26th International Congress on Modelling and Simulation, Adelaide, SA, Australia, 30 Nov–4 Dec 2025
mssanz.org.au/modsim2025

Development of an Australia-wide extreme storms database for hydrologic risk assessments

C. Dykman[1], Y. Kim[2], R. Nathan[3], A. Sharma[1]

[1]*School of Civil Engineering, University of New South Wales, Kensington, New South Wales, Australia*
[2]*Climate Change Research Centre, University of New South Wales, Kensington, New South Wales, Australia*
[3]*Department of Infrastructure Engineering, The University of Melbourne, Parkville, Victoria, Australia*
Email: c.dykman@unsw.edu.au

Abstract: It is now well understood that anthropogenic induced global warming is increasing extreme rainfalls, with the more extreme the rainfall, the greater the intensification. This in turn is increasing the magnitude of rare floods. Less well understood, however, is the changing impact of spatial and temporal patterns of extreme rainfall on flooding under a warming climate. The spatial and temporal distribution of rainfalls during storm events has a significant influence on flood peaks, runoff volumes and water availability. Hence, robust datasets are required to model hydrologic risk with changes in storm spatial and temporal patterns.

We have developed Australia-Wide Extreme Storms Database (or AWESD) which characterises storm patterns with high spatial (12km) and temporal (hourly) resolution for hydrologic risk assessments. Whilst the record length for such high-resolution data is currently 30 years, the availability of information at a high resolution over large homogeneous regions allows the trading of space for time, which has the potential to provide equivalent independent record lengths that are much longer than 30 years. To develop such a database, we first identify and track storms using two data sets: a low resolution (daily) gridded dataset based on observations, and a higher resolution (hourly) reanalysis dataset. Identified storms are then filtered to ensure they are independent in both space and time. Storms identified in the high-resolution reanalysis dataset are checked for consistency with the observation-based data set to ensure a grounding in reality. Subsequently, we have created a software for storm selection. Storms are selected based on an input catchment location plus a prespecified buffer region and within a range of prespecified ratios of catchment size. Storms are then transposed to the catchment centre. The final storms selection can then be formatted to facilitate input to event-based flood modelling software.

The development of the extreme storms database and storm selection software facilitates the undertaking of hydrologic risk assessments by sampling storms on depth/rarity, spatial homogeneity and temporal homogeneity. Furthermore, it can be used for climate impact assessments by sampling storms based on characteristics associated with a warmer climate e.g. higher depths and shifting spatio-temporal pattern distributions. With this storm database we believe it will enhance hydrological risk assessments performed for both present and future climate scenarios and deepen understanding of the role of spatio-temporal distributions on extremes.

Keywords: *Precipitation, extremes, spatio-temporal patterns, ensembles, datasets*

1. INTRODUCTION

A design storm typically consists of a design storm burst depth, for a specified duration, a temporal pattern, that enables detailing it over the duration considered, and an indicator of expected wetness conditions prior to the storm burst. Given the uncertainty associated with the specification of a design storm, the Australian Rainfall and Runoff (ARR) recommends multiple temporal patterns be assessed while using the design storm for derived analyses. In addition to multiple temporal patterns, in applications over larger and more complex catchments, spatial patterns can also be used. As the storm increases in rarity, defining all aspects (depth, spatial and temporal patterns, and pre-storm characteristics) becomes more challenging. Doing so for extreme design storms approaching the Probable Maximum Precipitation poses perhaps the biggest challenge, as storm events are unlikely to ever be a part of the historical record.

Currently, guidance for estimating the probable maximum precipitation from the Australian Bureau of Meteorology (BoM) prescribes generalised methods that rely on storm standardisation and transposition (B.o.M 2003, 2005, 2006). There are several limitations to this method. To begin with, to obtain large enough datasets that are likely to include extreme (outlier) storms, data must be taken across very large regions with, what are considered to be, consistent hydro-meteorological conditions. However, because of the requirement for storms to cover a range of durations and areal extents from all geographic areas, many storms in the dataset are not truly extreme events. For example, the Average Recurrence Interval (ARI) for the 110 storms in the Generalised Southeast Australia Method (GSAM) database is about 50 years (Minty and Meighen 1996; Meighen and Minty 1998).

Moreover, for storms to be able to be transposed to any point in the hydro-meteorologically consistent region, site specific storm characteristics must be separated from those storm characteristics that are consistent with the region's meteorological conditions i.e. generalised. Site specific storm characteristics such as moisture inflow direction, and storm movement, i.e. spatial and temporal patterns, are effectively removed during the generalising process when quantifying depth-duration-area curves. As such storms developed with generalised PMP methods are no longer identifiable with any one real event, and spatial and temporal distributions of the PMP depths must be designed. Design spatial patterns only infer the most probable distribution of the topographic component of rainfall within the catchment as knowledge about the spatial distribution of the convergence component of the rainfall within extreme storms has been removed. Design Temporal distributions are created using the Average Variability Method (Pilgrim et al. 1969) and smoothed by either selecting a subsample of the top 10 storms (Green et al. 2003) or using Nathan (1992). The aim of the AVM is to produce a single design temporal pattern with an average variability in rainfall intensity such that the AEP of the rainfall is maintained in the resultant flood i.e. probability neutral. However, it is not representative of any real storm's temporal pattern. Furthermore, the number of individual temporal distributions used to derive design (AVM) temporal distributions at each standard area and burst duration varies greatly e.g. using GSAM approximately 40 storms are available for durations ≤ 24hrs and standard areas ≤ 10,000 km2, while < 3 are available for durations ≥ 96hrs and standard areas ≥ 40,000 km2.

These limitations, amongst other requirements, such as the need to incorporate recent storm data, motivate the need for development of an Australia-Wide Extreme Storms Database (AWESD) utilising modern high-resolution datasets, moving away from the rain gauge density limitations imposed on ARR/BoM procedures. Modern, gridded datasets with high spatial resolution create extended equivalent independent record lengths (EIRL), despite typically having shorter record lengths, by trading space for time when undertaking regional analysis. Therefore, storm data can be gathered from more relevant, localised regions around a catchment of interest. Additionally, sub-daily resolution enables detailed temporal distribution of rainfall events. Herein, we detail the development of the AWES database built upon high resolution reanalysis data that is ground-truthed with gridded observation data.

The development of the extreme storms database and storm selection software facilitates the undertaking of hydrologic risk assessments as storms may be sampled on depth/rarity, spatial homogeneity and temporal homogeneity. For example, it can be used to investigate how spatial and temporal patterns of rainfall may vary with event severity, and this could be used to inform estimates of dam failure risks. Furthermore, it can be used for climate impact assessments by sampling storms based on characteristics associated with a warmer climate e.g. higher depths and shifting spatio-temporal pattern distributions.

2. COMPILATION OF EXTREME STORMS

2.1. Datasets

There are two principal datasets utilised for the development of the AWESD, namely the Australian Gridded Climate Data version 1 (AGCDv1), previously published under the Australian Water Availability Project (AWAP) (Jones et al. 2009) and the Bureau's Atmospheric high-resolution Regional Reanalysis for Australia, known as BARRA-Rv1.0 (Su et al. 2019). The motivation for using two datasets is to exploit the high resolution of reanalysis data (BARRA-Rv1.0) while ground-truthing with lower resolution observation-based data (AGCDv1). AGCDv1 precipitation analysis data is derived from high quality in situ observations of rainfall gauges managed by the Australian Bureau of Meteorology. Robust topography resolving analysis methods have been used to produce gridded data at spatial resolutions of 0.05 x 0.05 degrees (~5km x 5km). Further details of the analysis can be found in Jones et al. (2009). BARRA-Rv1.0 is an atmospheric regional reanalysis product covering a large region over Australia, New Zealand and Southeast Asia with horizontal resolution of approximation 12km (BARRA-Rv1.0). BARRA-Rv1.0 utilises the Australian Community Climate Earth-System Simulator (ACCESS) computer weather model and provides reanalysis products for a 29-year period spanning 1990-2019 at an hourly resolution across 70 vertical levels up to 80 km into the atmosphere. A detailed description of the reanalysis can be found at Su et al. (2019). A summary of the advantages and disadvantages of each dataset is shown in Table 1.

Table 1: Dataset Advantages and Disadvantages

Dataset	Advantages	Disadvantages
AGCDv1	• High quality historical observations • Long record (>100 years)	• Areas of limited station density • Temporal resolution restricted to daily
BARRA-Rv1.0	• High Temporal resolution of 1 hour • High Spatial resolution 12 km	• Biases associated with model parametrisation. • Short record (<30 years)

2.2. Storm identification

A Storm identification procedure is undertaken using both AGCDv1 and BARRA-Rv1.0 datasets. The motivation being that storms identified in the BARRA-Rv1.0 dataset that are also identified (within space and time thresholds) in the observation based AGCDv1 dataset are more likely to represent real storms. Thus, the detailed hourly outputs available from BARRA-Rv1.0 can be used with more confidence. The conditional storm identification procedure is summarised below:

1. Identify storms using AGCDv1 data.
2. Identify storms using BARRA-Rv1.0 data.
3. Conditional selection of storms from BARRA-Rv1.0 within space and time thresholds of those in AGCDv1. (Space and time thresholds are user inputs, however default values used are 100km and 1 day).
4. Extract hourly storm characteristics from BARRA-Rv1.0.

Extreme storm identification is carried out in two parts: determining independence in time and independence in space. To achieve independence in time the highest independent six events per year at each grid cell of the datasets are identified as follows:

1. Take 95th percentile at each year.
2. Apply pseudo independent in time:

$$X_t : (X_{t-1} = 0 \land X_{t+1} = 0) \lor (X_t > X_{t-1} \land X_t > X_{t+1}) \quad 1$$

3. Take the highest six events.

For each of the independent in time events, independence in space is determined by grouping nearby regions into "rainstorm segments" using a technique based on an algorithm termed "almost-connected component labelling" (Chang et al. 2016). The procedure can be summarised in four steps

1. Identify contiguous regions of rainfall above a predefined threshold.
2. Form segments using almost-connected-component labelling for only the large contiguous regions from step 1.
3. Assign small regions to segments from Step 2 if that are sufficiently close.
4. Form segments with remaining small regions.

2.3. Storm tracking

Once storms have been identified, they are tracked across time steps using a method called "storm tracking and regional characterization" (STARCH) that utilises the "overlapping ratio method" (Liu and Wright 2022). The fundamental premise of the overlapping ratio method is that if two storm objects at t_i and t_{i+1} have a sufficiently high ratio of overlapping area then they are deemed to be a single storm.

2.4. Storm database

Extreme storms identified and tracked using the aforementioned procedures are characterised and compiled into an extreme storms database. Attributes are calculated for each storm and include but are not limited to storm centre location, timestamp, storm area, average intensity, maximum intensity and duration. An extreme storms database is compiled for both the AGCDv1 and BARRA-Rv1.0 datasets. A final AGCDv1-conditioned-BARRA-Rv1.0 database is compiled from storms identified in BARRA Rv1.0 within defined distance and time thresholds of AGCDv1 storms (default 1 day and 100 km).

3. STORM SELECTION

Having compiled an Australia-wide extreme storms database, a storm selection software was developed for implementation at the catchment scale.

3.1. Identify target storms

Target storms for a catchment of interest are identified by using the catchment boundary (in shapefile format) plus a user specified buffer (default 1 degree). This will identify all storms in the extreme storms database whose centre falls within the catchment boundary plus buffer. To ensure a suitable balance between regional hydroclimate similarity and event rarity, it is recommended that a catchment buffer is specified such that a minimum of 100 storms are selected. An example of storms selected for a particular catchment boundary is shown in Figure 1.

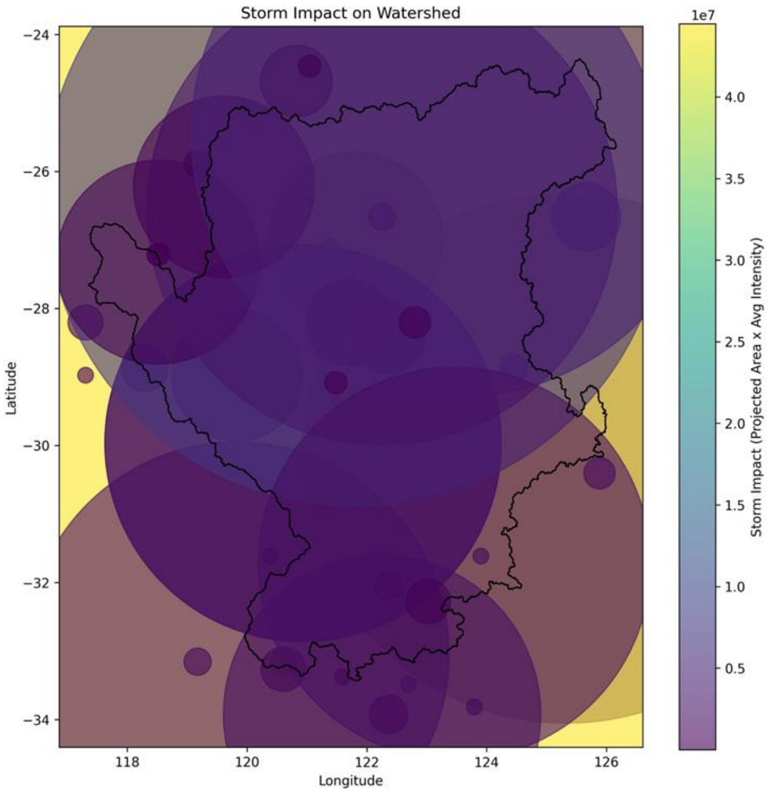

Figure 1: Selected Storms for Example Catchment Boundary. The centre of each circle represents the storm centre with the radius proportional to the storm area. The colour of each circle is related to the volume of the storm. NB: some circles are so large they are outside the boundaries; however, this image is for indicative purposes only.

3.2. Conditional selection

Selection of the identified target storms can be further conditioned. To ensure close to full storm coverage of the target catchment, storm selection can be conditioned on storm area (default 1-3x catchment area). Furthermore, to prevent selection of highly localised storms, selection can be conditioned on a rain intensity ratio where the ratio of storm maximum intensity to average intensity does not exceed a defined threshold (default 10). Finally, target storms may also be conditionally selected on volume/rarity spatial homogeneity (Ghanghas et al. 2023) and temporal homogeneity (Visser et al. 2023).

3.3. Obtain data

The dates and locations of identified and conditionally selected target storms are then used to download and select the relevant BARRA-RV1.0 files from the National Computational Infrastructure (NCI) database. Storm data is then transposed so that all storms are centred on the catchment centre (Scorah et al. 2021). Storm data is normalised such that all spatial and temporal data are in the form of percentages. Spatio-temporal patterns are saved as netCDF (.nc) files and lumped storm, burst and pre-bust temporal patterns are saved as CSV (.csv) files.

4. STORM DATA PROCESSING

4.1. Pre-processing

With the objective of analysing extreme storms, storm pre-processing is undertaken with the following knowledge:

1. Generalised Probable Maximum Precipitation (PMP) methods assume full catchment coverage for a storm.
2. As storms become more extreme, temporal patterns are likely to be more uniform (Scorah et al. 2025).

Thus, to prevent selected extreme storms from having uncharacteristically high variability across spatial patterns and or temporal patterns, that are likely associated with storms of higher frequency, both a spatial and temporal smoothing process is undertaken. It is worth noting however that by conditioning selection of storms on area, rain intensity ratio and rarity most storms of high spatial and temporal variability are likely to be filtered out prior to this stage.

Spatial smoothing

If an event has a subarea with weighting greater than a specified threshold (default 200%) then spatial smoothing is applied. A default of 200% was chosen as both the Generalised South-east Australia Method (GSAM) (B.o.M 2006) and Generalised Tropical Storm Method (GTSMR) (B.o.M 2005) for PMP spatial patterns use the function-modified ratios of the intensity frequency-duration (IFD) for a 72-hour, 50-year ARI (Average Recurrence Interval) to the corresponding flat-land IFD over the catchment. This ratio has a lower limit of 1 and an upper limit of 2, meaning that the theoretical maximum distribution any subarea could have is <200%. The spatial pattern is scaled such that the maximum percentage equals 100% for any subarea and the remaining weight of the spatial pattern is added from the PMP spatial pattern; either GSAM or GTSMR. E.g.

$$SP_{smooth} = [SP_{BARRA} \times w] + [SP_{PMP} \times (1 - w)] \qquad 2$$

Where SP_{BARRA} is the raw spatial pattern from the BARRA-Rv1.0 data, SP_{PMP} is the spatial pattern derived from either GSAM or GTSMR, w is the weighting factor, and SP_{smooth} is the resulting smoothed spatial pattern.

Temporal smoothing

If an event has a temporal pattern with a D50 (the percent of time required to achieve 50% of event rainfall volume) outside a specified range (default 40%-60%) then temporal smoothing is applied. A default range of 40%-60% was chosen as a uniform temporal pattern will have a D50 = 50%, and a D50 ≤ 40% is generally considered front loaded while a D50 ≥ 60% is generally considered back loaded. The event D50 is scaled to the D50 of the PMP Average Variability Method (AVM) temporal pattern for the specific duration and the remaining weight of the temporal pattern is added from the PMP AVM temporal pattern for the specific duration; either GSAM or GTSMR. E.g.

$$if\ D_{50} < 40: w = 1 - \frac{50}{D_{50}} \qquad 3$$

$$if\ D_{50} > 60: w = \frac{50}{D_{50}} \qquad 4$$

$$TP_{smooth} = [TP_{Barra} \times w] + \left[TP_{pmp} \times (1 - w)\right] \qquad 5$$

Where TP_{Barra} is the raw temporal pattern from the BARRA-Rv1.0 data, TP_{pmp} is the AVM temporal pattern derived from either GSAM or GTSMR, w is the weighting factor, and TP_{smooth} is the resulting smoothed temporal pattern.

4.2. Formatting

Using the storm selection software, BARRA-Rv1.0 storm pattern data can be formatted for further modelling in three primary ways as outlined below. Pattern nomenclature used below is specific to the RORB software, however, the general principles are adaptable to any rainfall-runoff, runoff-routing or hydrodynamic model that takes rainfall as an input.

Lumped burst patterns

The first option is lumped burst patterns (refer Figure 2a). These patterns are calculated by 'lumping' or adding event rainfall bursts across the catchment at each time step thus aggregating into a single temporal pattern to be applied across the whole catchment. The spatial pattern is calculated by finding the proportion of total catchment burst rainfall that each 'subarea' is responsible for.

Lumped storm patterns

The second option is lumped storm patterns (refer Figure 2b). These are as per lumped burst temporal patterns, however, with the addition of a storm specific pre-burst aggregated at the catchment scale. The spatial pattern is calculated by finding the proportion of total catchment storm (pre-burst + burst) rainfall that each 'subarea' is responsible for.

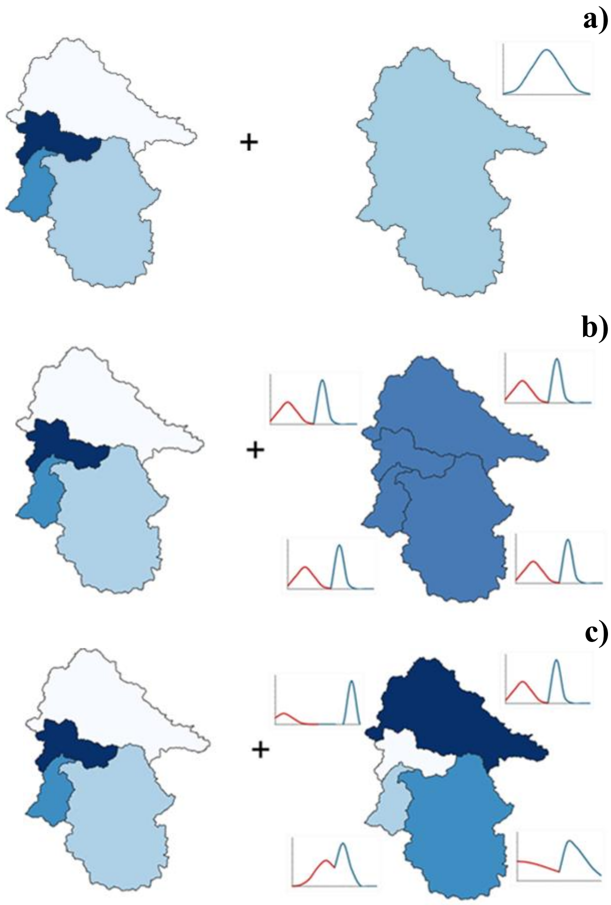

Figure 2: Spatial (left) and temporal (right) pattern formatting. Panel a) illustrates a lumped *burst* temporal pattern, b) illustrates a lumped *storm* temporal pattern and c) illustrates a distributed *storm* temporal pattern or storm space-time pattern.

Space-time patterns

The third option is storm space-time, or spatio-temporal, patterns (refer Figure 2c). These patterns are calculated by aggregating storm rainfall at the 'subarea' level such that each 'subarea' has a unique storm (pre-burst +burst) temporal pattern. The spatial pattern is as per lumped storm patterns.

5. CONCLUSION

Herein, we have detailed the development of the AWESD built upon the Australian Bureau of Meteorology's high-resolution BARRA-Rv1.0 reanalysis dataset and ground-truthed with AGCDv1 gridded observation-based precipitation data. The use of high-resolution gridded datasets creates extended equivalent independent record lengths with greater detailing of temporal patterns, allowing storm data to be used from more relevant, localised regions around a catchment of interest. This is a step towards moving away from the rain gauge density restrictions imposed on ARR/BoM procedures. Whilst it is acknowledged that the depths in reanalysis and observation datasets may differ, BARRA-R has been shown to have broad agreement with AGCDv1 (Su et al. 2019; Su et al. 2021). Incorporating newer datasets such as BARRA2 (Su et al. 2022a) and BARPA (Su et al. 2022b) for future climate projections in subsequent studies would further enhance the applicability and robustness of the developed software. The development of the AWESD and the associated storm selection software improves confidence in sampling extreme storm patterns with opportunities for use in flood studies, impact assessments, probabilistic rick assessment, dam breach assessments, and climate change studies.

ACKNOWLEDGMENTS

This work was funded by WaterNSW. The authors acknowledge input from the Supply and Forecast Modelling team at WaterNSW including but not limited to: Narendra Tuteja, Angelo Breda, Ha Nguyen, Mark Powell, Ashok Raut, Raj Upreti, Yiliang Du and Lavendini Sivaneasharajah.

REFERENCES

B.o.M (2003) Guidebook to the Estimation of Probable Maximum Precipitation: Generalised Short-Duration Method. In: Bureau of Meteorology (ed.).

B.o.M (2005) Guidebook to the Estimation of Probable Maximum Precipitation: Generalised Tropical Storm Method. In: Bureau of Meteorology (ed.).

B.o.M (2006) Guidebook to the Estimation of Probable Maximum Precipitation: Generalised Southeast Australia Method. In: Bureau of Meteorology (ed.).

Chang W, Stein ML, Wang J, Kotamarthi VR and Moyer EJ (2016) Changes in Spatiotemporal Precipitation Patterns in Changing Climate Conditions. Journal of Climate 29(23), 8355-8376. Doi: 10.1175/jcli-d-15-0844.1.

Ghanghas A, Sharma A, Dey S and Merwade V (2023) How Is Spatial Homogeneity in Precipitation Extremes Changing Globally? Geophysical Research Letters 50(16). Doi: 10.1029/2023gl103233.

Green J, Walland D, Nandakumar N and Nathan R (2003) Temporal Patterns for the Derivation of PMPDF and PMF Estimates in the GTSM Region of Australia. 28th International Hydrology and Water Resources Symposium: About Water; Symposium Proceedings. Institution of Engineers, Australia, Barton, A.C.T.

Jones D, Wang W and Fawcett R (2009) High-quality spatial climate data-sets for Australia. Australian Meteorological and Oceanographic Journal 58. Doi: 10.22499/2.5804.003.

Liu Y and Wright DB (2022) A storm-centered multivariate modeling of extreme precipitation frequency based on atmospheric water balance. Hydrology and Earth System Sciences 26(20), 5241-5267. Doi: 10.5194/hess-26-5241-2022.

Meighen J and Minty LJ (1998) Temporal Distributions of Large and Extreme Design Rainfall Bursts Over Southeast Australia HRS Report No. 5. Bureau of Meteorology, Melbourne, Australia.

Minty LJ and Meighen J (1996) Development of the Generalised Southeast Australia Method for Estimating Probable Maximum Precipitation, HRS Report No. 4. Bureau of Meteorology, Melbourne, Australia.

Nathan R (1992) The Derivation of Design Temporal Patterns for Use with the Generalised Estimates of Probable Maximum Precipitation. Civil Engineering Transactions, Institution of Engineers, Australia 34(2), 139-150.

Pilgrim DH, Cordery I and French R (1969) Temporal Patterns of Design Rainfall for Sydney. Civil Engineering Transactions, Institution of Engineers, Australia 11, 9-14.

Scorah M, Nathan R and Lang S (2021) Incorporating stochastic variation of storm space time pattern centring in a Monte Carlo hydrologic model. In: HWRS 2021: Digital Water: Hydrology and Water Resources Symposium 2021. Engineers Australia. Viewed 2025/04/29, https://search.informit.org/doi/10.3316/informit.338574053169070.

Scorah M, Nathan R and Wasko C (2025) Rainfall events become temporally more uniform with increasing storm depth. Journal of Hydrology 657, 133101. Doi: https://doi.org/10.1016/j.jhydrol.2025.133101.

Su CH, Eizenberg N, Steinle P, Jakob D, Fox-Hughes P, White CJ, Rennie S, Franklin C, Dharssi I and Zhu H (2019) BARRA v1.0: the Bureau of Meteorology Atmospheric high-resolution Regional Reanalysis for Australia. Geosci. Model Dev. 12(5), 2049-2068. Doi: 10.5194/gmd-12-2049-2019.

Su CH, Ye H, Dowdy AJ, Pepler A, Stassen C, Brown A, Tucker SO and Steinle PJ (2021) Towards ACCESS-based regional climate projections for Australia. Australian Bureau of Meteorology, Melbourne, Australia.

Su CH, Rennie S, Dharssi I, Torrance J, Smith A, Le T, Steinle P, Stassen C, Warren RA, Wang C and Le Marshall JF (2022a) BARRA2: Development of the next-generation Australian regional atmospheric reanalysis. Bureau of Meteorology.

Su CH, Stassen C, Howard E, Ye H, Bell S, Pepler A, Dowdy AJ, Tucker SO and Franklin C (2022b) BARPA: New development of ACCESS-based regional climate modelling for Australian Climate Service. Australian Bureau of Meteorology.

Visser JB, Wasko C, Sharma A and Nathan R (2023) Changing storm temporal patterns with increasing temperatures across Australia. Journal of Climate, 1-26. Doi: 10.1175/jcli-d-22-0694.1.

