Journal of Hydrology: Regional Studies 16 (2018) 67–79

Contents lists available at ScienceDirect

Journal of Hydrology: Regional Studies

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

A comprehensive continent-wide regionalisation investigation for daily design rainfall

F. Johnson[a,*], J. Green[b]

[a] School of Civil and Environmental Engineering, University of New South Wales, Sydney, 2052, Australia
[b] Environment and Research Division, Bureau of Meteorology, Australia

ARTICLE INFO

Keywords:
Design rainfall
Regionalisation
Region of influence

ABSTRACT

Study region: Australia.

Study focus: Design rainfalls, in the form of Intensity Duration Frequency curves, are the standard input for most flood studies. Methods to combine rainfall data across space are required to provide optimal estimates of design rainfalls and constrain their uncertainty. This paper robustly investigates the use of a variety of regionalization methods to provide Australia wide design rainfall estimates using 8619 high quality rainfall stations. The influence of an individual large rainfall event in March 2012 on the regionalised design rainfalls is also investigated to guide decisions on when design rainfall data should be updated following record-breaking rainfall events.

New hydrological insights for the region: The optimum approach was found to be circular regions of influence with a size of 500 station-years. The regions of influence are least biased when defined using spatial proximity and including elevation in the calculating spatial proximity can help to reduce the bias in design rainfall estimates in areas of strong topographic relief. The March 2012 rainfall event increased site rainfall quantiles by 14% on average. When regionalisation is used the increases were substantially moderated and were within the uncertainty bounds of the design rainfalls. It is still not routine to provide uncertainty estimates with design rainfall data, but such information is particularly important given likely future changes in extreme rainfalls due to anthropogenic climate change.

1. Introduction

Design rainfalls (commonly known as Intensity-Duration-Frequency relationships) are used by engineers to calculate flood risk and correctly size flood protection and stormwater infrastructure. Common design standards for such infrastructure include the 1% and 0.5% Annual Exceedance Probability (AEP) events which are used in many parts of the world as flood planning levels (Plate, 2002). It is important that the rainfall totals or intensities associated with these large events are estimated accurately and precisely to enable designs to proceed with known confidence. However in many locations rainfall gauges have operated for periods of time shorter than the return periods of interest and there may be many missing records. In these cases the uncertainty in the estimates of the events of interest can be very large and the fitted probability distributions poorly constrained (Madsen et al., 2002).

To overcome these issues, regionalisation is often adopted, whereby a number of nearby stations are pooled together to better define the extreme value distribution (Hosking and Wallis, 1997). Regionalisation has been commonly used in the analysis of regional flood frequency relationships (e.g. Burn, 1990; Haddad and Rahman, 2012; Merz et al., 2001; Rosbjerg and Madsen, 1995) and also in

* Corresponding author.
E-mail address: f.johnson@unsw.edu.au (F. Johnson).

https://doi.org/10.1016/j.ejrh.2018.03.001
Received 1 August 2017; Received in revised form 21 December 2017; Accepted 1 March 2018
2214-5818/

design rainfalls (Madsen et al., 2002; Nguyen et al., 1998; Norbiato et al., 2007; Wallis et al., 2007). In both cases the general assumption is that a common probability distribution can be used to describe the data at all the sites in a region, possibly with a site-specific scaling factor. The aim of the regionalisation is to define the parameters of this common probability distribution and then use it for design.

The regionalisation methods are based around grouping stations together that have similar distributions for the extreme rainfall statistics. The grouping increases the amount of data that are available to fit the probability distribution. This extra information can lead to increases in both the accuracy and precision of the quantile estimates compared to only using the data from a single site. One way to regionalise data is to delineate a number of different regions and assign every station to a region. The regions can be defined using geographic coordinates or alternatively defined using other features. In either case the clustering seeks to find the stations that are closest to each other in the space of the features used. This approach is often referred to as a fixed region approach as the total number of regions is known. Ways of delineating the regions have also included arbitrary geographic boundaries such as states or counties. An alternative approach to the clustering is that each station is located at the centre of its own Region of Influence (ROI) (Burn, 1990). The ROI includes all the neighbours for the station of interest. An adjacent station will have its own ROI which may be similar but not contain exactly the same subset of stations.

One of the important questions in either approach is how large the regions should be. If the region is too small then there is limited benefit from undertaking the regionalisation. On the other hand if the region is too large then the stations included in the region may be too different from each other or the target station and will lead to biases in the fitted probability distribution. How big is the optimal region? This work seeks to answer this question in the context of daily rainfall data. In the regionalisation of streamflow data region sizes of 5Y where Y is the return period of the quantile of interest have been found to be useful (Centre for Ecology and Hydrology, 2008). However it is not clear how this translates to rainfall data given the higher spatial density of stations and generally longer record lengths that are available at rainfall gauges compared to streamflow gauges. The higher spatial density of stations is likely to lead to more spatial correlations between adjacent gauges, reducing the effective independent data available for regionalisation.

Much of the previous work on regionalisation has been related to streamflow measurements where gauged catchments are located relatively far apart compared to the density of daily rainfall gauge networks. Previous work has also generally focused on regionalisation applied over limited areas. Estimating design rainfalls at a national scale is a large undertaking and as a result the estimates are not regularly updated. For example in Australia design rainfalls were estimated in the mid-1980s and were updated in 2013, whilst in the United States Atlas 14 design rainfall reports have been issued between 2004 and 2015, replacing Atlas 2 estimates from the mid-1970s (NOAA Hydrometeorological Design Studies Center, 2016). Given these long time frames, what happens if a significant rainfall event occurs just after the publication of new design rainfall estimates? When data is regionalised, how strong is the influence of an individual rainfall event on the design rainfalls?

In this paper different regionalisation strategies for Australia are considered using an extensive network of daily read rainfall gauges and recommendations for deriving design rainfalls are made. The impact of isolated historical records on design rainfalls is also considered using techniques that have not previously been tested in a regionalised context. Information on the data sources are presented in the next section, then the different regionalisation strategies that have been considered are described in Section 3. Finally recommendations are made for regionalisation studies in Section 4.

2. Data

A database of daily read rainfall gauges was compiled by the Australian Bureau of Meteorology for the purposes of revising design rainfalls for Australia. Gauges with at least 30 years of data were adopted for this study. The Annual Maximum Series (AMS) was extracted for each station based on years where at least 10 months had less than 25% missing data. Papalexiou and Koutsoyiannis (2013) suggest that when forming the AMS an additional criteria should be implemented that considers the rank of the maximum value in a year that is classified as “missing”. Testing this criterion on the dataset for this study only led to small changes (1%) in the estimated site quantiles. The total number of stations used in determining the optimal regionalisation strategy as discussed here was 8074. The number of stations used in the final design rainfall product was 8619 which included some stations with 20 years of data and some continuous (sub-daily) stations in areas which were sparsely gauged. The data from each station was quality controlled using both automated processes and extensive manual checks (Green et al., 2012a). The AMS data extracted from daily read gauges is expected to underestimate the true 24 h rainfall totals by 10–15% (Dwyer and Reed, 1995). The AMS values were therefore scaled by 1.15 to create equivalent 24 h rainfall maximums (Jakob et al., 2005).

The locations of the 8074 daily rainfall stations used for testing are shown in Fig. 1a and b. The study period covered the full period of record at each station until the end of 2011. The stations with more than 100 years of data are shown in both panels for reference, whilst Fig. 1b shows the locations of all 8074 gauges. The average record length was 65 years and the maximum record length was 156 years. It is clear that despite the relatively sparse population densities in many parts of the country, the coverage of daily rainfall gauges is relatively good although there are very few gauges located in the arid centre of the country. The stations with the longest records tend to be located on the east coast and around Adelaide (South Australia). There are few long records in the north-west of the country which suggests that regionalisation of the rainfall data in this area will be particularly important. Shown in the background of Fig. 1a is the topography of Australia. The highest elevations are found in the Great Dividing Range and Australian Alps which run parallel to the eastern and southern coast line. Relatively large relief is also found in Tasmania, which has significant impacts on rainfall distribution with large annual rainfall totals occurring in the western part of the island due to the moist air flows from the Southern Ocean. Significant orographic effects are known to exist in the coastal hinterland of Queensland and New South