Proc. 3rd International Conference on Water and Environmental Engineering (iCWEE2022)
27-30 Nov 2022, Sydney, Australia
ISBN: 978-0-6456692-0-6

Updating Australian Rainfall and Runoff Regional Flood Estimation Database till 2022

Xiao Pan[1], Ataur Rahman[2], Khaled Haddad[3] and Mohammed A Alim[4]

[1]PhD student, School of Engineering, Design and Built Environment, Western Sydney University, Australia

[2]Professor, School of Engineering, Design and Built Environment, Western Sydney University, Australia

[3]Manager, Research Reforms, Department of Education, NSW, Australia

[4]Academic Course Adviser, School of Engineering, Design and Built Environment, Australia

Corresponding author's E-mail: Tony.xiao.pan@gmail.com

Abstract

Flood is one of the most destructive and costly natural disasters in Australia. It often results in a heavy economic burden due to managing the damaged infrastructures afterward and disruption of day-to-day life. Hydrological studies on flood-prone areas are conducted to decrease the consequences of floods. In hydrological studies, estimation of design flood is needed. In this regard, the time series of the streamflow discharge dataset is the primary source and essential for fitting the probability distribution for flood frequency analysis (FFA). At present, the Australian Rainfall and Runoff recommended regional flood estimation method called RFFE Model that used annual maximum flood data till 2011, which is 10 years old as of 2022. This study extracted the up-to-date daily maximum discharge records from state authorities across Australia and used them for the construction of the annual maximum time series up to the year 2022. A total of 747 gauging station discharge records in Australia are cross-checked, with the data extracted up to 2022 for validation purposes. It is expected that this study will benefit both research and industry practices. An extraction routine is also built using the modern computational science language, R, for future data extraction and cross-validation deployment. The proposed extraction routine using R is also designed to suit different purposes based on the requirements, such as constrained by time interval, magnitude, and total counts of peaks.

Keywords: *Flood database, annual maximum flood, floods, R, Australia.*

1. INTRODUCTION

Floods are one of the worst natural disasters, which results in significant economic burdens due to managing the consequences of flooding (Acosta et al., 2016). To minimise the consequences of flooding, hydrologists and engineers must estimate the design floods accurately by conducting hydrological studies in the flood-prone area (Haddad et al., 2010; Karim et al., 2017; Pan & Rahman, 2022). Design flood is often estimated through at-site flood frequency analysis (FFA) adopting annual maximum flood (AMF), which consists of one highest discharge value per year from streamflow records (Haddad & Rahman, 2011, 2015). Following the formation of the AMF, the probability distribution function is then fitted to the extracted AMF.

Streamflow records and extracted flood series (based on AMF or other approaches) are the foundation of FFA. However, the shortage of the streamflow record is one of the most raised concerns by hydrologists

(Haddad et al., 2010). Traditionally, a 2T rule is recommended for the at-site FFA (i.e., 100 years streamflow record is needed for estimating a 50-year design flood). However, most of the streamflow gauging stations in Australia do not satisfy this requirement for a higher return period. In this regard, the regional flood frequency estimation (RFFE) is developed to overcome the data shortage issue. RFFE aims to transfer flood information from data-rich sites to data-poor sites based on linear (quantile regression) (Rahman et al., 2010) or nonlinear assumptions (generalised additive model) (Rahman et al., 2018). Nonetheless, the accuracy of the estimated flood quantile (based on various RFFE methods) is highly dependent on the constructed flood series derived from the primary streamflow record (Noor & Rahman, 2016). Currently, Australian Rainfall and Runoff (ARR) employs an RFFE method based on extracted AMF till 2011 (AMF2011), which is 10 years old in 2022.

This study takes the initiative to construct the new AMF database up to 2022 (AMF2022). In this context, the purposes of this study are: (a) to develop a systematic extraction and validation routine for catchments/sites adopted in the ARR RFFE module; (b) to update the current AMF2011 database to AMF2022 where possible; (c) Investigate and compare the flood database between AMF2011 and AMF2022. It is expected that the outcome of this study will benefit both research and industry practice.

2. METHODS

The adopted methodology of this study is illustrated in Figure 1. Firstly, the complete list of catchments currently employed in ARR based on AMF2011 is obtained and tabulated with geographical locations (longitude and latitude). After the list of catchments is constructed, maximum daily streamflow records up to 2022 are obtained through the state’s water authority portal. For each station listed in the AMF2011 database, a new AMF flood series is constructed by selecting one maximum discharge value per year from the primary daily maximum streamflow record up to 2022. During the iterative process of selecting gauged stations, there may be scenarios where a particular station cannot be found anymore in the active station list (closure of station). In this case, the AMF remains the same as per the AMF2011 database. For other stations, it is required to undergo a validation process.

The validation process aims to ensure the quality of the newly constructed AMF database. It is implemented using modern computational science language, R. Considering the future use of AMF, a systematic programming routine must be developed. In other words, the developed R script must be reusable for future AMF construction by adopting a parametric input/output coding structure. The criteria employed to construct the AMF2011 database by Haddad et al. (2010) are firstly programmed, with the addition of the below criteria and assumptions.

- Regardless of the start year of the streamflow record, any of AMF2022 must have the same start year as AMF2011.
- For AMF data value up to 2011, the absolute relative error in percentage for any given year must be less than or equal to 5%. Maximum discharge value should be taken (between AMF2011 and AMF2022) for any given year if the absolute relative error is between 5% and 10%.
- Considering any AMF value between 2011 and 2022, the gap in the streamflow record needs to be identified and logged for manual processing.
- If the continuous difference between AMF2011 and AMF2022 (absolute relative error greater than 10%) is observed before 2011, the AMF2011 should replace the value with a continuous difference.

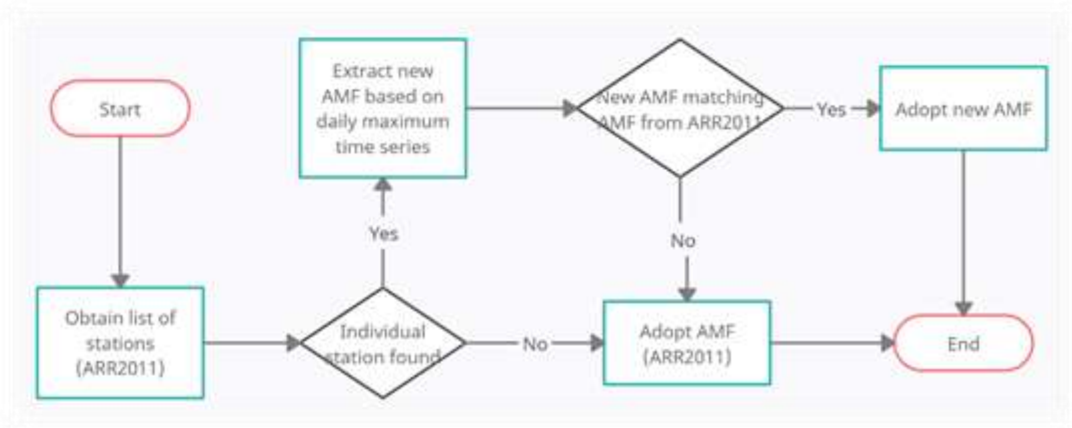

Figure 1 Illustration of adopted methodology updating ARR RFFE flood database

3. DATA

This study checks and validates 746 catchments' AMF data, which is listed in AMF2011, and plotted the location of the catchment in Fig. 2. The selected catchments cover all states of Australia except for Tasmania (currently work in progress). Most of the selected catchments (75% of the total) are located along Australia's northeastern and southeastern coastal lines. These catchments are from the states of New South Wales (NSW), Victoria (VIC), and Queensland (QLD) and will be further investigated in a later section. Overall, 535 catchments are validated and updated in the AMF2022 database, which accounts for 72% of the total catchments listed in ARR.

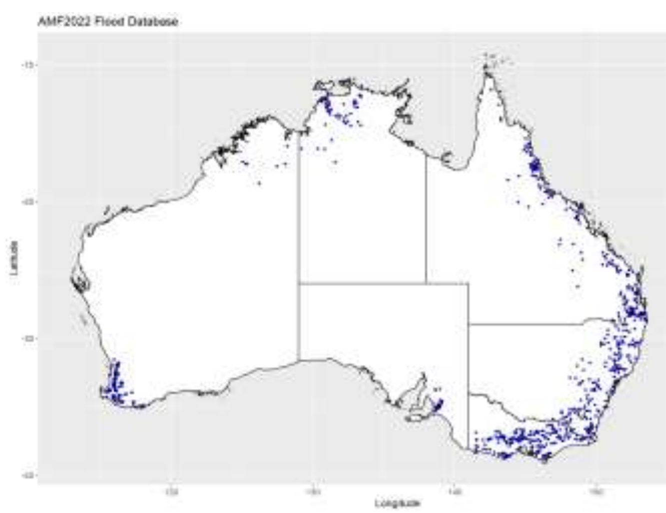

Figure 2 Geographical location catchments using AMF2011

4. RESULTS AND DISCUSSION

Table 1 summarises the comparison of descriptive statistics between the AMF2011 and AMF2022 databases. It is observed that there is a significant increase in the record length among most of the states in Australia. For example, NSW increases from 36 years of flood record length to 44 years on average while having a median of 44 years (AMF2022 database).

Table 1 Descriptive statistics of comparison between AMF2011 and AMF2022 databases

State	AMF up to	Count of catchments	Minimum year of record	Maximum year of record	Mean year of record	Standard deviation (year of record)	Median year of record
New South Wales	2022	176	20	93	44	15	44
	2011		20	82	36	12	34
Northern Territory	2022	50	19	57	38	13	42
	2011		19	69	42	17	45
Queensland	2022	195	20	113	51	19	51
	2011		20	102	43	17	41
South Australia	2022	28	20	63	37	9	37
	2011		28	54	43	9	46
Victoria	2022	186	20	71	46	9	48
	2011		20	60	37	7	38
Western Australia	2022	111	20	60	32	10	30
	2011		20	71	38	13	36

This study further examines the impact of updating the AMF2022 database along the eastern coastal line (states of NSW, VIC and QLD in Australia), which accounts for most of the AMF2011 database. A total of 89 basins are tabulated based on the longest record length within the same basin (i.e., only one catchment is selected in the same basin with the longest record length and AMF updated to 2022). After the catchment is selected for each basin, the peak flow value is classed based on bin width of 18 years and plotted in Figure 3. It is observed that 19 out of 89 basins (21%) have the highest discharge occurring after 2011, as shown in Figure 3.

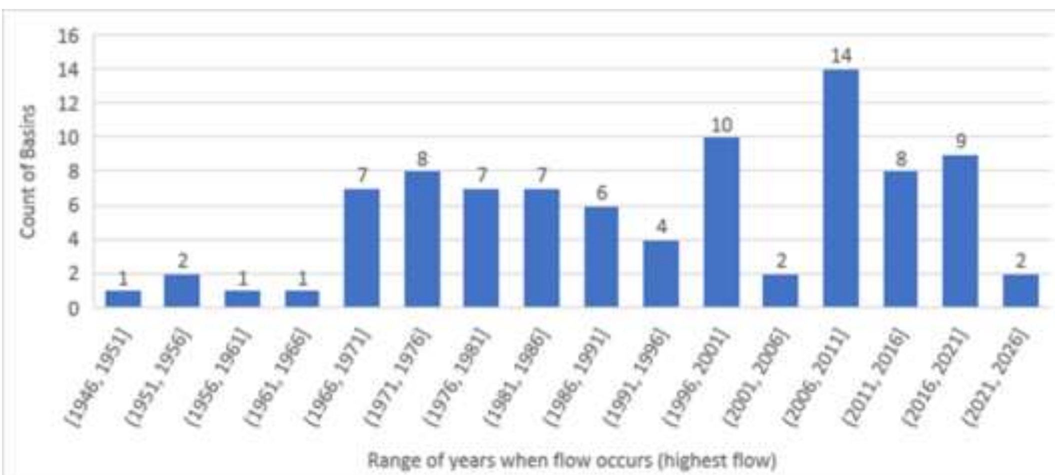

Figure 3 Histogram of peak flow occurs (highest flow of selected basins)

Figure 4 plots the geographical location of basins in red where the highest peak flow occurs after 2011. Most of the basins with the highest peak flow occurring after 2011 are located along the coastal line of Australia, and only a few basins are further inland. However, there is no spatial coherence observed in Figure 4.

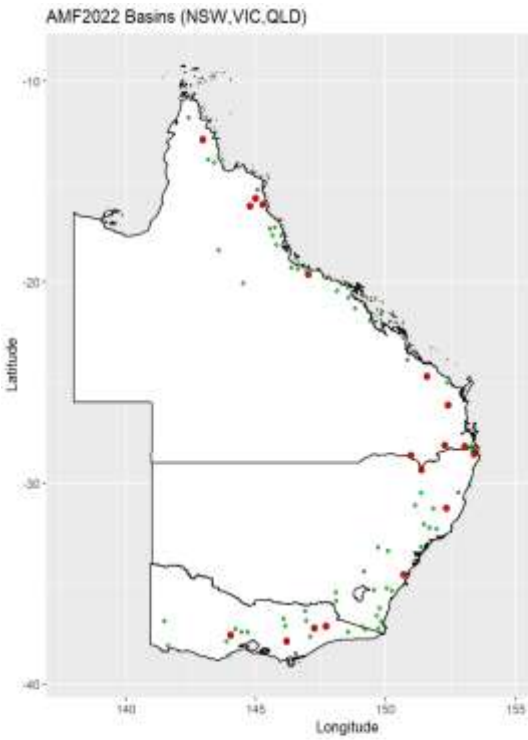

Figure 4 Geographical location of the highest peak flow occurs (red: highest peak flow occurs after 2011; Green: highest peak flow occurs before 2011)

5. SUMMARY

The study examines the current AMF2011 database in ARR and takes the initiative to update the existing AMF2011 database till 2022. A total of 747 catchments listed in the AMF2011 database are examined, and 535 catchments' AMF (75% of total catchments) are updated till 2022. This study also investigates the eastern coastal line of Australia regarding the impact of updated AMF2022. It is observed that 21% of selected 89 basins have the highest peak flow occurring after 2011, which was not considered in the previous AMF2011. It is also noted that the mentioned basins are mainly located along the coastal area of Australia, and no spatial coherence is identified.

REFERENCES

Acosta, L. A., Eugenio, E. A., Macandog, P. B. M., Magcale-Macandog, D. B., Lin, E. K. H., Abucay, E. R., Cura, A. L., & Primavera, M. G. (2016). Loss and damage from typhoon-induced floods and landslides in the Philippines: Community perceptions on climate impacts and adaptation options [Article]. *International Journal of Global Warming*, *9*(1), 33-65. https://doi.org/10.1504/IJGW.2016.074307

Haddad, K., & Rahman, A. (2011). Selection of the best fit flood frequency distribution and parameter estimation procedure: A case study for Tasmania in Australia [Article]. *Stochastic Environmental Research and Risk Assessment*, *25*(3), 415-428. https://doi.org/10.1007/s00477-010-0412-1

Haddad, K., & Rahman, A. (2015). Estimation of large to extreme floods using a regionalization model. In *Landscape Dynamics, Soils and Hydrological Processes in Varied Climates* (pp. 279-292). https://doi.org/10.1007/978-3-319-18787-7_14

Haddad, K., Rahman, A., & Weinmann, P. E. (2010). Estimation of major floods: Applicability of a simple probabilistic model [Article]. *Australian Journal of Water Resources*, *14*(2), 117-126. https://www.scopus.com/inward/record.uri?eid=2-s2.0-79952590865&partnerID=40&md5=60effb7ffb3be7b8a389e01fc79c638b

Haddad, K., Rahman, A., Weinmann, P. E., Kuczera, G., & Ball, J. (2010). Streamf ow data preparation for regional food frequency analysis: Lessons from southeast Australia [Article]. *Australian Journal of Water Resources*, *14*(1), 17-32. https://doi.org/10.1080/13241583.2010.11465371

Karim, F., Hasan, M., & Marvanek, S. (2017). Evaluating annual maximum and partial duration series for estimating frequency of small magnitude floods [Article]. *Water (Switzerland)*, *9*(7), Article 481. https://doi.org/10.3390/w9070481

Noor, S., & Rahman, A. (2016). Regional flood modelling in Tasmania: Comparison of ARR RFFE model and kriging. 37th Hydrology and Water Resources Symposium 2016: Water, Infrastructure and the Environment, HWRS 2016,

Pan, X., & Rahman, A. (2022). Comparison of annual maximum and peaks-over-threshold methods with automated threshold selection in flood frequency analysis: a case study for Australia. *Natural Hazards*, *111*(2), 1219-1244. https://doi.org/10.1007/s11069-021-05092-y

Rahman, A., Charron, C., Ouarda, T. B. M. J., & Chebana, F. (2018). Development of regional flood frequency analysis techniques using generalized additive models for Australia [Article]. *Stochastic Environmental Research and Risk Assessment*, *32*(1), 123-139. https://doi.org/10.1007/s00477-017-1384-1

Rahman, A., Haddad, K., Zaman, M., Kuczera, G., & Weinmann, P. E. (2010). Design flood estimation in ungauged catchments: A comparison between the probabilistic rational method and quantile regression technique for NSW [Article]. *Australian Journal of Water Resources*, *14*(2), 127-140. https://www.scopus.com/inward/record.uri?eid=2-s2.0-79952583614&partnerID=40&md5=6c0ee146f2cc81478f6cf268a80f9500