Blending radar, NWP and satellite data for real-time ensemble rainfall analysis: a scale-dependent method

Luigi J. Renzullo 1,*, Carlos A. Velasco-Forero 2, and Alan W. Seed 2

1 CSIRO Land and Water, Canberra, ACT, Australia
2 Bureau of Meteorology, Melbourne, Vic., Australia
* Luigi.Renzullo@csiro.au; +61 2 6246 5785
## Contents

**Contents**

i

**Figures** iii

**Tables** v

**Acknowledgments** vi

**Executive summary** vii

1 **Introduction** 8

   The issue ........................................................................................................... 8

   Real-time applications ..................................................................................... 8

   Delayed mode and near real-time applications .............................................. 8

   Research motivation ....................................................................................... 9

   Study objectives ............................................................................................. 10

2 **Data and method** 11

   2.1 Rainfall data sources .................................................................................. 11

   2.2 A multifractal blending method: adaptation of STEPS ............................... 13

       2.2.1 Spatial scale decomposition into cascade components ....................... 15

       2.2.2 Scale dependent blending of rainfall data ............................................ 18

       2.2.3 Calculating the blending weights ......................................................... 18

       2.2.4 Rainfall ensemble generation .............................................................. 21

       2.2.5 Blending method summary ................................................................. 24

3 **Results** 25

   3.1 Selection of NWP forecasts and satellite products .................................... 25

   3.2 Spatial structure and power spectral density ............................................. 27

   3.3 Spectral decomposition into cascade components ..................................... 29

   3.4 Blending weights ....................................................................................... 32

   3.5 Blended rainfall analyses ........................................................................... 34

Blending radar, NWP and satellite data for real-time ensemble rainfall analysis: a scale-dependent method | i
3.6 Evaluation of blending results

4 Discussion

4.1 Source real-time rainfall data

4.2 Spatial structure and blending weights

4.3 Non-stationarity and ensemble generation

4.4 Evaluation

5 Summary and recommendations

5.1 Summary

Data

Structure

Weights

Evaluation

5.2 Next steps

6 References
Blending radar, NWP and satellite data for real-time ensemble rainfall analysis: a scale-dependent method

Figures

Figure 1. Gridded rainfall data for 1400 hrs UTC on 2 September 2016. Highlighted are Study Areas 1 and 2 in eastern Australia, corresponding to a 512 x 512 km² and 1600 x 1600 km² areas in eastern Australia. Rainfall estimates are: (a) Rainfields v3.0, showing the distribution and coverage of radar data for Australia (grey indicates no data areas); (b) Rainfields estimate for Study Area 1; (c) ACCESS-R forecast (T + 8 hours) for Study Area 2; and (d) GPM IMERG 3B-HHR-L for Study Area 2.

Figure 2. Rainfields estimates of rainfall over Study Area 1 (Fig. 1) for 2000 hrs UTC on 20 September 2016: (a) rain rate R in mm hr⁻¹; (b) transformed into dBZ units, i.e. log Z, using Eq. (1) and pixel frequency histograms for R and log Z in (c) and (d), respectively.

Figure 3. Power spectral density (PSD) function for Rainfields rainfall estimate in Fig. 2: (a) 2D PSD; and (b) isotropic (ISO) PSD derived through radial averaging of the 2D PSD (subfigure a) for wavenumbers from 0.25 km⁻¹ (4 km the Nyquist critical frequency) to 0.002 km⁻¹ (i.e. 512 km). Inset: Power law, distribution as a function of wavenumber, $\omega$.

Figure 4. Decomposition of 1400 hr UTC 2 September 2016 Rainfields estimate for Study Area 1. Spectral components for cascade levels: (a) >512 km; (b) 256 – 512 km; (c) 128 – 256 km; (d) 64 – 128 km; (e) 32 – 64 km; (f) 16 – 32 km; (g) 8 – 16 km; (h) 4 – 8 km; and (i) < 4 km.

Figure 5. Process of generating random noise with the spatial structure: (a) two dimensional power spectrum of the Rainfields radar estimate for 1400 hr UTC on 2 September 2016 (inset); (b) 2D power spectrum of spatially independent zero-mean unit-variance random numbers (white noise, in inset); (c) isotropic power spectral density of radar (grey), white noise (blue lines x 2 realisations), and noise transformed to have spatial structure of radar data; (d) standardised cascade level 4 (32-64 km scale); (e) structure noise for cascade level 5 (16-32 km scale); and (f) structured noise created by summing noise for each cascade level 0-8 in standardised dBZ units.

Figure 6. Isotropic power spectral density curves for radar (Rainfields), NWP (ACCESS-R) and satellite (GPM-IMERG) hourly rainfall products for 1400 hrs UTC on 2 September 2016 over (a) 512 x 512 km² area (Study Area 1) and (b) 1024 x 1024 km study area. Spectral slopes $\beta$ (± standard error) correspond to the 25 – 125 km spatial scale highlighted in grey.

Figure 7. Time series of spectral slope, $\beta$, values for the 25 – 125 km scale for Rainfields (Radar), ACCESS-R (NWP) and GPM IMERG (Satellite) data over September 2016 time period: (a) spectral slopes derived for 512 x 512 km² study area for the three rainfall estimates; (b) and (c) spectral slope time series for the Study Area 1 (solid lines) and Study Areas 2 for ACCESS-R and GPM IMERG rainfall, respectively.

Figure 8. Three sources of rainfall data (in dBZ) over the 512 x 512 km² study area at 1000 hrs UTC on 14 September 2016: (a) Rainfields v3.0; (b) GPM IMERG 3B-HHR ‘late’ and (c) ACCESS_R forecast.
(T+10 hrs). In (d) isotropic power spectral density for the three data sets showing the extent of each cascade level. ................................................................. 31

Figure 9. Correlation of the cascade component for each rainfall data set against the ‘truth’ (in Eq. 8) for each level (Study Area 2). Value calculated for every hour between 0100 hrs UTC 1 September – 2300 hrs UTC 30 September 2016. Shaded area represents the interquartile range of the values and the solid line represents the median value over the time period. ................................................................. 32

Figure 10. The frequency with which the rainfall data is given highest weight amongst the sources (radar, NWP and satellite R, N and S respectively) over the study period for each cascade level (defined in Table 2) for study areas (a) 1 and (b) 2. Highlighted are the cascade levels spanning the 50 – 400 km spatial scales for both study areas. ............................................................................................................ 33

Figure 11. Hourly accumulated rainfall (in dBZ) for 2000 hrs UTC on 29 September 2016. The source data: (a) Rainfields, (b) ACCESS-R T+8 hrs, and (c) GPM IMERG_L; and (d) the blended result corresponding to the ensemble mean.......................................................................................... 35

Figure 12. Rainfall analyses for 0600 hrs UTC 14 September 2016. Source products: (a) Rainfields, (b) ACCESS-R T+12 hrs, and (c) GPM IMERG-L. Blended analyses: (d) radar-NWP; (e) radar-satellite; and (f) radar-NWP-satellite.................................................................................................................. 36

Figure 13. Blended radar-NWP rainfall analysis for 1400 hrs UTC on 2 September 2016, Study Area 1: (a) and (b) two members from the 50 member ensemble; (c) ensemble mean; and (d) isotropic PSD for the source NWP (blue) and radar (black), along with the PSD of the 50 members of the ensemble (grey). Original Rainfields (e) and ACCESS-R (f) rainfall estimates included for comparison................................. 38

Figure 14. Rainfall analysis (mm) for 2000 hrs UTC 20 September 2016 derived from the 50-member ensemble of blending Rainfields and ACCESS-R T+8 hrs. Displayed are the ensemble (a) mean and (b) standard deviation................................................................................................................. 39

Figure 15. Location of the rain gauges (blue dots) reporting 24 accumulated rainfall in relation to the radar coverage (white areas) and Study Area 1. Note that the gauge data are independent of the derived Rainfields, ACCESS-R and GPM rainfall products. The distribution of gauges is typical for any given day. In this figure there are a total 1,070 gauges over the 1600 x 1600 km² area: 220 inside 512 x 512 km² area (black square); 905 inside radar coverage; and 105 outside of radar coverage (grey areas). .......... 40
Tables

Table 1. Summary of error statistics for satellite (GPM IMERG) and ACCESS-R rainfall estimates over Australia between 1-30 September 2016. Evaluation used the Rainfields v3.0 data mosaic for the continent. ................................................................................................................................................................. 26

Table 2. Cascade level definition for Study Areas 1 (512 x 512 km$^2$) and 2 (1600 x 1600 km$^2$). ........... 30

Table 3. Comparison of source and blended rainfall analyses (based on ensemble mean) with 24-hour accumulated rainfall from gauge-based analysis extracted for individual gauge locations during September 2016. Statistics calculated are difference (product – observation) in areal average (DAA) and mean absolute error (MAE). Highlighted are the best statistic in each column. ......................................................... 40
Acknowledgments

This work was conducted as part of the Water Information Research and Development Alliance (WIRADA) between the Bureau of Meteorology and CSIRO.
Executive summary

This report is the second in a series on the development path towards a national-scale real-time rainfall analysis system for Australia. The document provides a description of a method suitable for real-time and near real-time rainfall analysis and ensemble generation. A preliminary evaluation of the method for a study area and period is also presented.

The method developed uses components of the short-term ensemble prediction system (STEPS; Seed et al. 2013) adapted for rainfall analysis (as opposed to forecasting). Specifically the method employs a multiplicative cascade representation of rainfall, blends multiple sources of rainfall data at the cascade level, derives weights for the blending based on mutual correlation between data at the cascade level, and computes ensembles of rainfall analyses in which each member possess the expected scaling properties of rainfall across a wide range of spatial scales. We call the method scale-dependent blending.

We developed, applied and tested the scale-dependent blending method for three sources of real-time and near real-time gridded estimates of hourly accumulated rainfall from (a) Rainfields v3.0 mosaic of radar rainfall (radar), (b) ACCESS-R forecasts of rainfall accumulations (NWP) and (c) merged multi-satellite rainfall product from the GPM IMERG system (satellite). The study focused on a 1600 x 1600 km² area of southeast Australia and the month of September 2016.

Our investigation found that:

- all three of data sets examined were very good at capturing rainfall spatial structure over a range of scales, with radar and satellite most similar in their ability, while NWP consistently showed less structure in the range 25 - 125 km scale;
- the weights derived from the correlation of cascade components for each data related strongly to the ability of the data to represent structure at the corresponding spatial scale; and
- the preliminary evaluation against independent data showed that the blended results often compensated for poor performance of individual data, and were almost always better than satellite data alone in areas of the country without radar coverage.

Recommended future work focusses primarily on more thorough evaluation of the method, over different parts of the country (e.g. Snowy Mountains), and detailed examination of the ensembles as a true representation of rainfall uncertainty. Specific items include:

(Immediate tasks in order of priority)

- Assessing the impact of a revised structured noise perturbation scheme on blended ensemble means, especially in areas of transition from radar coverage to no radar coverage.
- Extending the verification statistics to include metrics on detection (categorical statistics) and skill scores on related to how well the ensembles represent uncertainty.
- Conducting localised verification targeting (i) coast zones with high levels of gauge coverage; (ii) in-land areas with poor or no radar coverage; and (iii) mountainous areas that pose challenge to radar and satellite retrievals.
- Experimenting with temporally static weights versus those derived instantaneously, or dynamically, to quantify the expected difference in performance in real-time and near real-time settings.

(Further development, especially ahead of national-scale implementation)

- Exploring impact of non-stationarity on method and consider the approaches of concatenating a series of localised analyses, or using spatially varying weights in the blending procedure.
- Incorporating the result of independent evaluation of the source data though either an a priori bias correction or guidance on the appropriate use choice of data to serve as the reference of each cascade level.
1 Introduction

The issue

The benefits to the Australian community of accurate, frequently updated (~ hourly) gridded observations of accumulated rainfall with national coverage are numerous. A real-time assessment of the spatial distribution and intensity of rainfall events is of critical importance for timely public announcements, as well as for the coordinated response of emergency services to severe weather events such as heavy storms, flash and riverine flooding. While there are multiple observing networks that provide information on precipitation (e.g. rain gauge, weather RADAR, satellite) no single data source currently possesses the required accuracy, spatial resolution, spatial coverage and data latency for such a comprehensive real-time national-scale product. However, it is expected that a combination, or blend, of multiple rainfall data sources, in a way that the best characteristics of each one of the available data sources are extracted to create an enhanced product, may be able to meet the needs of real-time and near real-time applications for the whole country. The identification of a suitable blending method and the assessment of its characteristics and limitations are the core objectives of this research project.

Real-time applications

Spatially explicit accumulated rainfall estimates, with quantified uncertainty, at 30-minute or hourly update intervals will greatly improve on rainfall information products currently available for flash-flood forecasting and severe weather warning services. While accuracy and timeliness (i.e. low data latency) are characteristics of utmost importance for real-time hydrometeorological applications, it has been recognised that the ability of a rainfall analysis to accurately capture spatial (and temporal) patterns across a range of scales is often of equal importance, especially in capturing how rainfall is distributed across the region of interest (e.g. catchment).

Delayed mode and near real-time applications

Gridded rainfall analyses at hourly or sub-hourly time steps provide spatially-explicit and temporally-dense synopses of rainfall distribution over a given geographic domain (e.g. continent). Applications of these data abound in hydrology and include: catchment runoff estimation (Chiew and McMahon, 1993; Chiew et al., 2002; Vaze et al., 2011; Bennett et al., 2016); global land surface modelling (Anderson et al., 2007; Yilmaz et al., 2014; Bierkens et al., 2015), climate process studies (Ban et al., 2014); ecophysiology (Donohue et al., 2012); soil erosion and sediment transport modelling (Wilkinson et al., 2009; Baartman et al., 2012; De Vente et al., 2013; Gao and Zhang, 2016).

Evaluating performance of a numerical weather prediction (NWP) model forecast of rainfall is another important application of these data (Ebert et al., 2013). Of course, the rainfall analyses must be independent of the forecast for the verification to be meaningful. The feedback of the verification results to
the modellers provides an assessment of the reliability of the forecasts which is a necessary part of the quality control and assurance.

**Research motivation**

Real-time rainfall data sources in Australia include gauges, weather radar, NWP and satellite estimates. Rain gauges provide a direct estimation of rainfall amounts at ground but the existing network lacks the number and spatial distribution to provide reliable estimates over the vast expanse of the Australian continent, except in close proximity (~1-2 km) to the gauge location. Weather radar provides the local coverage (~100-200 km radius) at high resolution with low data latency, but are indirect estimates of rainfall (i.e. retrievals) and as such require extensive correction, including calibration against rain gauges. The weather radar coverage in Australia is also limited to the more populous areas on the continent leaving vast areas uncovered. NWP provide the required spatial coverage at moderate (~4 km city scale) to coarse resolution (~10 km resolution at continental scale), but rainfall amounts are model estimates that are not directly constrained by surface observation and often have difficulty in accurately representing specific types of rain over some geographic domains, e.g. convective systems in the tropics. Satellite rainfall estimate too possess the needed coverage (at spatial resolution > 10-km resolution for passive microwave retrievals and between 2 – 4 km for geostationary thermal-based retrievals), however they have high data latency (~ few hours) and can have high errors due to challenges in the retrieval algorithms, e.g. orographic rainfall, or inadequate sampling, e.g. missing short-lived small scale events.

Despite the deficiencies in any one of these sources of rainfall data, it is reasonable to expect that it is possible to merge these data in a way that results in a gridded rainfall product that has all the positives (pros) of the individual data and meets the needs of real-time and near real-time applications for the whole country. For example the combination of rain gauge data with weather radar rainfall estimates is an international common practice in meteorological organisations across the world to improve the quality and accuracy of the quantitative rainfall estimates and allow their hydrological use (IAC, 2010). In addition, some countries with an extensive network of weather radars across its territory (such as the USA) incorporate also satellite precipitation estimates in areas where there is limited or no radar coverage (Kondragunta et al., 2005, Zhang et al., 2016)

The blending of gauge, radar, NWP and satellite rainfall estimates could provide a comprehensive real-time national-scale product. Such a product is expected to: 1) provide a seamless coverage of quantitative rainfall estimations across the whole continent filling the gaps in radar and gauge networks, 2) give gridded estimates of accumulated rainfall with at least an hourly update frequency, 3) provide realistic patterns of rainfall distribution and structure across a very wide range of spatial scales, and 4) represent the uncertainty in rainfall estimate explicitly through an ensemble of equally plausible rainfall grids and distributions derived therefrom.
**Study objectives**

This work describes a method of blending sources of rainfall data with a focus on honouring the structure of rainfall across a wide range of spatial scales, i.e. the scaling properties of rainfall, in the resulting gridded rainfall analysis. In presenting a scale-dependent method for blending three rainfall data sets, our objectives in this investigation are to:

- examine the spatial structure of the three candidate rainfall data to see how they represent rainfall patterns over a range of spatial scales;
- develop a weighting scheme and implement a weighted averaging method that exploits the data ability to represent spatial scale;
- conduct a preliminary evaluation of blending method against independent data; and
- recommend a way forward in the development of blending method.
2 Data and method

2.1 Rainfall data sources

Australia has a limited number of rainfall data that possess both national-scale coverage and availability in real-time. For example, a network approximately 1,700 rain gauges reporting 24-hour accumulated rainfall totals in real-time is concentrated around populated areas of the continent (see e.g. Chappell et al., 2013; Renzullo et al., 2016). The relatively small number of gauges and the highly skewed spatial distribution means a gauge-based analysis of hourly rainfall will be hugely unreliable except in areas where the density of gauges is high. Here we look at some alternatives to rain gauge data, namely radar, NWP and satellite data, to provide hourly gridded rainfall analyses for Australia.

The Australian Bureau of Meteorology (BoM) manages an Australia-wide network of more than 50 weather radars1. Reflectivity data for these radars are processed by the BoM to produce a variety of rainfall products, including a 1-km resolution qualitative rainfall intensity (indicating light, moderate and heavy rainfall) at 6 minute intervals (so-called radar loops), a short-term rainfall forecasts, and a continental mosaic of gauge-correct radar data, called Rainfields which provide real-time accumulated rainfall at 30 minute intervals at 2-km resolution (Seed et al., 2007; Seed et al., 2008). For the purposes of this study, we have used Rainfields version 3.0 as the data have undergone the strictest quality control and best-practice gauge correction to produce the highest quality radar rainfall estimate for Australia. Figure 1 gives an example of Rainfields data for 1400 hrs UTC on 2 September 2016 at the continental scale (Fig. 1a) and smaller 512 x 512 km² region (Fig. 1b) that we called Study Area 1.

The Australian Community Climate and Earth-System Simulator (ACCESS) is an implementation of the UK Met Office’s Unified Model (UM) which is used by the BoM to provide weather forecast variables, including rainfall, for the Australian region at a range of spatial scales and lead times. The model generates forecasts every 6 hours, with a latency of 2 – 4 hours. The continental-scale version, ACCESS-R, provides forecasts of hourly rainfall at ~11 km resolution out to 72 hours. Figure 1c provides a T+8 hour forecast for 1400 hrs UTC on 2 September 2016 for a 1600 x 1600 km² area called Study Area 2.

The Integrated Multi-satellitE Retrieval for GPM (IMERG) system generates a series of multi-source rainfall products with quasi-global coverage (Huffman et al., 2013; Huffman, 2015). The most relevant IMERG products for RT applications, which utilises the strengths of both IR and PMW sensing technologies, are the half-hourly ~10-km resolution 3B-HHR products known as ‘early’ (3B-HHR-E with ~ 4 hr latency) and ‘late’ (3B-HHR-L with ~ 12 hr latency). The late IMERG product has been enhanced through the so-called morphing technique (Joyce et al. 2004; 2011) where the microwave rainfall retrievals are advected both forward and backwards in time to the given time of interest. In this study we have used GPM IMERG

v03. There is also a final product from IMERG which utilises monthly composites of satellite data and gauge analyses to give a higher quality product. However because of the latency of this product being on the order of $\sim$ month(s) it will not be considered here.

**Figure 1.** Gridded rainfall data for 1400 hrs UTC on 2 September 2016. Highlighted are Study Areas 1 and 2 in eastern Australia, corresponding to a 512 x 512 km$^2$ and 1600 x 1600 km$^2$ areas in eastern Australia. Rainfall estimates are: (a) Rainfields v3.0, showing the distribution and coverage of radar data for Australia (grey indicates no data areas); (b) Rainfields estimate for Study Area 1; (c) ACCESS-R forecast (T + 8 hours) for Study Area 2; and (d) GPM IMERG 3B-HHR-L for Study Area 2.
2.2 A multifractal blending method: adaptation of STEPS

In this section we present a method for blending rainfall data sources based on their respective ability to observe rainfall at different spatial scales. The method is based on the multiplicative cascade scale decomposition approach featured in the Short-Term Ensemble Prediction System (STEPS) for generating ensembles of short-term rainfall forecasts (nowcasts) from radar observations and NWP (Seed, 2003; Bowler et al., 2006; Seed et al. 2013). While STEPS is primarily a forecasting system, there are components of the method that can be adapted to the problem of generating ensembles of blended rainfall analyses, specifically: decomposition of rainfall 2D field into spatial components; blending of multiple sources of rainfall at the components level; and generating structured noise for ensemble generation. The following sections provide the detail of how these components of STEPS can be used for the analysis problem. Beforehand we describe the data transformation and fundamental tool (the Fourier transform) for the investigation.

The gridded estimates of rain rate, \( R \) (mm hr\(^{-1}\)) used in our investigation are first transformed into reflectivity (or power) space, \( Z \), in decibels (dBZ) using a \( Z-R \) relationship commonly used in radar processing (e.g. Wilson and Brandes, 1979) namely,

\[
\log Z = 10 \log_{10}(200R^{1.6}).
\]

For areas where \( R = 0 \) mm hr\(^{-1}\) we set \( \log Z = 0 \) dBZ. Similarly areas where \( \log Z < 0.93 \) dBZ were set to 0 dBZ, noting that \( \log Z = 0.93 \) dBZ corresponds to a rain rate \( R = 0.042 \) mm hr\(^{-1}\). The coefficients in Eq. (1) follow those used in Bowler et al. 2006, while the 0.93 dBZ thresholds was chosen such that an accumulation of 24 rainfall values of 0.042 mm is 1 mm – our rain/no rain threshold. Figure 2 illustrates the two-fold effect of the transformation: namely it increases the contrast of low intensity pixels, which then highlights spatial patterns of light rainfall (i.e. < 2 mm hr\(^{-1}\)); and it makes the spatial distribution ‘more normal’.

The 2D Discrete Fourier transform (DFT) is a well-established technique for transforming image data from the spatial into frequency domain. It is used in the method described below to provide the power spectral density information of the rainfall data (see Figure 3) and to decompose the rainfall data into scale components (Section 2.2.1). Figure 3 show the power spectra of a Rainfields estimate of hourly accumulated rainfall for 1400 hrs UTC on 2 September 2016 (i.e. Fig. 1b). In the first instance (Fig 3a) we have the 2D power spectra in 10 log\(_{10}\) units, oriented such that the low frequency (low wavenumber) spatial information is at the centre, while the outer edges correspond to high frequency features, including noise, up to and including the Nyquist frequency of 0.25 km\(^{-1}\) (4 km in the spatial domain). The radially averaged (a.k.a isotropic) power spectral density (PSD) of the image is displayed in Fig 3b. The linear feature of the curve on the log-log plot indicates the well-known power-law distribution (inset Fig. 3b) that is characteristic of rainfall and demonstrates structure in the measured rainfall across a wide range of spatial scales. Figure 3b also two district spectral slopes, \( \beta_1 \) and \( \beta_2 \), that are often observed in radar data and is linked to the multifractal nature of rainfall. For the rain system depicted in Fig. 3 that field appears to have greater structure (slightly smaller spectral slope) at the larger scales (i.e. smaller wavenumbers) than at the small scale.
**Figure 2.** Rainfields estimates of rainfall over Study Area 1 (Fig. 1) for 2000 hrs UTC on 20 September 2016: (a) rain rate $R$ in mm hr$^{-1}$; (b) transformed in to ddBZ units, i.e. log $Z$, using Eq. (1) and pixel frequency histograms for $R$ and log $Z$ in (c) and (d), respectively.
2.2.1 Spatial scale decomposition into cascade components

The conceptualisation of rainfall spatial and temporal behaviour as multifractal processes has been accepted for decades (Lovejoy and Schertzer, 1986; Tessier et al., 1993), and the representation via a multiplicative cascade of components of varying spatial scales has been well studied (Gupta and Waymire, 1993; Veneziano et al, 1996; Gupta et al., 2006; Veneziano et al., 2006; Rupp et al., 2012; Foresti and Seed, 2014; Niemi et al., 2014; Pui et al, 2014).

A cascade model represents the spatial structure of rainfall as a sum\(^2\) of components, each corresponding to a band of spatial frequencies, or wavenumbers, from the small scale (high wavenumber) to the large scale (low wavenumber) for a given geographic domain. Here, for simplicity, we assume the geographic domain is a square with extent \(L\) (in km). We index each of our candidate rainfall data sources by \(j \in \{R,N,S\}\), where \(R\), \(N\) and \(S\) denote radar, NWP and satellite rainfall estimates, respectively. Thus a cascade component for rainfall data \(j\) is denoted \(Y_j^k\), and corresponds to spatial scales represented by

\(^2\) A multiplicative cascade model in terms of rain rate becomes an additive cascade model through the log transformation in Eq. (1).
wavenumbers \((\omega_k)\) in the range, \(\frac{2^{k-1}}{L} < \omega_k < \frac{2^k}{L}\), where \(k = 0, \ldots, K\) indicates the component level. The cascade model for rainfall is thus,

\[
\log Z_j = \sum_{k=0}^{K} Y_j^k. \tag{2}
\]

A cascade component is derived via the discrete two-dimensional Fourier transformation (DFT) in forward, \(\mathcal{F}\), and inverse, \(\mathcal{F}^{-1}\), modes. First, in the forward direction, we obtain the 2D power spectrum (e.g., Fig. 3a). We then filter the spectrum using a Gaussian bandpass filter, \(H\), only passing those wavenumbers appropriate to the given cascade level centred on wavenumber, \(\omega_k\). Finally we perform the inverse DFT of the filtered spectrum to get spatial domain representation. That is,

\[
Y_j^k = \mathcal{F}^{-1}\{\mathcal{F}\{\log Z_j\} H(\omega_k)\}, \tag{3}
\]

Cascade components, with \(K = 8\), for the radar-derived Rainfields estimate for 1400 hrs UTC on 2 September 2016 over Study Area 1 (Fig. 1) are displayed in Figure 4. The components are displayed in order from lowest to highest wavenumber, Fig. 4a-i. The low wavenumber components capture the large scale features of the rainfall structure, while the high wavenumber capture the fine scale structure up to the Nyquist critical frequency of 0.25 km\(^{-1}\) (4 km scale). Note that the cascade components displayed have been standardised by the spatial mean and variance at each cascade level. That is, components normalised as,

\[
X_j^k = \frac{Y_j^k - M_j^k}{\sqrt{V_j^k}} \sim N(0,1)
\]

where \(M_j^k\) and \(V_j^k\) are the spatial mean and variance of the \(k\)th cascade for the \(j\)th data product. This standardisation is important, not only for display purposes, but for the blending procedure in the following sections as it removes any systematic difference between data sets. To reconstruct the original rainfall data we sum the non-standardised components, namely \(Y_j^k\), using Eq. 2.
Figure 4. Decomposition of 1400 hr UTC 2 September 2016 Rainfields estimate for Study Area 1. Spectral components for cascade levels: (a) >512 km, (b) 256 – 512 km; (c) 128 – 256 km; (d) 64 – 128 km (e) 32 – 64 km; (f) 16 – 32 km; (g) 8 – 16 km; (h) 4 – 8 km; and (i) < 4 km.
2.2.2 Scale dependent blending of rainfall data

The motivation for the method of blending presented here is that some rainfall data are better at representing rainfall of particularly spatial scale than others. It follows then that the blending of data ought to reflect the ‘skill’ of a given data source at a given scale. Those data with high skill should be weighted higher than those with poorer skill at a given spatial scale. We define this approach as scale-dependent blending and describe in Section 2.2.3 a weighting scheme to be applied at each cascade level.

We follow the approach of STEPS and perform the blending on the standardised cascade components, $X^k_j$. This has the advantage that any systematic differences between different rainfall data sources are eliminated. However converting the weighted sum of components back into rainfall dBZ requires a specification of a spatial mean and variance. To this end, we can assume that one of the candidate rainfall data can serve as a reference. For illustrative purposes, we assume here that the radar rainfall product is the reference. The weighted sum of cascade components is expressed as,

$$ Y^k_B = M^k_R + \sqrt{V^k_R} \sum_{j \in \{R,N,S\}} w^k_j X^k_j, $$

where $M^k_R$ and $V^k_R$ are the spatial mean and variance of the $k$th cascade level of the radar product. The final blended rainfall analysis is then simply a sum over all cascade levels as in Eq. (2). Determining the weights, $w^k_j$, is the topic of the next section.

2.2.3 Calculating the blending weights

A multiplicative error model is often used to describe rainfall uncertainty. In log-space this translates into an additive error model. Thus for each cascade level, $k$, and data set, $j \in \{R,N,S\}$, we can assume a linear error model of the form

$$ Y^k_j = a_{j,T} + b_{j,T} T^k + \epsilon^k_j, $$

where $T^k$ represents the cascade component of the true underlying rainfall for that spatial scale, and $\epsilon^k_j$ is the error in the cascade components representation of the truth at that scale. Coefficients $a_{j,T}$ and $b_{j,T}$ are the intercept and slope of the linear relationship between of component $j$ and the truth, respectively. We further assume a Gaussian error model, that is $\epsilon^k_{ref} \sim N[0, (\sigma^k_j)^2]$, where $(\sigma^k_j)^2$ is a constant error variance. It is this error variance that we seek, as the weights for blending our rainfall data at the cascade level are defined simply as the reciprocal of these error variances. In the following we provide a mathematical description of how the error variances, and thus weights, are obtained. Note that the idea we have implemented is based on triple collocation (TC) (Caires and Sterl, 2003; McColl et al., 2014), where three independent observations are used to infer the error variances in each respectively. Traditionally TC is applied to time series (Dorigo et
al., 2010; Draper et al., 2013; Alemohammad et al., 2015), and error variances represent the uncertainty for a given location across all time. Our approach is to swap space for time and use three data sources to provide uncertainty estimates for a given time over the spatial domain.

The spatial variance of each cascade component is derived from Eq. (5) as

$$\text{VAR}(Y^k_j) = b^2_j, \text{VAR}(T^k) + (\sigma^k_j)^2, \quad (6)$$

where $\text{VAR}(T^k)$ is the variance of the cascade component of the true rainfall spatial pattern for the $k^{th}$ spatial scale. Through ordinary least squares (OLS) regression we determine the slope of the Eq. (5) as

$$b_{j,T} = \rho^k_{j,T} \sqrt{\text{VAR}(Y^k_j) / \text{VAR}(T^k)}, \quad (7)$$

where $\rho^k_{j,T}$ is the correlation between of $Y^k_j$ and $T^k$. Substituting this into Eq. (6) and rearranging to give the error variance,

$$(\sigma^k_j)^2 = \left[1 - (\rho^k_{j,T})^2\right] \text{VAR}(Y^k_j). \quad (8)$$

For our standardised cascade components, this simplifies to $(\sigma^k_j)^2 = 1 - (\rho^k_{j,T})^2$. Equations (8) is attractive in that uncertainty for each cascade components is simply in terms of correlation. The challenge, however, is in having a suitable value for the typically unknown $\rho^k_{j,T}$.

A possible approach to determine $\rho^k_{j,T}$ is through independent means, such as an ‘offline’ assessment against components derived from gauge analyses or other independent, prior or expert information. We may know through an independent means that the radar rainfall data, for example, the value $\rho^k_{R,T}$, is as high as 0.95 for $k = 1$ but reduces to 0.1 at $k = 6$. Given an estimate of $\rho^k_{R,T}$ enables $\rho^k_{j,T}$ for the other two other data, $j = N$ and $S$, to be inferred. That is, using radar data as a reference for cascade level $k$, we have from Eq. (5),

$$\text{COV}(Y^k_j, Y^k_R) = b_{j,T} b_{R,T} \text{VAR}(T^k) \quad (9)$$

with the OLS regression slope values given by Eq. (7). In addition we have,

$$\text{COV}(Y^k_j, Y^k_R) = \rho^k_{j,R} \sqrt{\text{VAR}(Y^k_j) \text{VAR}(Y^k_R)}, \quad (10)$$

where $\rho^k_{j,R}$ is the correlation between the source rainfall data and the reference for cascade level $k$. By equating Eqs. (9) and (10) and rearranging we obtain,

$$\rho^k_{j,T} = \rho^k_{j,R} / \rho^k_{R,T}, \quad (11)$$
which gives the error variance of data $j$, cascade level $k$ expressed in terms of correlation with the reference (radar in this case) as,

$$(\sigma_j^k)^2 = \left[ 1 - \left( \frac{\rho_{jR}^k}{\rho_{R^k}} \right)^2 \right] \text{VAR}(Y_j^k). \quad (12)$$

If it can be assumed, or argued, that the reference gives the best representation (compared with the other data) of rainfall spatial structure for the given cascade level, then the ratio of the correlations in Eq. (12) will be less than 1. However, due to inaccurate characterisation of the reference correlation with truth, or indeed inappropriate choice of the data as a reference, it is possible that the ratio will exceed 1 leading to meaningless estimate of error variance. In such cases a practical work around could be to assume that both data be treated with equal weight.

Alternatively, we can continue with the above arguments and extend to all three rainfall sets. From Eq. (11) we have

$$
\rho_{R,N}^k = \rho_{R,T}^k \rho_{N,T}^k,
\rho_{R,S}^k = \rho_{R,T}^k \rho_{S,T}^k, \quad \text{and}
\rho_{N,S}^k = \rho_{N,T}^k \rho_{S,T}^k. \quad (13)
$$

Rearranging Eqs. (13) and substituting into Eq. (8) we obtain:

$$(\sigma_{R}^k)^2 = \left[ 1 - \frac{\rho_{R,N}^k}{\rho_{R,S}^k} \right] \text{VAR}(Y_{R}^k),
(\sigma_{N}^k)^2 = \left[ 1 - \frac{\rho_{R,N}^k}{\rho_{N,S}^k} \right] \text{VAR}(Y_{N}^k), \quad \text{and}
(\sigma_{S}^k)^2 = \left[ 1 - \frac{\rho_{R,S}^k}{\rho_{R,N}^k} \right] \text{VAR}(Y_{S}^k), \quad (14)
$$

as estimates of the error variance of cascade component $k$ derived from radar, NWP and satellite data, respectively. For the standardised components, we eliminate the ‘VAR’ terms from Eqs. (14). The above are similar to expressions found in TC literature when error variances are expressed in terms of correlation (McColl et al., 2014).

We use the error variances of Eqs. (14) (without the ‘VAR’ terms) for the weights in Eq. (4) for the blended spatial component of cascade level $k$. Given the linear and Gaussian error model, the optimal blended sum is given using weights of the form

$$
w_j^k = \frac{(\sigma_j^k)^{-2}}{\sum_j (\sigma_j^k)^{-2}}, \quad (15)
$$

where $j$ represents radar, NWP and satellite data.
where \( j \in \{ R, S, N \} \). The weighted blend of standardised components is written as

\[
X_B^k = \sum_{j \in \{ R, N, S \}} w_j^k X_j^k,
\]

with an error variance \( \sigma_B^k \) given by,

\[
(\sigma_B^k)^2 = \frac{1}{\sum_j (\sigma_j^k)^2}.
\]

Equations (15) – (17) stem from linear approximation theory for the best unbiased, minimum variance linear combination of data. The error variance of the blend is always smaller than any one of the individual rainfall data sets, and it is smallest when \( \sigma_R^k = \sigma_N^k = \sigma_S^k \), i.e. one third of the error variance. In the next section we describe how the error variance is used in the generation of ensembles of blended cascade components and ultimately the blended rainfall estimates we seek.

### 2.2.4 Rainfall ensemble generation

An ensemble of gridded rainfall estimates provides a measure of the uncertainty in our analysed rainfall. However for accurate spatial representation of catchment processes, particularly the routing of flows in across landscapes in river modelling, it is critical that each ensemble member exhibits the same scaling behaviour of the rainfall analysis\(^3\). This is achieved by adding spatially correlated random noise to the weighted sum of the cascade components for each level. Random noise with spatial structure appropriate to the given cascade level is achieved by convolving a 2D field of Gaussian white noise with the band-passed Fourier spectrum of the rainfall data for that level (Seed et al., 2013; Niemi et al., 2014; Nerini et al., 2017). The result is a set of gridded rainfall estimates, each possessing very similar spectral scaling properties.

The noise generation process is illustrated in Figure 5 using the example of the radar Rainfields estimate for 1400 hrs UTC on 2 September 2016. Subfigures 5a and b show the 2D power spectrum of the rainfall estimates and white noise respectively (insets show the data over the 512 x 512 km\(^2\) spatial domain). Subfigure 5c presents the radially-averaged PSD of the radar (in grey) and random noise fields for two random realisations. Firstly, we have the PSD’s of white noise displayed in blue, whose flat, near constant curve over the entire range of spatial scales is indicative of the ‘structureless’ (i.e. no variation in structure) nature of spatially uncorrelated noise. Displayed in red are the PSD’s of the noise resulting from the imposition of radar spatial correlation structure on white noise. The cascade components of the structure

---

\(^3\) The alternative of, say, white noise introduces spatially incoherent perturbations and the impact of the input to the hydrological model will be spatially inconsistent distributions of water in the landscape.
noise are obtained by taking the inverse Fourier transform of the product of the 2D spectrum of the radar image (Fig. 5a), white noise (Fig. 5b) and applying a Gaussian band-pass filter specific to each cascade level. That is,

\[ Y^p_k = \mathcal{F}^{-1}\{\mathcal{F}\{\log Z^*_R\}\mathcal{F}\{N(0,1)\} H(\omega_k)\}, \tag{18} \]

where \( Y^p_k \) is the noise component for cascade level \( k \), \( N(0,1) \) denotes the 2D image of white noise, and where other terms as defined in Eq. (3). The (standardised) structured noise for the 32-64 km and 16-32 km spatial scales are displayed in Figs. 5d and e, respectively. Finally, an image of structured noise resulting from the sum across all cascade components, with equal weights, is shown in Fig. 5f.

Noise generated using Eq. (18) will possess the spatial structure of the radar data cascade component \( Y^R_k \). More specific to our objectives, we want an ensemble to be generated through perturbations by noise with the same spatial structure of our blended rainfall analysis. To this end, we employ the linear property of the Fourier transform\(^4\) to modify Eq. (18) as,

\[ Y^p_k = \mathcal{F}^{-1}\{[w^R_k \mathcal{F}\{X^R_k\} + w^N_k \mathcal{F}\{X^N_k\} + w^S_k \mathcal{F}\{X^S_k\}] \mathcal{F}\{N(0,1)\}\}, \tag{19} \]

resulting in random noise component for cascade level \( k \) with the structure we desire. The next step is to give the perturbations a magnitude that corresponds to the error variance of the blended cascade component. This is achieved by scaling the standardised noise component as \( X^p_k \sim N[0, (\sigma^k_B)^2] \) with an error variance given by Eq. 17.

By generating multiple, say \( N_E \), realisations of random noise in Eq. (19) we obtain an ensemble of perturbed blended rainfall components,

\[ X^E_k(l) = X^*_B + X^p_k(l), \quad l = 1, \ldots, N_E \tag{20} \]

where each member has spatial structure and error variance characterised by the blended analysis for that cascade level. We use a reference data set (either \( R \), \( N \) or \( S \)) to provide the spatial mean, \( M^E_{\text{ref}} \), and variance, \( V^E_{\text{ref}} \), for that cascade level to convert the standardised perturbed components (of Eq. 20) into unstandardized space, giving \( Y^E_k(l) \). Finally we sum over cascade levels and obtain an ensemble of blended rainfall analyses,

\[ \log Z^*_B(l) = \sum_{k=0}^{K} Y^E_k, \quad l = 1, \ldots, N_E. \tag{21} \]

---

\(^4\) Linear property in the Fourier transform states that if \( X = \mathcal{F}(x) \) and \( Y = \mathcal{F}(y) \), then \( \mathcal{F}(ax + by) = aX + bY \), where \( a \) and \( b \) are constants.
Figure 5. Process of generating random noise with the spatial structure: (a) two dimensional power spectrum of the Rainfields radar estimate for 1400 hr UTC on 2 September 2016 (in inset); (b) 2D power spectrum of spatially independent zero-mean unit-variance random numbers (white noise, in inset); (c) isotropic power spectral density of radar (grey), white noise (blue lines x 2 realisations), and noise transformed to have spatial structure of radar data; (d) standardised cascade level 4 (32-64 km scale); (e) structure noise for cascade level 5 (16-32 km scale); and (f) structured noise created by summing noise for each cascade level 0-8 in standardised dBZ units.
2.2.5 Blending method summary

We proposed a scale-dependent method of generating ensembles of rainfall analyses of hourly accumulated rainfall by blending three sources of (near) real-time rainfall data, namely radar (R), NWP forecast (N) and satellite product (S). The method is summarised as follows:

1. Transform rainfall data, via Eq. (1) into power space, i.e. \( \log Z_R, \log Z_N \) and \( \log Z_S \).

2. Derive cascade components \( Y_R^k, Y_N^k \) and \( Y_S^k, \ k = 1, \ldots, K \) through spectral decomposition (Eq. 3).

3. For \( k = 1, \ldots, K \) (number of cascade levels):
   - standardise cascade components such that \( X_R^k \sim X_N^k \sim X_S^k \sim N(0,1) \) and compute pairwise spatial correlations \( \rho_{RN}, \rho_{RS}, \rho_{NS} \);
   - calculate the error variances (Eq. 14),
     \[
     (\sigma_R^k)^2 = \left[ 1 - \frac{\rho_{RN}\rho_{RS}}{\rho_{NS}} \right], \quad (\sigma_N^k)^2 = \left[ 1 - \frac{\rho_{RN}\rho_{NS}}{\rho_{RS}} \right] \quad \text{and} \quad (\sigma_S^k)^2 = \left[ 1 - \frac{\rho_{RS}\rho_{NS}}{\rho_{RN}} \right],
     \]
     and corresponding weights (Eq. 15) \( w_R^k, w_N^k \) and \( w_S^k \);
   - identify a reference rainfall data set for the given cascade level based on \( \max(w_R^k, w_N^k, w_S^k) \);
   - calculate the weighted sum of the standardised components, giving the blended rainfall component (Eq. 16), \( X_B^k = w_R^k X_R^k + w_N^k X_N^k + w_S^k X_S^k \) and error variance \( (\sigma_B^k)^2 \) (Eq. 17);
   - For \( l = 1, \ldots, N_e \) (number of ensemble members):
     - construct 2D grid of spatially white normally distributed noise (zero mean unit variance) and compute the 2D discrete Fourier transform;
     - generate structured noise, \( Y_P^k(l) \), by convolving the blended cascade component with the noise and the Gaussian band-pass filter specific to this level (Eq. 19);
     - scale the noise such that \( X_P^k(l) \sim N(0, (\sigma_P^k)^2) \);
     - calculate ensemble member and standardise, i.e.
       \[
       X_E^k(l) = X_B^k + X_P^k(l) \sim N(0,1);
       \]
     - calculate ensemble member \( Y_E^k(l) = M_{ref} + \sqrt{V_{ref}} X_E^k(l) \) given the spatial mean \( (M) \) and variance \( (V) \) of the reference rainfall for the cascade level.

4. Generate ensemble of rainfall analyses,
   \[
   \log Z_B(l) = \sum_{k=0}^{K} Y_E^k(l), \quad l = 1, \ldots, N_e.
   \]
3 Results

Results of our application of proposed rainfall blending method over the identified study areas of south eastern Australia (Fig. 1a) are presented in this section. We describe our approach to data selection, spectral analysis, the characteristics of the blending weights, and a preliminary evaluation of the method. We limited our investigations to the month of September 2016, however we recognise that other periods may be better suited for assessing the method under different climate conditions. Nevertheless, September was a particularly wet month for Australia in 2016, with monthly total across the country being 2nd highest on record (highest on record for NSW), as well as having a large number of significant weather events, including flooding and thunderstorm damage across vast areas of eastern Australia (http://www.bom.gov.au/climate/mwr/).

3.1 Selection of NWP forecasts and satellite products

One of our first tasks was to choose the appropriate ACCESS-R forecast and version of GPM IMERG product to use for the investigation. ACCESS-R provides forecast of one hour accumulated rainfall out to 72 hours, updated four times a day. Therefore, for any time of the day there can be up to 12 forecast estimates to serve as our NWP rainfall analysis. It may be reasonable to expect that ‘older’ forecasts (those with greater lead times) could be less reliable than more recent forecasts. In this study we considered the most recent ACCESS-R forecasts (T + 0 – 5 hrs) and the previous forecasts (T + 6 – 11 hrs), noting that the most recent forecasts may not be available for real-time analysis (i.e. latency between 2 – 4 hours). Similarly, neither GPM product is available for real-time application. However our choice of either early (E) or late (L) products could influence the post real-time generation with latency of either 4 or 12 hours.

To help make the appropriate choice of which forecast and product to use, we used estimates of hourly accumulated rainfall for every hour from 0100 hrs UTC 1 September – 2300 hrs UTC 30 September 2016 and evaluated it against coincident Rainfields data over the whole continent. Note that Rainfields data were aggregated from their original 2-km resolution to 10-km for comparison with the GPM and ACCESS-R products. The statistical results are summarised in Table 1.

The distribution of the error statistics is expressed in terms of percentiles from the minimum (taken here to be 5th percentile) to the maximum (95th percentile) over the September time period. Both categories of rainfall product (satellite and NWP) show minor difference in terms of the root-mean-squared error (RMSE) and bias. The biggest difference in performance between the rainfall products is seen in the categorical statistics, where both the probability of detection (POD) and false alarm ratio (FAR) are higher for the ACCESS-R estimates than they are for GPM. The higher proportion of missed rainfall events for GPM is evident in the low POD (< 35%) and moderately high FAR (> 38%). Between GPM_E and GPM_L product,
however, the late product performed better in terms of POD, which is also reflected in the best relative mean-squared error (rMAE)\(^5\) of all the rainfall products.

With regard to addressing which ACCESS forecasts to use, the statistics in Table 1 suggest that ACCESS-R forecast T + 6 – 11 hours perform as well as those from the most recent forecast (T + 0 – 5 hrs). Thus for the application of the blending method we used forecasts with lead time from 6 – 11 hours as the NWP rainfall analyses. And while the GPM statistics are comparable to those of ACCESS, with the exception of the categorical statistics, the slightly better POD and rMAE of late product led to our choice of GPM-L for the investigation. Finally note that all rainfall data in this investigation were reprojected from their original projections into Albers Equal Area for Australia. Furthermore, in the analysis that follows, we resampled coarser resolution NWP and satellite data (~10-km) to the resolution of Rainfields v3.0 (i.e. 2-km) using nearest neighbour sampling.

### Table 1. Summary of error statistics for satellite (GPM IMERG) and ACCESS-R rainfall estimates over Australia between 1-30 September 2016. Evaluation used the Rainfields v3.0 data mosaic for the continent.

<table>
<thead>
<tr>
<th>Percentiles</th>
<th>RMSE (mm hr(^{-1}))</th>
<th>rMAE* (%)</th>
<th>Bias (mm hr(^{-1}))</th>
<th>Corr</th>
<th>POD</th>
<th>FAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPM_E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5%</td>
<td>0.41</td>
<td>80.8</td>
<td>-0.15</td>
<td>0.01</td>
<td>0.00</td>
<td>0.19</td>
</tr>
<tr>
<td>25%</td>
<td>0.56</td>
<td>108.5</td>
<td>-0.09</td>
<td>0.12</td>
<td>0.12</td>
<td>0.32</td>
</tr>
<tr>
<td>50%</td>
<td>0.72</td>
<td>154.1</td>
<td>-0.05</td>
<td>0.26</td>
<td>0.29</td>
<td>0.39</td>
</tr>
<tr>
<td>75%</td>
<td>0.96</td>
<td>218.0</td>
<td>-0.02</td>
<td>0.38</td>
<td>0.43</td>
<td>0.52</td>
</tr>
<tr>
<td>95%</td>
<td>1.27</td>
<td>360.2</td>
<td>0.03</td>
<td>0.54</td>
<td>0.64</td>
<td>0.87</td>
</tr>
<tr>
<td>GPM_L</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5%</td>
<td>0.41</td>
<td>79.0</td>
<td>-0.14</td>
<td>0.01</td>
<td>0.00</td>
<td>0.20</td>
</tr>
<tr>
<td>25%</td>
<td>0.56</td>
<td>104.9</td>
<td>-0.08</td>
<td>0.17</td>
<td>0.15</td>
<td>0.30</td>
</tr>
<tr>
<td>50%</td>
<td>0.72</td>
<td>145.1</td>
<td>-0.04</td>
<td>0.29</td>
<td>0.35</td>
<td>0.38</td>
</tr>
<tr>
<td>75%</td>
<td>0.97</td>
<td>209.1</td>
<td>-0.02</td>
<td>0.40</td>
<td>0.49</td>
<td>0.50</td>
</tr>
<tr>
<td>95%</td>
<td>1.26</td>
<td>355.5</td>
<td>0.06</td>
<td>0.56</td>
<td>0.70</td>
<td>0.88</td>
</tr>
<tr>
<td>ACCESS_R</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25%</td>
<td>0.60</td>
<td>118.0</td>
<td>-0.05</td>
<td>0.15</td>
<td>0.46</td>
<td>0.57</td>
</tr>
<tr>
<td>T + 0 – 5 hrs 50%</td>
<td>0.71</td>
<td>155.0</td>
<td>-0.03</td>
<td>0.26</td>
<td>0.60</td>
<td>0.64</td>
</tr>
<tr>
<td>75%</td>
<td>0.92</td>
<td>199.8</td>
<td>-0.01</td>
<td>0.38</td>
<td>0.71</td>
<td>0.75</td>
</tr>
<tr>
<td>95%</td>
<td>1.29</td>
<td>281.4</td>
<td>0.02</td>
<td>0.55</td>
<td>0.74</td>
<td>0.94</td>
</tr>
<tr>
<td>ACCESS-R</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25%</td>
<td>0.56</td>
<td>119.4</td>
<td>-0.05</td>
<td>0.15</td>
<td>0.44</td>
<td>0.57</td>
</tr>
<tr>
<td>T + 6 – 11 hrs 50%</td>
<td>0.71</td>
<td>155.7</td>
<td>-0.03</td>
<td>0.24</td>
<td>0.59</td>
<td>0.65</td>
</tr>
<tr>
<td>75%</td>
<td>0.92</td>
<td>204.8</td>
<td>-0.01</td>
<td>0.36</td>
<td>0.69</td>
<td>0.75</td>
</tr>
<tr>
<td>95%</td>
<td>1.26</td>
<td>287.6</td>
<td>0.02</td>
<td>0.54</td>
<td>0.79</td>
<td>0.94</td>
</tr>
</tbody>
</table>

\(^5\) This is MAE expressed as percentage of rainfall amount calculated only for hits, i.e. both product and reference detect rain (\(> 0.05\) mm hr\(^{-1}\)).
Spatial structure and power spectral density

The isotropic (ISO) power spectral density (PSD) provides a simple way of examining the spatial structure of rainfall as captured by our three rainfall data sources. Figure 6 shows the ISO PSD for Rainfields (Radar), ACCESS-R (NWP) and GPM-IMERG (Satellite) data over Study Area 1 (512-km domain) and Study Area 2 (1600-km domain) in Figs. 6a and b, respectively. They were derived by radially averaging the 2D Fourier spectra of the 60 min rainfall accumulation at 1400 hrs UTC on 2 September 2016. We have not included a PSD for radar over Study Area 2 to keep Fig. 6b uncluttered.

All three data possess the linear feature that is indicative of the power-law scaling of rainfall over a broad spatial extent of the areas. Recall that the slope of the curves, the spectral slope $\beta$, is the exponent of the power law distribution (see Fig. 3b) and can be used to assess the organisation of the structure in our rainfall data over a range of spatial scale (Harris et al., 2001). Small values of $|\beta|$ indicate disorder and random structure, whereas high values of $|\beta|$ are indicative of smooth, highly organised structure.

In Study Area 1 (Fig. 6a) the radar curves appears to have greater power (i.e. offset above the others) over most of the area but flattens out considerably at the large scale (i.e. small wavenumber), suggesting a lack of structure at the very large scale. Satellite data over both study areas appears to be consistently linear from the small (high wavenumber) to the large scale (small wavenumbers). Note that any structure inferred for scales $< 10$ km ($> 0.1$ km$^{-1}$) for the NWP and satellite data is likely an artefact of the resampling and is therefore unreliable.
In Fig. 6 a broad-scale feature spanning 25 – 125 km scales (highlighted in grey) is observed in the PSD of NWP in both study area. Over this region the spectral slopes for radar and satellite are statistically the same, $\beta_R \approx \beta_S \approx -2.3$. For NWP we have $\beta_N = -3.4$, which is significantly less than those of the radar and satellite, suggesting a smoothness over the scale which perhaps is not unexpected for a model derived rainfall estimate. We note that there is also agreement (within statistical uncertainty) in the spectral slopes of the respective rainfall data between study areas. This consistency in the inferred structure for scale in the range 25 – 125 km across over study areas suggests spatial stationarity of the statistical properties for the particular weather system.

Figure 7. Time series of spectral slope, $\beta$, values for the 25 – 125 km scale for Rainfields (Radar), ACCESS-R (NWP) and GPM IMERG (Satellite) data over September 2016 time period: (a) spectral slopes derived for 512 x 512 km study area for the three rainfall estimates; (b) and (c) spectral slope time series for the Study Area 1 (solid lines) and Study Areas 2 for ACCESS-R and GPM IMERG rainfall, respectively.
To see if the ‘smoothness’ in the NWP rainfall data over the spatial scale 25 – 125 km is a recurring feature in the data, we extended the analysis to all data over the September study period for both study areas. Time series of the spectral slopes derived for the 25 – 125 km scale from the ISO PSD of the radar, NWP and satellite data are displayed in Figure 7. In Fig. 7a we have the time series of $\beta$ for the radar (black), NWP (blue) and satellite (red) data for Study Area 1. Gaps in the time series where the curves drop below the x axis correspond to those times when no rain was detected over the 512 x 512 km$^2$ area. At this scale, over this time period, there is general agreement between the spectral slope of radar and satellite data, with $|\beta_R| \approx |\beta_S| \approx 2.7$, while the slope for NWP is generally higher ($|\beta_N| = 3.2$). When we extended this analysis to Study Area 2 for the NWP and satellite data (dashed lines in Figs. 7b and 7c, respectively) we observed the same outcome, namely that $|\beta_N| > |\beta_S|$, with averages over September of 3.2 and 2.8, respectively. In addition we found that when only slopes were calculated for Study Areas 1 and 2, i.e. rain event detected in both, there was a high degree of consistency between the spatial domains.

Having identified a consistent smoothness in the ACCESS-R representation of rainfall spatial structure for features in the 25 – 125 km scales, it may be of interest to know how this, and other structural information, impacts the weighting strategy (Section 2.2.3). The expectation is that the pairwise correlation of data sets will reveal the two data most similar in their ability to represent the structure of rainfall over a given range of spatial scales and, as such, be assigned higher weights in the blended product than that of the other data set. Agreement (high correlation) between data need not imply increased ability to identify spatial structure. In the following we define the spatial scale of each cascade level for the two study areas.

### 3.3 Spectral decomposition into cascade components

For each of the study areas we decomposed the transformed rainfall data (from Eq. 1) into spatial components corresponding to the cascade levels provided in Table 2. Each level is defined in terms of upper and lower bounds (used to define the full-width half-maximum limits of the Gaussian bandpass filter) that were obtained by partitioning the study domains into quarters, then each quarter being partitioned into quarters, and repeating the process until the Nyquist sampling limit (4 km resolution) is reached. For Study Area 1 this resulted in 9 levels and for Study Area 2 it resulted in 10.

To see where these cascade levels sit on the ISO PSD we take the rainfall data for 1000 hrs UTC on 14 September 2016 over Study Area 1 given in Fig. 8. The Rainfields data (Fig. 8a) show an extensive rainfall system covering the central northern part of the area, extending eastward into fine-scale patches of rain, with another large rain area off the NSW coast. The GPM product (Fig. 8b) shows coverage similar to Rainfields albeit less patchy, extending more to the northwest and missing the band of rain in the southwest. The ACCESS-R data shows similar coverage to the Rainfields data but with considerably less power. The ISO PSD for each data set is given in Fig. 8d and reveals the smoothness feature we observed in Section 3.2. This feature corresponds to cascade levels 3 – 5 for Study Area 1. From Table 2 we see that this same range is spanned by cascade levels 4 – 6 in Study Area 2.
Table 2. Cascade level definition for Study Areas 1 (512 x 512 km$^2$) and 2 (1600 x 1600 km$^2$).

<table>
<thead>
<tr>
<th>Cascade level</th>
<th>512 x 512 km$^2$ study area</th>
<th>1600 x 1600 km$^2$ study area</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>km</td>
<td>km$^{-1}$</td>
</tr>
<tr>
<td></td>
<td>upper</td>
<td>lower</td>
</tr>
<tr>
<td>0$^6$</td>
<td>&gt; 512</td>
<td>&lt; 0.002</td>
</tr>
<tr>
<td>1</td>
<td>512</td>
<td>256</td>
</tr>
<tr>
<td>2</td>
<td>256</td>
<td>128</td>
</tr>
<tr>
<td>3</td>
<td>128</td>
<td>64</td>
</tr>
<tr>
<td>4</td>
<td>64</td>
<td>32</td>
</tr>
<tr>
<td>5</td>
<td>32</td>
<td>16</td>
</tr>
<tr>
<td>6</td>
<td>16</td>
<td>8</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>9</td>
<td>---</td>
<td>---</td>
</tr>
</tbody>
</table>

$^6$ DC level in the Fourier spectrum.
We performed the spatial scale decomposition and computed standardised cascade components $X_R^k$, $X_N^k$ and $X_S^k$ for all available rainfall data over the period 0100 hrs UTC on 1 September – 30 September 2016. For Study Area 1 with maximal radar coverage missing data in the Rainfields product were filled with log $Z = 0$ dBZ prior to conducting the decomposition. For Study Area 2 this too was done, however the cascade components were masked out in these areas so as not to introduce spurious values in the calculated statistics that may inadvertently influence blending results.
3.4 Blending weights

For those times over the study period were all three rainfall data were available we computed the error variances using Eqs. (14). Recall that another way of expressing the errors is in terms of data correlation with the truth (Eq. 8) for each cascade level, $\rho^k_R$, $\rho^k_N$ and $\rho^k_S$. The distributions of the calculated correlations over the time period are displayed in Figure 9 for Study Area 2. The median (50th percentile) correlations for the Rainfields (R), ACCESS-R (N) and GPM (S) data are displayed for each cascade level in Fig. 9 as solid lines, with the interquartile range (IQR) of values indicated by shaded polygons. The correlation for each follows the general trend from high correlation (therefore, low error variance) at the large spatial scales (i.e. low cascade values) and decreasing to low correlation (high error variance) at the small scale.

For most of the cascade levels in Fig. 9 there is significant overlap in the IQR’s, with each spanning a wide range of correlation values which indicates considerable variation over time. For most of the cascade levels 0 – 5 the radar and satellite data have similar median values but diverge for $k \geq 6$, corresponding to spatial scales < 50 km. Noticeably there is a drop in the correlation of the ACCESS-R data for cascade level 4, which corresponds to spatial scales 100 – 200 km. This would indicate a large error variance and consequently lower weight given the NWP data in the blended product.

![Figure 9](image-url)  
*Figure 9. Correlation of the cascade component for each rainfall data set against the ‘truth’ (in Eq. 8) for each level (Study Area 2). Value calculated for every hour between 0100 hrs UTC 1 September – 2300 hrs UTC 30 September 2016. Shaded area represents the interquartile range of the values and the solid line represents the median value over the time period.*
To see how these correlations influence the blending, we calculated weights for each data set using Eq. (15) and the error variances derived over the study period for both study areas. We note that for high cascade levels there may be high correlation between NWP and Satellite, particularly at scale < 11-km resolution as the data have been resampled to sub-grid resolution of the respective data. This may lead to high weighting of the data which may be prudent to ignore. Nevertheless at each time step we calculated the weights and tallied for each cascade level which of the three data was given highest weight. The number of times a given data set was assigned the highest weight (expressed as a percentage of times) is plotted in Figure 10 for Study Areas 1 (Fig. 10a) and 2 (Fig. 10b). Recall that the cascade levels (Table 2) for the two study areas represent difference spatial scales. We have indicated on both Figs. 10a and b the scales in the range 50 – 400 km for the respective study areas. The figure shows that for the scales in the range 50 – 400 km that ACCESS_R is frequently given the lowest weighting, in other words assigned the highest weight only ~20% of the time. The Rainfields data are assigned the highest weight more often than the GPM data over this range, while ACCESS_R is frequently assigned the highest weight for scales above 500 km.

Figure 10. The frequency with which the rainfall data is given highest weight amongst the sources (radar, NWP and satellite R, N and S respectively) over the study period for each cascade level (defined in Table 2) for study areas (a) 1 and (b) 2. Highlighted are the cascade levels spanning the 50 – 400 km spatial scales for both study areas.
3.5 Blended rainfall analyses

Scale-dependent blended rainfall analyses were generated using the derived weights for Study Areas 1 and 2 for the times when all three source data (Rainfields, ACCESS-R and GPM) were available. Disruption to the GPM mission IMERG processing system meant that there were some times that the 3B-HHR-L products were not available. The impact for this study was that of the possible 720 time steps in the study period, there were 28 for which we did not conduct the blending. Ensembles were comprised of 50 members.

An example of the blended results is the hourly rainfall analysis for 2000 hrs UTC on 29 September 2016 is provided in Figure 11. Figures 11a-c are the source rainfall data from Rainfields, ACCESS-R and GPM respectively. The data shows three distinct, moderate-sized rainfall events over the 1600 x 1600 km$^2$ study area: the first, off the southwest coast of the region near the state border between South Australia and Victoria; the second straddling the southern New South Wales and northern Victoria boarder; and the third is off the southern Queensland coast in the northeastern part of the study area. All three data broadly identify the same events, however they differ primarily in the amount of rainfall but also the structure, especially at the fine scale. For example in the southwest, ACCESS-R captures quite a large distribution of rainfall that extends up into Victor Harbour, South Australia, that is not as expansive in the GPM data nor in the Rainfields data due to lack of radar coverage. In the blended product (Fig. 11d) the event has both the coverage of the NWP estimate as well as the small scale structure imparted by both satellite and radar data. The rainband over the northern Victoria and southern New South Wales border is not captured by Rainfields, however it appears as a smooth, narrow rainband is in the ACCESS-R data (Fig. 11 b) but a lot more clustered in GPM (Fig. 11c). The blended for this event shows enhanced fine scale structure with higher rainfall values than in the original GPM, and an overall structure that appears quite realistic. In the northeast of the region the blended result provides a good illustration of the usefulness of the scale dependent blending, which allows the NWP and satellite to influence the structure at the large scales and the radar to provide the very fine scale information. Overall, the spatial smoothness of the ACCESS-R data is augmented by the Rainfields and GPM data via the scale-dependent blending.

A situation that one is likely to encounter is when only two of the three data are available. This may be due to data latency or dropout in the case of real-time applications, e.g. satellite data are not available in the time constraints of the application or its production system failed, leaving only Rainfields and ACCESS-R. Alternatively a data set may be deliberately withheld from the blending for validation purposes, e.g. only Rainfields and GPM data were used in the blend to evaluate ACCESS-R forecasts. In these cases, the simultaneous estimation of blending weights from Eqs. (14) for the cascade levels is not possible. As a potential work around, we could use the median value of inferred correlations with the ‘truth’ (Fig. 9) derived for the two available data, in a ‘climatological’-based blending. This approach was trialled over the September study period using the correlation values from Section 3.4 mimicking the scenarios when either only Rainfields and ACCESS-R or Rainfields and GPM data are available. Example results are provided in Figure 12.
Figure 11. Hourly accumulated rainfall (in dBZ) for 2000 hrs UTC on 29 September 2016. The source data: (a) Rainfields, (b) ACCESS-R T+8 hrs, and (c) GPM IMERG_L; and (d) the blended result corresponding to the ensemble mean.

The source rainfall data for 0600 hrs UTC on 16 September are presented in Fig. 12 a – c: Rainfields, ACCESS-R T+12 hrs and GPM, respectively. The blended analysis using Rainfields - ACCESS-R only (radar-NWP) and Rainfields - GPM only (radar-satellite) based on the climatological derived weights for each cascade level is given in Figs. 12 d and e, respectively. These blends resemble, as one would expect, the weighted average of the original data most notably in the circular feature in the centre of the north boundary of the area in Fig. 12d, and in the features in the central eastern coast in Fig. 12e. The scale-dependent blending with weights derived from Eqs. (14) simultaneously is displayed in Fig. 12f.
**Figure 12.** Rainfall analyses for 0600 hrs UTC 14 September 2016. Source products: (a) Rainfields, (b) ACCESS-R T+12 hrs, and (c) GPM IMERG-L. Blended analyses: (d) radar-NWP; (e) radar-satellite; and (f) radar-NWP-satellite.
Each of the analyses presented in Figs. 11 and 12 correspond to the mean of a 50 member ensemble of rainfall for a given hour in September 2016. Individual members of an ensemble, recall from Section 2.2.4, are generated from random noise convolved with the Fourier spectrum of the blended rainfall at each cascade level (Eq. 19). Two arbitrarily chosen members from the rainfall analyses of 1400 hrs UTC on 2 September 2016, derived from the climatological blending of Rainfields and ACCESS-R for Study Area 1, are displayed in Fig. 13 a and b. The ensemble mean is given in Fig. 13 c and we can see that the perturbations have introduced small-scale noise in parts of the study area (e.g. northwest corner) where, at least according to radar data in Fig. 13e, there is likely no rain. Furthermore, the resolution of the ACCESS-R data, of ~ 11 km, is evident by the apparent blockiness (Fig. 13f). This suggests that the weighted contribution of the NWP cascade at this resolution is too strong, and as such we took the decision to remove the influence of NWP and satellite products for subsequent analysis by assigning weights of zero to cascade components with scales < 10 km.

To examine the spatial structure of the ensemble we computed the ISO PSD of each of the 50 members and plotted then against the source Rainfields and ACCESS-R data in Fig. 13d. Note that at each cascade level the standardised perturbed components are rescaled by the spatial mean and variance of the Rainfields components, thus the overlap of the radar PSD (black) with the ensemble (grey). The influence of the NWP at the large scale can be seen by the alignment of the slope of the ensemble PSD with that of the NWP PSD (blue).

The spread of the ensemble members at the larger scales is a feature requiring further analysis as the expectation is that these levels in the cascade have higher correlations with the ‘truth’ (Fig. 9), which result in smaller error variances and even smaller blended error variance, $(\sigma_B^2)$ (see Eq. 17). We observed, however, a sizable spread at the large scale which suggests an inconsistency between our expectations and way the method was implemented that requires further investigation. We took the decision to give no perturbation on cascade levels 0 – 2 for subsequent analyses (note however that Figs 11-13 show results that include perturbed cascade components for levels 0 – 2). Also, we observed at the high cascade levels, i.e. small scale < 10 km, that the PSD in Fig. 13d flattens out which suggests a white noise structure. Given that the method often weights the influence of NWP and satellite higher than radar at this scale (Fig. 10), which is finer than the original resolution of the data, we thought it prudent to leave the radar components unperturbed at this scale as the derived correlations are likely to be meaningless.

Finally, we converted the analysed rainfall for September 2016 from power units (dBZ) into rainfall (mm hr$^{-1}$) by inverting Eq. (1). Figure 14 illustrates the impact of assigning zero weights to scales < 10 km for NWP and satellite rainfall, as well as leaving the blended analyses unperturbed for cascade levels 0 – 2 and > 7 on the ensemble estimates. The mean and standard deviation of the ensemble for 2000 hrs UTC on 20 September 2016, resulting from blending NWP and radar is displayed in Fig. 14a and b, respectively. Note that both figures show the distinct pattern of the radar coverage in blended product. This is an artefact that we expect could be eliminated with a more thoughtful treatment of the weighting of NWP, satellite and noise contribution to the small scale cascade levels.
Figure 13. Blended radar-NWP rainfall analysis for 1400 hrs UTC on 2 September 2016, Study Area 1: (a) and (b) two members from the 50 member ensemble; (c) ensemble mean; and (d) isotropic PSD for the source NWP (blue) and radar (black), along with the PSD of the 50 members of the ensemble (grey). Original Rainfields (e) and ACCESS-R (f) rainfall estimates included for comparison.
3.6 Evaluation of blending results

As a preliminary evaluation of the blending method, we compared the 24-hour accumulated total from the ensemble means of our rainfall analyses against gauge-derived gridded analyses at the location of real-time gauges (Fig. 15). Specifically we used the Bureau’s near real-time gridded estimates of daily rainfall developed under the Australian Water Availability Project (AWAP, Jones et al., 2009), which represent the 24 hour to 9 am local time at ~5-km resolution across Australia. We chose these gridded data as they avoid issues of change-of-support (i.e. point-to-pixel comparisons) and, more importantly, are derived from a set of gauges largely independent of those used in the processing of Rainfields. In addition, we used rainfall values of only those grid cells coincident with the network of daily gauges data (independent of those used in calibrating Rainfields) as they are more representative of observed rainfall than those at locations away from gauges.

Figure 15 shows a typical distribution of these validation data relative to the radar coverage for Study Area 2. On any given day, there are approximately 1,050 gauges reporting 24 accumulated rainfall in real-time for over the area comprising between 850 – 950 gauge locations within radar network coverage (noting that coverage can be intermittent), and 100 – 200 gauges outside of radar coverage. For Study Area 1 (delineated in Fig. 15) there are ~ 220 gauges most of which are inside radar coverage.

---

5 At the time of writing this report the authors believed that these data were independent of Rainfields. Subsequently we have confirmed that approximately half of the gauges displayed are indeed independent. Future investigations will focus validation on only those truly independent data.
Figure 15. Location of the rain gauges (blue dots) reporting 24 accumulated rainfall in relation to the radar coverage (white areas) and Study Area 1. Note that the gauge data are independent of the derived Rainfields, ACCESS-R and GPM rainfall products. The distribution of gauges is typical for any given day. In this figure there are a total 1,070 gauges over the 1600 x 1600 km\(^2\) area: 220 inside 512 x 512 km\(^2\) area (black square); 905 inside radar coverage; and 105 outside of radar coverage (grey areas).

Table 3. Comparison of source and blended rainfall analyses (based on ensemble mean) with 24-hour accumulated rainfall from gauge-based analysis extracted for individual gauge locations during September 2016. Statistics calculated are difference (product – observation) in areal average (DAA) and mean absolute error (MAE). Highlighted are the best statistic in each column.
The evaluation results are summarised in Table 3 for the three source rainfall data and the various blended estimates over the month of September 2016. Two metrics are presented: (i) the difference in areal average (DAA); and (ii) the mean absolute error (MAE). The first (DAA) provide a measure of the difference in representation of the mean rainfall (extracted at the gauge locations) over the area compared with that derived from the gauge data. The second (MAE) is a measure of the average error in the estimates of rainfall from the source and blended products over the area. The metrics were calculated for Study Area 1 for all gauges in the area, and Study Area 2 for the set of gauges within radar coverage and for those outside coverage.

For Study Area 1, the DAA values in Table 3 show that ACCESS-R alone provides an estimate of mean rainfall that is closest to that derived by the gauge data, differing by 0.43 mm which is on average about an eighth of the mean rainfall daily total. ACCESS-R estimates also have the lowest MAE (2.19 mm). The estimate with the biggest difference is the blend of Rainfields and GPM (R+S) with a DAA of 2.5 mm which is likely due to the high MAE of the GPM product (3.17 mm). Generally speaking, for Study Area 1, there appears to be little difference in performance between the source data and the blended products, however we can see that the addition of NWP into the blend of radar and satellite (R+N+S) is an improvement on radar and satellite (R+S).

For Study Area 2 in the area with radar coverage GPM provides the estimate of average areal rainfall that is in closest agreement with gauge-based estimate, underestimating on average by 0.42 mm, followed by ACCESS-R which overestimates mean rainfall by 0.78 mm on average. The best performing source data product in terms of MAE is ACCESS-R with a value of 3.66 mm, which reduces to 3.28 mm in the blended R+N estimate. The blends which include ACCESS-R (R+N, R+N+S) are shown to result in the greatest improvement with, for example, a decrease in MAE of 1.9 mm from that of Rainfields. Note too that Rainfields shows a DAA that is comparable in size to the MAE of the blended analyses in this area.

For Study Area 2 with no radar coverage the best estimate of average areal rainfall is provided in the blend of Rainfields, ACCESS-R and GPM (R+N+S) with DAA and MAE of -0.45 and 3.63 mm, respectively, compared to -1.75 and 3.61 mm for the blend R+N. This is a reduction in MAE over GPM alone of 2.8 mm. Generally, for Study Area 2, blending estimates are an improvement on the data alone, especially when the blend includes NWP. From Fig. 10 we see than ACCESS-R is most often weighted highest in the blend for cascade levels 0-3, i.e. spatial scales > 300 km, and levels 6-8 (12 – 50 km). Given that the ACCESS-R data have low DAA and MAE, it is reasonable to conclude that the NWP is responsible for the improvements in the blends here. Rainfields is used to determine the blending weights to be applied to the whole study area, but are determined for those areas where all three data have coverage. What remains to be assessed is how much of an improvement comes from radar and or satellite given they dominate the blend for cascade levels 4 and 5.
4 Discussion

We have explored a scale-dependent blending of radar, NWP and satellite rainfall estimates where the blending weights are calculated in a way that is related to how well a data source captures rainfall structure at a given spatial scale. Using a multiplicative cascade model representation of rainfall and decomposing the source data into common spatial component (Section 2.2.1) provided a means of examining the weighting behaviour across a range of scales. At each spatial scale weights were determined through the pairwise correlation between the components of the source data (Section 2.2.3; Eqs. 14). Furthermore we described a method of generating structured noise (Section 2.2.4) with which to perturb the blended components and results in an ensemble where each member has the scaling properties of the rainfall analysis.

4.1 Source real-time rainfall data

The choice of rainfall data used in this study was guided by what is available to in real- to near-real-time in Australia (Section 2.1), however they are typical of the sources of rainfall information available to most national weather agencies around the world. Specific to the Australian Bureau of Meteorology are the Rainfields national radar mosaic and the ACCESS-R NWP rainfall forecasts. These two data are likely to provide the first level of blended nation-wide real-time rainfall analysis because of their low data latency, while a second level of rainfall analysis more suited for near real-time applications will result from the blending of radar, NWP and satellite rainfall.

Gauge-corrected Rainfields v3.0 represents the best quality radar estimates to-date of 60-minute accumulated rainfall (at 30-min time steps) and is available within 20 minutes of the time of interest. In addition, the correction of radar reflectivities with observed accumulated rainfall from the network of real-time gauges is considered the most appropriate use of these data, as gridded analysis based on gauge data alone will be a highly unreliable and thus inadequate for national scale applications. While Rainfields provides real-time rainfall data at 2-km resolution, the coverage is concentrated around the population centres of the country.

ACCESS-R provides forecasts of 60 minute accumulated rainfall out to 72 hours 4 times a day, each with latency up to 4 hours. The cursory evaluation in Section 3.1 showed that there was negligible difference in performance between most recent ACCESS-R rainfall forecasts, i.e. forecasts with 0 – 5 hour lead times, and those from 6 hours prior, i.e. 6 – 11 hours lead time. Provided the results hold across different seasons and geographic domains, this means that the slightly older ACCESS-R forecasts are suitable for use without delay to ‘fill the gap’ in the radar coverage.

The satellite rainfall data source used was GPM IMERG v03 ‘Late’ multi-satellite product. These data were chosen purely as being representative of the current state-of-the-art in near real-time (quasi-)global mapping of precipitation. Alternatives sources could include the Tropical Rainfall Measuring Mission (TRMM) multi-satellite analysis system (TMPA, Huffman et al., 2007), NOAA’s Climate Prediction Center
morphing technique (CMORPH) and QMORPH (Joyce et al., 2004; Joyce and Xie, 2011) and Global Satellite Mapping of Precipitation (GSMaP) and GSMaP-NOW (Kubota et al., 2007). Each product provides ‘instantaneous’ rain rate at time steps varying between 8 – 48 times per day. They possess comparable spatial resolution ~10 – 25 km across most of the globe. The quality of satellite rain rate retrievals varies, particularly with season and intensity, and has been the topic of much research over the last decade (e.g. Ebert et al., 2007; Tian et al., 2009; Sappiano et al, 2010; Liu 2015; Guo et al., 2016). However, the latency of the data is one of the main limitations to real-time application. With latency ~ 4 hour for most systems, the satellite rainfall estimates are more suitable for near real-time analyses. A possibility could be to use rain rate retrievals from infrared sensors carried by geostationary satellites (e.g. Scofield and Kuligowski, 2003) as these data are available in real-time. While such rain rate retrievals have many well-known issues they could potentially provide better estimate of convective rainfall in the tropics than NWP, and serve as an interim stopgap in rainfall information needs. The choice of which satellite rainfall data to use in a blended rainfall analysis system requires on-going evaluation.

The blended rainfall analysis based on radar and satellite could be used as independent data to verify NWP forecasts. The quality of the R+S blended estimates need to carefully evaluated. The preliminary evaluation (Table 3) often showed Rainfields and GPM blended analysis to be the poorest performer. This was likely due to the high MAE in GPM estimates of the satellite product. A more detailed evaluation of the GPM data would help in developing correction (calibration) factors to adjust the satellite estimates to have better agreement with observed rainfall.

4.2 Spatial structure and blending weights

We characterised spatial structure using the isotropic power spectral density (PSD) for each data (e.g. Fig. 6). All three data demonstrated the expected scaling behaviour of rainfall, with the GPM and radar showing highest consistency for the 512 x 512 km² study area (Study Area 1) evident by similar spectral slopes over the widest range of scales (e.g. Figs. 6 and 7). The ACCESS-R rainfall gave the highest absolute value of spectral slope for the spatial scales in the range 25 – 125 km, indicative of smoothness, i.e. high level of structure, in the NWP forecasts at that scale. Furthermore, this smoothness appeared to be consistent across most of the analysis times over the study period and across the study areas (Fig. 7). This is a range typical of stratiform rainbands and the smoothness is discernible in the ACCESS-R forecasts (e.g. Fig. 8).

The blending weights were determined simultaneously when all three data sets were available in an approach similar to triple collocation (Eq. 14). While the values varied hour-to-hour, there was some consistent findings over the study period about how the weights varied with spatial scale (e.g. Figs 9 and 10). Specifically we found that for more than half of the time, ACCESS-R estimate were given the highest weights for scales 512 – 1600 km over Study Area 2. We also found that over the range 25 – 125 km, cascade levels over this range for ACCESS-R were given the lowest weight on average, which corresponds to the range where the rainfall forecast show greatest smoothness.
Some consistency in the weighting of cascade levels for the respective source data was observed across study areas (Fig. 10). It was reasonable, therefore, to develop a set of temporally ‘static’ weights (based on the median correlations in Fig. 9) in the event that one of the three rainfall data sets were not available. This was trialled as R+S and R+N in the evaluation (Table 3) and there was no apparent deterioration in performance when compared with the simultaneous, or ‘dynamic’, retrieval (R+N+S). In fact, often R+N outperformed R+N+S. We note that this was a preliminary evaluation and a more controlled trial is required to see if the static approach is valid. And while the static approach eliminates the need to compute the weights at each time step, resulting in minor improvement in computational efficiency, it is the suitability of the approach for real-time applications when only two data sets are available that has the most important implication.

4.3 Non-stationarity and ensemble generation

The method was applied to two study areas, the largest (Study Area 2) corresponds to just under 33% of Australia’s land area. The method makes the assumption that the one power spectrum, single mean and variance standardisation parameters, and set of cascade weights, are valid over the study area(s). While the consistency in behaviour of the spectral slope (Fig. 7) and weights (Fig. 10) between study areas suggests a degree of validity of the assumption, at least in terms of structure, the variable nature of rainfall regimes and associated large non-stationarity in the spatial statistical properties over large geographic extents (e.g. due to topography, prevailing meteorological directions, proximity to coast, and land cover) may compromise the method’s suitability over continental domains.

A possibility for future development of the method could include a series of localised analyses over a number of small geographic areas for which the assumption of stationarity holds. Neirini et al. (2017), for example, employed just such an approach to tackle non-stationarity in their new approach to the generation of structured noise over large areas. They conduct a series of small-scale Fourier analyses which they then concatenate over the larger geographic domain. In the context of our proposed method, this will result in spatially varying weights for each cascade level, as well as noise perturbations that are reflective of the variability in rainfall scaling properties over the continent. Fourier-based methods for generating noise are demonstrably faster to implement than, for example, geostatistical methods of stochastic simulation (Goovaerts, 1997), and as such more suited to real-time applications. Assessing the suitability of the Neirini et al. (2017) approach to that of single PSD approach (Section 2.2.4) will need to be made in the context of trade-off between added skill and algorithmic complexity.

The error variance of the blended analysis (Eq. 17) is used to scale the structured noise of the standardised components of each cascade level and generate the ensemble of rainfall analyses (Eq. 20). While the ensemble members were observed to have the desired spatial structure (e.g. Fig. 13) the magnitude of the variances at the extremes in spatial scale led to some unexpected results. For example, initially perturbations were adding too much noise at large scales, resulting in rain appearing where no rain was likely
to have occurred. In addition the higher weights that the method assigned to NWP and satellite rainfall for cascade levels > 7 (i.e. < ~10 km) resulted in structureless white noise, which dominated those levels (see Fig. 13d) and swamped the radar influence at the fine spatial scale. The decision was made for the current study to leave the extreme levels \(k = 0 – 2\) and \(k > 7\) unperturbed. The impact on the resulting rain rates is shown in Fig. 15 in which the discontinuity in radar coverage is often discernible. The treatment of the error variance of the large and small scale components needs revisiting in the method.

Finally, converting the blended cascade levels into original units requires scaling with the spatial mean and variance of the reference data, i.e. \(M^k_{\text{ref}}\) and \(V^k_{\text{ref}}\), respectively. These scaling factors were chosen based on which data was given the highest weight. Given the weights are related to structure they are not necessarily related to the bias in rainfall product. A better approach might be to identify through independent offline investigation which of the data are better correlated with gauge-based analysis for the different cascade levels. This would be conducted using a gauge-based analysis over an area with sufficiently high density of gauges.

### 4.4 Evaluation

Most of the blending results presented have focussed on the analysis ensemble mean. Visually (Figs. 11 – 14) the blending has achieved the primary motivation for this work and that is to fill the gaps in the radar network coverage with data that honour the spatial structure of rainfall across wide range of spatial scales (i.e. provide realistic rainfall patterns). So too, the individual rainfall analyses that comprise the ensemble possess the same scaling properties of the blend (Fig. 13).

The quantitative evaluation presented here was a simplistic and incomplete first-pass assessment of the method based on a narrow selection of performance metrics (Section 3.6). Comparisons of the source and blended rainfall analyses were conducted on accumulated ensemble means to be consistent with the independent daily validation data (i.e. gauge-based analysis rainfall values at set of gauge locations). ACCESS-R gave the best estimates on average for Study Area 1, but broadly there was little difference between the performance between the source data and the blended estimates. For Study Area 1, Rainfields is most often chosen as the reference, followed by the GPM data (Fig. 10). However, these two data are seen, on average, to have high bias (DAA) and mean absolute error compared with ACCESS-R estimates (Table 3). This results in little improvement in the blended analysis. For Study Area 2 ACCESS-R is most often chosen as the reference for spatial scales > 300 km and between 12 – 50 km (Fig. 10). This may explain why blended MAE reduces as much as 1.9 mm compared to Rainfields in radar coverage areas, and the reduction of 2.8 mm from GPM MAE in areas with no radar coverage. A more considered approach to selection of spatial mean and variance, as discussed in 4.3 (last paragraph) is likely to demonstrate further improvement of the blended analysis.
We stress that this evaluation is preliminary and there is a great deal more work needed for a more thoroughly evaluation of the blending method. For example, we have yet to explore a wider range of performance metrics, including those given in Table 1. It may be that through the ensemble generation process the ensemble mean gives a high POD on average than the individual data, but also leads to a higher FAR. In addition, we need to assess whether the ensemble is indeed an accurate representation of the uncertainty in rainfall estimation. This can be assessed through a variety of measures including spread-skill relationships (Grimit and Mass, 2007), as well as probabilistic metrics such as the Brier score and continuous ranked probability score. Finally, extending the evaluation to longer time series, say a year, will identify seasonal performance which may highlight the well-known complimentary nature in model (reanalysis) and satellite rainfall. Conducting more localised evaluation over selected regions, e.g. mountainous regions, desert area or the tropics, can assess performance over challenging areas.
5 Summary and recommendations

5.1 Summary

We presented a method of blending three sources of real-time and near real-time rainfall data to produce seamless gridded estimates of hourly accumulated rainfall covering large parts of the continent. The method focusses on the data ability to capture spatial structure in rainfall across a wide range of spatial scale. A simple weighting scheme was devised that weighted spatial components of the source data (derived through spectral decomposition) by their respective ability to capture structure based on the mutual correlation between spatial components. Furthermore, the method can produce ensembles of gridded rainfall analysis to represent the uncertainty in rainfall estimation, and where each member possesses the scaling properties of the blended rainfall analysis.

Data

Three sources of rainfall were considered: radar (Rainfields) high resolution (~2-km) mosaic with coverage limited mainly to population centres; NWP (ACCESS-R) forecasts of accumulated rainfall at 11 km resolution; satellite (GPM IMERG 3B-HHR) rain rates summed to give hourly accumulation at 10-km resolution. The method was developed and applied to the month of September 2016 for a 1600 x 1600 km² study area of southeast Australia (a third of the continental land area) with a smaller 512 x 512 km² area embedded where there was maximum coverage of radar data.

Structure

Examination of the power spectral density (PSD) of all three data showed that they were all very good at capturing rainfall spatial structure over a wide range of spatial scales with each having the expected power-law distribution. Satellite and radar rainfall compared most closely in the shape of spectral slopes of their PSDs. There was a recurring feature in the NWP for the spatial scale 25 - 125 km in which the spectral slope was very often higher in absolute value than the other data over this range. This is due to the NWP forecast rainfall possessing a high level of smoothness, which suggests an inability of the data to represent spatial irregularities in the rainfall pattern over this range of scales.

Weights

The method assumes a multiplicative cascade model. As such, spectral decomposition of the data sources was conducted with the aim of deriving weighted a sum of cascade components as a way of blending our rainfall products. Furthermore, the assigned weights correspond to the respective data ability to capture spatial structure at that cascade level (i.e. range of scales). The method calculates weights, based on triple collocation, to each cascade components when all three data are available. Variation of weights across spatial scales exhibited strong correspondence with spatial structure, notably the previously mentioned smoothness.
of NWP forecasts. This meant that the NWP spatial components over that range of scale were consistently down weighted in the blended analysis. More detailed investigation is needed, e.g. through comparison of the method to that of assigning equal weights, to verify if this down-weighting results in increased skill.

Weights exhibited some consistent temporal behaviour, suggesting the possible use of temporally ‘static’ values to blend two rainfall data when one of the three rainfall sources is missing. We envisage three scenarios:

- near real-time applications where, due to the latency of the satellite rainfall estimates (~ 4 hours) we wait until all three data (R+N+S) are present to compute the weights as per the described method;
- real-time, where we use the mean (or median) of the weights for each cascade levels derived from previous (month or so) analyses to blend radar and NWP (R+N); or
- post real-time, where we use climatological weights to blend radar and satellite (R+S) for the purposes of evaluating NWP forecasts.

The preliminary evaluation of the methods showed little difference in performance between the blending based on weights extracted from climatology and those derived from three data simultaneously.

**Evaluation**

An initial comparison of blending results was conducted against independent daily gauge-derived rainfall analyses. The results showed that the most notable improvements in rainfall estimation over the source data alone were observed in the large (1600 x 1600 km$^2$) study area, and especially for the areas with no radar coverage. The comparison also revealed some significant bias in the radar and satellite data, with direct implications on the approach to selection of a reference data set, which requires further investigation and could explain the null result in the smaller (512 x 512 km$^2$) study area.

**5.2 Next steps**

The work to-date has primarily focussed on developing the blending method. Some pragmatic decisions were made around the implementation of the method in an attempt to generate some results in a relatively short time frame. So too, a cursory evaluation of the method was conducted in an attempt to see if the blended results were in the ‘ball park’. There is certainly more work to be done, especially around the method verification, and in the following we itemise a set of next steps, from the immediate tasks to ideas for future development.
Immediate tasks (in order of priority):

- Assessing the impact of a revised structured noise perturbation scheme on blended ensemble means, especially in areas of transition from radar coverage to no radar coverage.
- Extend the verification statistics to include metrics on detection (categorical statistics) and skill scores related to how well the ensembles represent uncertainty.
- Conduct localised verification targeting (i) coast zones with high levels of gauge coverage; (ii) inland areas with poor or no radar coverage; and (iii) mountainous areas that pose challenges to radar and satellite retrievals.
- Experiment with temporally static weights versus instantaneously or ‘dynamically’-derived weights to assess the expected difference in performance in real-time and near real-time settings.

Further development, particularly ahead of national-scale implementation:

- Explore the impact of non-stationarity on the method and consider the approaches of concatenating a series of localised analyses, or using spatially varying weights in the blending procedure.
- Incorporate the result of independent evaluation of the source data though either an \textit{a priori} bias correction or guidance on the appropriate use choice of data to serve as the reference of each cascade level.
6 References


IAC (2010), Seventh Report of the Inter-Agency Committee on the Hydrological Use of Weather Radar. CEH Wallingford, United Kingdom


