol. 222, 93–108.
Hofstra, N., Haylock, M., New, M., Jones, P., Frei, C., 2008. The comparison of six methods for the interpolation of daily European climate data. J. Geophys. Res. 113, D21110. http://dx.doi.org/10.1029/2008JD010100.
Huffman, G.F., Adler, R.F., Arkin, P., Chang, A., Ferraro, R., Gruber, A., Janowiak, J., McNab, A., Rudolf, B., Schneider, U., 1997. The global precipitation climatology project (GPCP) combined precipitation dataset. Bull. Am. Meteorol. Soc. 78 (1), 5–20.
Huffman, G.F., Adler, R.F., Bolvin, D.T., Gu, G., Nelkin, E.J., Bowman, K.P., Hong, Y., Stocker, E.F., Wolff, D.B., 2007. The TRMM multisatellite precipitation analysis (TMPA): quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. J. Hydrometeorol. 8, 38–55.
IPWG, 2011. International Precipitation Working Group, <http://www.isac.cnr.it/~ipwg/validation.html>.
Isaaks, E.H., Srivastava, R.M., 1989. An Introduction to Applied Geostatistics. OUP, New York.
Jeffrey, S.J., Carter, J.O., Moodie, K.B., Beswick, A.R., 2001. Using spatial interpolation to construct a comprehensive archive of Australian climate data. Environ. Model. Software 66, 309–330.
Jones, D.A., Wang, W., Fawcett, R., 2009. High-quality spatial climate data-sets for Australia. Aust. Meteorol. Oceanogr. J. 58 (4), 233–248.
Journel, A.G., Huijbregts, C.J., 1978. Mining Geostatistics. Academic Press, London.
Joyce, R.J., Janowiak, J.E., Arkin, P.A., Xie, P., 2004. CMORPH: a method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. J. Hydrometeorol. 5, 487–503.
Krajewski, W.F., 1987. Cokriging radar-rainfall and rain gauge data. J. Geophys. Res. 92, 9571–9580.
Kubota, T., Shige, S., Hashizume, H., Aonashi, K., Takahashi, N., Seto, S., et al., 2007. Global precipitation map using satelliteborne microwave radiometers by the GSMaP Project: production and validation. IEEE Trans. Geosci. Remote Sens. 45 (7), 2259–2275.
Kummerow, C., Simpson, J., Thiele, O., Barnes, W., Chang, A.T.C., Stocker, E., Adler, R.A., Hou, A., Kakar, R., Wentz, F., Ashcroft, P., Kozu, T., Hong, Y., Okamoto, K., Iguchi, T., Kuroiwa, H., Im, E., Haddad, Z., Huffman, G., Ferrier, B., Olson, W.S., Zipser, E., Smith, E.A., Wilheit, T.T., North, G., Krishnamurti, T., Nakamura, K., 2000. The status of the tropical rainfall measuring mission (TRMM) after two years in orbit. J. Appl. Meteorol. 39, 1965–1982.
Kyriakidis, P., 2004. A geostatistical framework for area to point interpolation. Geograph. Anl. 36 (3), 259–289.
Laslett, G.M., 1994. Kriging and splines: an empirical comparison of their predictive performance in some applications. J. Am. Stat. Assoc. 89, 391–409.
Laslett, G.M., McBratney, A.B., 1990. Estimation and implications of instrumental drift, random measurement error and nugget variance of soil attributes – a case study for soil pH. J. Soil Sci. 41, 451–471.
Laslett, G.M., McBratney, A.B., Pahl, P.J., Hutchinson, M.F., 1987. Comparison of several spatial prediction methods for soil pH. J. Soil Sci. 38, 325–341.
Lebel, T., Bastin, G., Obled, C., Creutin, J.D., 1987. On the accuracy of areal rainfall estimation: a case study. Water Resour. Res. 23, 2123–2134.
Li, M., Shao, Q., 2010. An improved statistical approach to merge satellite rainfall estimates and raingauge data. J. Hydrol. 385, 51–64. http://dx.doi.org/10.1016/j.jhydrol.2010.01.023.
Renzullo, L.J., Chappell, A., Raupach, T., Dyce, P., Ming, L., Shao, Q., 2011. An assessment of statistically blended satellite-gauge precipitation data for daily rainfall analysis in Australia. In: International Remote Sensing of Environment Conference Proceedings, Sydney, Australia.
Tian, Y., Peters-Lidard, C.D., 2011. A global map of uncertainty in satellite-based precipitation measurements. Geophys. Res. Lett. 37, 6. http://dx.doi.org/10.1029/2010GL046008, L2440.
Velasco-Forero, C., Sempere-Torres, D., Cassiraga, E.F., Gomez-Hernandez, J.J., 2009. A non-parametric automatic blending methodology to estimate rainfall fields from rain gauge and radar data. Adv. Water Resour. 32, 986–1002.
Vila, D.A., de Goncalves, L.G.G., Toll, D.L., Rozante, J.R., 2009. Statistical evaluation of combined daily gauge observations and rainfall satellite estimates over continental South America. J. Hydrometeorol. 10, 533–543.
Webster, R., Oliver, M.A., 2001. Geostatistics for Environmental Scientists. Wiley, Chichester, UK.