Journal of Hydrology 533 (2016) 439–451

Contents lists available at ScienceDirect

Journal of Hydrology

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

Topographic relationships for design rainfalls over Australia

F. Johnson [a,*], M.F. Hutchinson [b], C. The [c], C. Beesley [c], J. Green [c]

[a] School of Civil and Environmental Engineering, University of New South Wales, Australia
[b] Fenner School of Environment and Society, Australian National University, Australia
[c] Environment and Research Division, Bureau of Meteorology, Australia

ARTICLE INFO

Article history:
Received 11 June 2015
Received in revised form 2 October 2015
Accepted 18 December 2015
Available online 23 December 2015
This manuscript was handled by Andras Bardossy, Editor-in-Chief, with the assistance of Felix Frances, Associate Editor

Keywords:
IDF rainfall relationships
Thin plate smoothing splines
Topographic dependence
Elevation scaling
Knots
Data transformation

SUMMARY

Design rainfall statistics are the primary inputs used to assess flood risk across river catchments. These statistics normally take the form of Intensity–Duration–Frequency (IDF) curves that are derived from extreme value probability distributions fitted to observed daily, and sub-daily, rainfall data. The design rainfall relationships are often required for catchments where there are limited rainfall records, particularly catchments in remote areas with high topographic relief and hence some form of interpolation is required to provide estimates in these areas. This paper assesses the topographic dependence of rainfall extremes by using elevation-dependent thin plate smoothing splines to interpolate the mean annual maximum rainfall, for periods from one to seven days, across Australia. The analyses confirm the important impact of topography in explaining the spatial patterns of these extreme rainfall statistics. Continent-wide residual and cross validation statistics are used to demonstrate the 100-fold impact of elevation in relation to horizontal coordinates in explaining the spatial patterns, consistent with previous rainfall scaling studies and observational evidence. The impact of the complexity of the fitted spline surfaces, as defined by the number of knots, and the impact of applying variance stabilising transformations to the data, were also assessed. It was found that a relatively large number of 3570 knots, suitably chosen from 8619 gauge locations, was required to minimise the summary error statistics. Square root and log data transformations were found to deliver marginally superior continent-wide cross validation statistics, in comparison to applying no data transformation, but detailed assessments of residuals in complex high rainfall regions with high topographic relief showed that no data transformation gave superior performance in these regions. These results are consistent with the understanding that in areas with modest topographic relief, as for most of the Australian continent, extreme rainfall is closely aligned with elevation, but in areas with high topographic relief the impacts of topography on rainfall extremes are more complex. The interpolated extreme rainfall statistics, using no data transformation, have been used by the Australian Bureau of Meteorology to produce new IDF data for the Australian continent. The comprehensive methods presented for the evaluation of gridded design rainfall statistics will be useful for similar studies, in particular the importance of balancing the need for a continentally-optimum solution that maintains sufficient definition at the local scale.

1. Introduction

The relationship between topography and climate, in particular rainfall, has been well studied, particularly at annual, monthly and daily time scales (Daly et al., 1994; Hutchinson, 1995; Phillips et al., 1992). Other work has considered the relationship of topography with the behaviour of individual storm systems (e.g. Tarolli et al., 2013). One area however that has received limited attention is the relationship between topography and extreme rainfall totals.

For hydrologic and hydraulic engineering problems, design rainfalls are the primary input for understanding flood risk in any particular catchment. Given the wide need for such flood studies it is quite likely that risks need to be calculated for catchments where there are limited or no rainfall records, particularly for sub-daily durations. It is therefore usual to interpolate point design rainfalls onto a regular grid so that researchers and practitioners have design rainfall estimates for any point of interest.

Thin plate smoothing splines have been a popular choice in undertaking such interpolations for other rainfall fields such as annual and daily rainfall due to their ease of fitting and good performance (Hutchinson, 1995; Hutchinson et al., 2009). This paper assesses whether thin plate smoothing splines can also provide

* Corresponding author at: School of Civil and Environmental Engineering, University of New South Wales, Sydney 2052, Australia.

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

http://dx.doi.org/10.1016/j.jhydrol.2015.12.035
0022-1694/

good results for design rainfalls. A primary motivation for the current work was the revision of design rainfalls by the Bureau of Meteorology for Australia (Green et al., 2012) as part of the revision of *Australian Rainfall and Runoff* (Engineers Australia, 1987). This project has provided design rainfall estimates for 1 min to 7 days for mainland Australia and Tasmania by analysing point rainfall records. The final stage of the project was to generate grids of design rainfalls that can be queried for any location and any duration.

The process of converting point rainfall statistics to a grid of information for design rainfalls is an interpolation problem. The aim is to fit a surface that interpolates the rainfall statistics, recognising that each data point will have an error due to the natural variability of the rainfall statistic, possible underlying observation and location errors and errors due to locations not being representative of the surrounding area. The last is essentially an allowance for the imperfections in the spatial statistical model. The goal is to find a surface that is consistent with the data with the right balance between variance and bias (Hastie et al., 2009). Bias will result if the fitted surface is too smooth whereas high variance would be caused by following individual data points too closely. One of the research gaps that this paper attempts to answer is whether the balance between bias and variance can be considered constant across Australia, leading to a simple national optimisation problem or if regional topographic and hydrologic characteristics require a more refined approach.

The relationship between rainfall and topography is reasonably well understood although most work has focused on daily, seasonal or annual rainfall characteristics with limited previous research into daily and sub-daily rainfall extremes. The notable exceptions are Prudhomme (1999) who considered the median of the annual maximum series (AMS) of daily rainfall in Scotland by residual kriging and Allamano et al. (2009) who investigated kriging and regression of maximum rainfall intensities for sub-daily durations. These methods are discussed below and compared to thin plate smoothing splines. This paper attempts to address the gap in the understanding of spatial patterns of daily extreme rainfall by conducting a continent-wide analysis for Australia whereas the previous studies had more limited geographical extents.

Previous analyses has shown that elevation and exposure to the prevailing wind are important controls on rainfall distribution (Basist et al., 1994). The physical processes that govern these relationships include vertical uplift of moist air by land masses, changes to the passage of synoptic scale systems such as low pressures systems or fronts and the promotion of local convection (Roe, 2005). Multivariate interpolation methods that include elevation have been found to outperform simpler bivariate methods for annual and monthly rainfall (Goovaerts, 2000; Hutchinson, 1995). It is not known how these previous findings transfer to daily extreme rainfalls. Geographic and topographic data can also lead to improvements in the mapping of daily extreme rainfall statistics (Prudhomme, 1999). The additional information is useful because ordinary bivariate kriging cannot account for the uneven spatial density of the gauge network; uplands and mountainous areas tend to be much more sparsely sampled than lowlands and valleys.

Many studies of rainfall and elevation have investigated kriging and its variants (Allamano et al., 2009; Goovaerts, 2000; Haberlandt, 2007; Phillips et al., 1992; Prudhomme, 1999; Wagner et al., 2012). Other methods have included copulas (Bárdossy and Li, 2008; Wasko et al., 2013), local elevation dependent methods (Daly et al., 1994; Nalder and Wein, 1998; Thornton et al., 1997) and thin plate smoothing splines (TPS) (Hutchinson, 1995; Jeffrey et al., 2001). Price et al. (2000) showed that TPS gave superior performance to the local elevation dependent method of Nalder and Wein (1998), particularly in mountainous data-sparse locations. On the other hand, Beesley et al. (2009) compared TPS and kriging for interpolating daily rainfall across Australia and found very similar performance. In a study in Mexico the difference between the methods was found to be smaller than the improvement obtained from including elevation in either model (Boer et al., 2001). One argument in favour of the use of splines is that the prediction errors can be accurately estimated without the need for assumptions about the correct model for the variogram parameters (Hutchinson and Gessler, 1994). Due to the large climatic and topographic variations over the study area and the heteroscedasticity of the design rainfall statistics it is difficult to obtain such variogram parameters. Therefore TPS has been adopted for the analyses.

The general question to be addressed in this paper is therefore how best to use the TPS for understanding design rainfall spatial relationships. There are also a number of specific questions that arise. What is the fundamental spatial scale of design rainfall in relation to topography in order to represent true spatial variations but not overfit the data or have unrealistic small scale variations? How should the smoothing spline parameters be specified? Should a transformation be applied to remove skewness in the precipitation statistics before application of thin plate spline interpolation, as is common practice? Does it matter whether the rainfall quantiles or the probability distribution parameters are interpolated? How can the gridded relationships be verified at local and continental scales? This paper addresses each of these questions.

2. The thin plate smoothing spline model

Thin plate smoothing splines are a multivariate generalisation of univariate cubic spline smoothers which are commonly used to fit a smooth relationship between a single predictor and a noisy predictand. Thin plate smoothing splines can similarly be used to fit a smooth surface through a set of noisy data distributed across a multivariate space. The noise in the data may have arisen from measurement error or other sources, including deficiencies in the smoothing spline model. This fitted relationship can be used to understand the spatial patterns in the data and underlying physical processes, as well as being used for prediction across space. Readers are referred to Wahba (1990) for more details on TPS.

The partial spline model for N observed data values z_i (in this case statistics summarising the extreme rainfall properties) can be defined by:

$$z_i = f(x_i) + b^T y_i + e_i \quad \text{for } i = 1, \ldots, N \tag{1}$$

where x_i is a d-dimensional vector of spline independent variables (described in detail in Section 2.1 below), f is an unknown smooth function of x_i, y_i is a p-dimensional vector of independent covariates with b an unknown p-dimensional vector of coefficients and e_i are independent normally distributed errors with variance σ^2 which is usually assumed to be constant across all rainfall stations. As noted above, the error term accounts for both errors in the data values and limitations in the spline representation of the field being interpolated. For the extreme value statistics considered here, year to year variability in the annual maximum precipitation values is a major contributor to this error term.

The spline function f and the coefficient vector b are estimated by minimising

$$\sum_{i=1}^{N} [z_i - f(x_i) - b^T y_i]^2 + \rho J_m(f) \tag{2}$$

where $J_m(f)$ is a non-negative measure of the complexity of f, defined in terms of an integral of mth order partial derivatives of f, and ρ is a positive number called the smoothing parameter. The smoothing parameter is the primary determinant of the trade-off between goodness of fit and simplicity of the fitted spline function,