ELSEVIER

Journal of Hydrology 256 (2002) 196–210

Journal of Hydrology

www.elsevier.com/locate/jhydrol

Monte Carlo simulation of flood frequency curves from rainfall

A. Rahman[a,*], P.E. Weinmann[b], T.M.T. Hoang[b,c], E.M. Laurenson[b]

[a]Physical Infrastructure Centre, School of Civil Engineering, Queensland University of Technology, Brisbane, Qld. 4001, Australia
[b]Department of Civil Engineering, Cooperative Research Centre for Catchment Hydrology, Monash University, Clayton, Vic. 3800, Australia
[c]Bureau of Meteorology, Melbourne, Vic. 3000, Australia

Received 6 March 2001; revised 29 August 2001; accepted 19 September 2001

Abstract

The currently adopted rainfall-based design flood estimation techniques, for example the Design Event Approach, do not account for the probabilistic nature of the key variables except for the rainfall depth. This arbitrary treatment of key inputs and parameters can lead to inconsistencies and significant bias in flood estimates for a given average recurrence interval. This paper presents a Monte Carlo simulation technique that makes explicit allowance for the probability-distributed nature of the key flood producing variables and the dependencies between them to determine derived flood frequency curves. The proposed approach employs joint probability principles to develop a design flood estimation technique that can incorporate commonly applied rainfall–runoff models and design data. The application of the proposed technique to three catchments in Victoria has shown that the new method provides a relatively precise reproduction of the observed frequency curves. The new technique is relatively easy to apply for catchments with good rainfall data and a limited streamflow record. The technique thus shows a strong potential to become a practical design tool; further work is needed to allow its routine application in a wider range of design situations.

Keywords: Monte Carlo simulation; Flood frequency; Joint Probability Approach; Derived distribution; Design rainfall; Rainfall–runoff

1. Introduction

Flood design generally requires the estimation of a frequency curve of a selected flood characteristic (e.g. peak flow rate, maximum flood level) at a site. If adequate streamflow data are available at or near the site of interest, the flood estimates can be derived directly from flood frequency analysis. In the more general case of catchments with limited streamflow data, or in catchments subject to major land use changes, design floods are generally estimated based on design rainfalls and, where applicable, other flood producing factors, such as snowmelt.

With all methods of rainfall-based design flood estimation currently applied in practice, a key issue to be resolved is how the design rainfall input for a given average recurrence interval (ARI) can be transformed into a design flood output of corresponding ARI. For the simple case of the Probabilistic Rational Method (IEAUST, 1998; Pilgrim and Cordery, 1993, Sec. 9.4), this problem has been directly addressed by using calibration data from gauged catchments to ensure that the derived design runoff coefficient for the gauged site correctly transforms the design rainfall frequency curve into the ‘observed’ flood frequency curve at the catchment

* Corresponding author. Address; Physical Infrastructure Centre, School of Civil Engineering, Queensland University of Technology, Brisbane, Qld. 4001, Australia. Tel.: +61-7-3864-1543; fax: +61-7-3864-1515.

E-mail address: a.rahman@qut.edu.au (A. Rahman).

0022-1694/02/$ - see front matter
PII: S0022-1694(01)00533-9

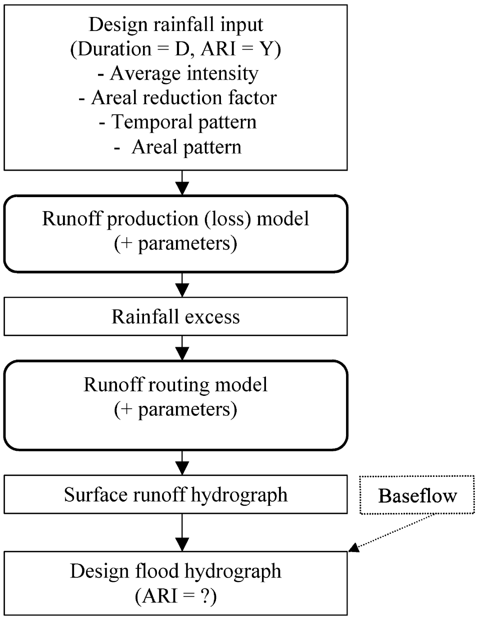

Fig. 1. Design Event Approach to rainfall-based design flood estimation.

outlet. However, the simplifying assumptions made when transferring the design runoff coefficients from gauged to ungauged catchments place severe constraints on the applicability and accuracy of the method.

For the unitgraph or runoff routing modelling approaches, such calibration is much more difficult to achieve, as several interacting inputs and parameters are involved. The current practice aims to define hypothetical *design events* of model inputs and model parameters that can be considered representative in a probability sense, that is they should transform the design rainfall input of a given ARI into a flood output of the same ARI. The problem of finding the critical rainfall duration for a specific design situation is addressed by a trial and error approach, adopting the rainfall duration that produces the largest flood outputs for a given input ARI. With this design practice, the equivalence of input and output ARI is not intrinsically assured but, unless checked against regional or at-site flood information, remains an assumption that is satisfied only for a limited set of conditions.

This paper presents a more holistic approach of design flood estimation, based on the principle of joint probability. The new method adopts some of the key modelling elements of the current Design Event Approach, to facilitate the development of practical design tools and their adoption by practitioners, but it accounts for the probability-distributed nature and correlations of the key model inputs/parameters. A large number of rainfall and runoff events are simulated (Monte Carlo simulation technique) using a calibrated rainfall–runoff model; these are then used to determine a derived flood frequency curve.

2. Limitations of current design flood estimation techniques and options for improvement

The commonly used rainfall-based flood estimation techniques are based on the Design Event Approach (shown in Fig. 1) in which a design rainfall intensity for specified duration and ARI is used in combination with 'typical values' of other relevant model inputs and parameters to obtain design flood estimates.

The key assumption involved in the Design Event Approach is that the representative design values of the inputs/parameters in the modelling can be defined in such a way that they are 'ARI-neutral', i.e. they result in a flood output that has the same ARI as the rainfall input. However, there are no definite guidelines on how to select the appropriate values of the inputs/parameters except for the average rainfall intensity, which is described by a conditional probability distribution with respect to rainfall event duration. In this regard, it is assumed that adoption of the rainfall intensity corresponding to the 'critical' rainfall duration adequately reflects the contribution of rainfall events of different durations to the total flood probability.

With regard to the selection of other design values, a designer is commonly in the situation to select a representative input/parameter value from a wide range of possible values. Due to the non-linearity of the transformation process involved, it is generally not possible to know a priori how a representative value should be selected to preserve the ARI, and the commonly used mean or median value from a sample of inputs or fitted parameter values may be a poor choice. The arbitrary treatment of various inputs/parameters in the Design Event Approach can lead to inconsistencies and significant bias in flood estimates for a given ARI. This is likely to result in systematic under- or over-design of engineering structures, both with important economic consequences.

To overcome the major limitations with the Design Event Approach, a number of more 'holistic' methods for design flood estimation have been proposed. They can be broadly classified as the Continuous Simulation Approach or the Joint Probability Approach. In the former, a complete time series of streamflows for an extended period is simulated, while the latter approach concentrates on the simulation of a large set of flood events. This paper focuses on methods that apply the Joint Probability Approach.

The Joint Probability Approach considers probability-distributed inputs and model parameters and their correlations to determine a derived distribution of a selected flood characteristic (e.g. flood peak, flood volume). This Derived Distribution Approach was pioneered by Eagleson (1972) who used an analytical method to derive the probability distribution of peak streamflow from an idealised V-shaped flow plane. A similar analytical approach has been adopted in some later applications (e.g. Wood, 1976; Hebson and Wood, 1982; Diaz-Granados et al., 1984) for idealised conditions, but it has limited applicability to real catchment situations.

A number of researchers (e.g. Beran, 1973; Laurenson, 1974) have used an approximate joint probability method, where the continuous distributions of hydrologic variables are discretised by dividing the possible range of a random variable into class intervals. The theorem of total probability is then applied to derive the joint probability distribution of the output in a discrete form.

Other investigators (e.g. Beran, 1973; Muzik, 1993; Durrans, 1995; Blöschl and Sivapalan, 1997) adopted a Monte Carlo simulation approach to determine a derived flood distribution. This involves random sampling from continuous distributions of input variables and parameters, and use of a rainfall–runoff model to obtain the flood hydrograph. The procedure is repeated N times (N in the order of thousands), and the N different values of the output variable are then used to determine the derived distribution.

Many of the previous applications employing the Joint Probability Approach were confined to theoretical studies; mathematical complexity, difficulties in parameter estimation and limited flexibility constrain the application of these techniques in practical situations. From the consideration of practical applicability, flexibility and the ability to account for the dependence between the input variables, the Monte Carlo simulation technique appears to be the most promising method to determine derived flood frequency distributions under practical situations, and thus has been adopted in this study. The new method offers scope for further development beyond the recent work by others. It is intended for use with commonly applied loss and runoff routing models, establishes a link with currently available design data and deals explicitly with any dependencies between key model inputs/parameters.

3. Outline of the adopted modelling framework

The proposed modelling framework is based on three principal elements: (a) a (deterministic) hydrologic modelling framework to simulate the flood formation process; (b) the key model variables (inputs and parameters) with their probability distributions and dependencies; and (c) a stochastic modelling framework to synthesise the derived flood distribution from the model input/parameter distributions. These elements are discussed in Sections 3.1–3.3.

3.1. Hydrologic modelling framework

The proposed hydrologic model of the flood formation process involves the same components as the models most commonly used with the current Design Event Approach (see Fig. 1): a runoff production function (or loss model), and a runoff transfer function (or runoff routing model).

3.1.1. Runoff production function — loss model

A runoff production function (or loss model) is needed to partition the gross rainfall input into effective runoff (or rainfall excess) and loss. Most of the previous derived distribution studies (e.g. Eagleson, 1972; Russell et al., 1979) have used an empirical infiltration equation (such as Horton's equation) or a more physically based equation (such as the Phillip and Green Ampt infiltration equations) to estimate the rainfall excess.

In design practice, use of simplified, lumped conceptual loss models is preferred over the mathematical equations because of their simplicity and ability to approximate catchment runoff behaviour. This is particularly true for design loss which is probabilistic in nature and for which complicated theoretical models may not be required. On this basis, the initial loss–continuing loss model has been adopted in the present study. In this model, it is assumed that no runoff is generated from a rainfall event until the cumulative rainfall exceeds the initial loss value; for the remainder of the event, loss is assumed to occur at a constant rate.

3.1.2. Transfer function — runoff routing model

A catchment response model is needed to convert the rainfall excess hyetograph produced by the loss model into a surface runoff hydrograph. The models commonly used in previous joint probability studies include: Kinematic Wave Model (e.g. Eagleson, 1972), Geomorphologic Unit Hydrograph Model (e.g. Diaz-Granados et al., 1984), Unit Hydrograph Method (e.g. Beran, 1973; Muzik, 1993), Clark's Model (Russell et al., 1979), and parallel linear storages (Blöschl and Sivapalan, 1997).

The proposed modelling framework is intended for application with a non-linear, semi-distributed runoff routing model like RORB (Laurenson and Mein, 1997), as the runoff and streamflow routing processes are generally non-linear and spatially variable. However, for the present study, a simpler conceptual runoff routing model (with a single non-linear storage concentrated at the catchment outlet) was adopted to reduce computational effort. This model can account for catchment non-linearity but not for the distributed nature of catchment storage; for medium size catchments up to about 500 km^2, it provides an indication of what could be achieved with a semi-distributed model such as RORB. For the adopted runoff routing model, the storage–discharge relationship is expressed by

$$S = kQ^m \tag{1}$$

where S is catchment storage in m^3, k a storage delay parameter in hour, Q the rate of outflow in m^3/s and m is a non-linearity parameter (taken as 0.8 here).

3.2. Key model variables

The major factors affecting storm runoff production are rainfall duration, rainfall intensity, temporal and areal patterns of rainfall, and storm losses. Factors affecting hydrograph formation are the catchment response characteristics embodied in the runoff routing model (model type, structure, and parameters) and design baseflow. Ideally, all the variables should be treated as random variables but, for practical reasons, application of a smaller number of random variables would be preferable, if it did not result in a significant loss of accuracy. Given the dominant role of rainfall and loss in the flood formation process for Australian conditions, it might be expected that the incorporation of the probabilistic nature of these variables would result in significant reduction of bias and uncertainties in design flood estimates. Thus, four variables have been considered here for probabilistic representation: rainfall duration, rainfall intensity, rainfall temporal pattern and initial loss. In contrast to this, the currently used Design Event Approach treats only the rainfall intensity for a given duration as a probabilistic variable.

The areal distribution of rainfall over the catchment is assumed to be uniform, and the average catchment rainfall intensity is obtained from the point design rainfall intensity using an areal reduction factor (e.g. Siriwardena and Weinmann, 1996). The continuing loss is assumed to be a constant; likewise, a constant baseflow is assumed, determined as the average baseflow at the start of surface runoff generation in observed events. A single set of parameter values for the runoff routing model is used here; the calibration procedure allows the determination of a set of model parameters for a given catchment, which can be applied with reasonable confidence.

Thus, the adopted Monte Carlo simulation technique considers probabilistic modelling related to the runoff production only, the hydrograph formation part (e.g. runoff routing) remains entirely deterministic. It has been left to future research efforts to determine if the probabilistic treatment of any of the above variables kept constant in the simulation might further improve the flood estimates.

3.3. Stochastic modelling framework

The basic idea underlying the proposed new modelling framework is that the distribution of the flood outputs can be directly determined by simulating the

Table 1
List of application catchments

Catchment	Catchment area (km^2)	Streamflow station no.	Period of streamflow data (years)	Pluviograph station and period of record (years)	Mean annual rainfall (mm)	Mean annual evaporation (mm)
Boggy Creek at Angleside	108	403226	1974–1998 (25)	Station 83031 1974–1993 (20)	1020	1400
Tarwin River east branch	127	227226	1957–1979 (22)	Station 85106 1957–1979 (22)	1005	1200
Avoca River at amphitheatre	78	408202	1975–1999 (25)	Station 81038 1974–1993 (20)	745	1333

possible combinations of hydrologic model inputs and parameter values. Here, we adopted a Monte Carlo simulation approach for its relative simplicity and flexibility.

For each run of the combined loss and runoff routing model, a specific set of input/parameter values is selected by randomly drawing a value from each of the respective distributions (for probability distributed variables) and by choosing a representative value (for other variables). Any significant correlation between the input variables is allowed for by using conditional probability distributions. For example, the strong correlation between rainfall duration and intensity is allowed for by first drawing a value of duration and then a value of intensity from the conditional distribution of rainfall intensity for that duration interval. The results of the run (e.g. flood peaks at the catchment outlet) are then stored and the Monte Carlo simulation process is repeated a sufficiently large number of times to fully reflect the range of variation of input/parameter values in the generated output. The output values of a selected flood characteristic (e.g. flood peak) can then be subjected to a frequency analysis to determine the derived flood frequency curve for the ARI range of interest.

4. Study catchments and data

The new technique was applied to three relatively small catchments in the sub-humid to humid parts of Victoria, Australia: Boggy Creek (catchment area, $A = 108\ km^2$), Tarwin River ($A = 127\ km^2$) and Avoca River ($A = 78\ km^2$). In the region represented by these catchments, heavy storm rainfall is the dominant flood producing factor, with snowmelt making only a relatively minor contribution to streamflow. The relevant details for these catchments are summarised in Table 1.

The available streamflow data for each catchment is sufficient to allow a flood frequency analysis, as a basis for an evaluation of the results of the new method. Pluviograph data from stations located in the catchments was supplemented by data from stations in a broader region around the catchments. The original pluviograph data was processed into hourly accumulations, as for the size of catchments considered here, a 1 h minimum time step was considered sufficient.

5. Calibration of runoff routing model

In calibration of the adopted runoff routing model (Section 3.1.2), m was taken to be 0.8 and a value of k was determined for each of the three selected application catchments (as listed in Section 4) to arrive at a satisfactory fit for the selected recorded rainfall and runoff events. Three to five events were used for calibration and validation of the runoff routing model parameters. Baseflow separation was achieved by application of a filtering technique (Lyne and Hollick, 1979). A typical result of the calibration for the Tarwin River catchment is shown in Fig. 2. The results reflect the limitations of the available storm rainfall data and the lumped treatment of the catchment response. The model calibration is judged to be satisfactory, as the specific interest in this study is on the simulation of peak flows, which are well reproduced.

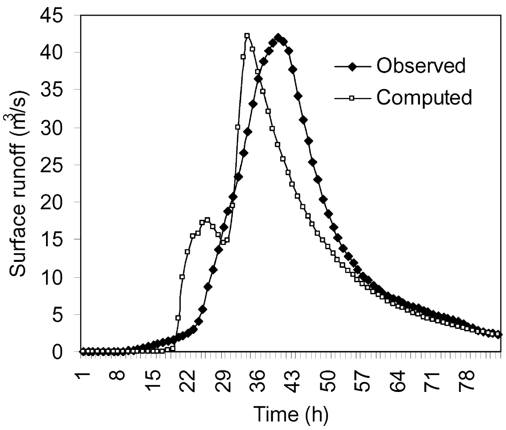

Fig. 2. Fitting the runoff routing model ($S = kQ^m$) at Tarwin River catchment ($m = 0.8$, $k = 30$ h).

6. Distributions of the key variables

6.1. Rainfall event definition

The proposed method requires rainfall events to be of random duration unlike the rainfall bursts of predetermined durations used in the Design Event Approach (IEAUST, 1998). For the purposes of the proposed method, 'complete storm' and 'storm-core' (most intense part of the storm) events are defined as follows.

A complete storm is a period of 'significant' rainfall that is separated from previous and subsequent rainfall events by a 'dry' period (Fig. 3). Here a period is defined as 'dry' if it lasts at least 6 h, the average rainfall intensity during the period is less than C_1 (mm/h), and all individual hourly rainfalls are less than C_2 (mm). The choice of the parameters defining a 'dry' period depends on the intended application of the selected storms; for a small urban catchment, the choice of a shorter separation time would ensure that the selected storms tend to be of shorter duration.

A complete storm is considered to be 'significant' if it has the potential to produce significant runoff. This is assessed by comparing its average rainfall intensity with a threshold intensity. Thus, the average rainfall intensity for a complete storm during the entire storm duration (I_D) or during a sub-storm duration (I_d) must satisfy one of the following conditions:

$$I_D \geq f_1 \times {}^2I_D \qquad \text{or} \qquad I_d \geq f_2 \times {}^2I_d$$

where f_1 and f_2 are reduction factors, the threshold intensity ${}^2I_D$ is the 2 year ARI design rainfall intensity for the selected storm duration D, and ${}^2I_d$ is the corresponding intensity for the sub-storm duration d. The design rainfall intensities are obtained from the available design rainfall information, commonly available in the form of intensity–frequency–duration (IFD) curves (e.g. IEAUST, 1998).

The use of appropriate values of f_1 and f_2 allows the selection of the events that have the potential to produce significant storm runoff. The use of smaller values of f_1 and f_2 captures a relatively larger number of events; appropriate values need to be selected such that events of very small average intensity are not included. In this study, the following parameter values have been found to be appropriate for rainfall stations in Victoria: $C_1 = 0.25$ (mm/h), and $C_2 = 1.2$ (mm), $f_1 = 0.4$ and $f_2 = 0.5$. This typically resulted in an average of 3–6 rainfall events per year being selected.

For each complete storm, a single storm-core can be identified, defined as 'the most intense rainfall burst within a complete storm' (Fig. 3). It is found by calculating the average intensities of all possible storm bursts, and the ratio with an index rainfall intensity ${}^2I_d$ for the relevant duration d, then selecting the burst of that duration which produces the highest ratio. Both complete storms and storm-cores can be used to obtain derived flood frequency curves. This paper, in particular, describes the techniques developed in relation to storm-cores. Hoang (2001) used complete storms to obtain derived flood frequency curves.

6.2. Rainfall duration and average rainfall intensity

The storm-cores are selected from the hourly pluviograph data of selected station(s) and analysed for storm-core duration (d_c), average rainfall intensity (I_c) and temporal patterns (TP_c). The distributions of d_c were examined for 29 pluviograph stations selected from Victoria and an exponential distribution was found to be appropriate to approximate the distribution. The observed and generated d_c data from the fitted exponential distribution for Station 83031 (used to obtain derived flood frequency curve for Boggy Creek, Fig. 8) are compared in Fig. 4. A similar degree of agreement between the observed and

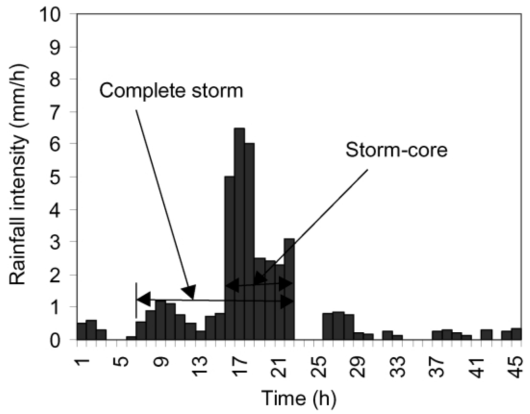

Fig. 3. Rainfall events: complete storm and storm-core.

generated d_c data was observed for the other selected stations in the study area.

As expected, a strong relationship between d_c and I_c was observed, confirming that the distribution of I_c needs to be conditioned on d_c. In practice, the conditional distribution of rainfall intensity is expressed in the form of IFD curves, in that rainfall intensity is plotted as a function of rainfall duration and frequency. The IFD curves for storm-core rainfall intensity were developed in a number of steps, as follows:

- The range of storm-core durations d_c was divided into a number of class intervals (with a representative duration being selected for each class except the 1 h class), e.g. 2–3 h (representative duration 2 h), 4–12 h (6 h), 13–36 h (24 h).
- For the data in each class interval (except the 1 h class), a linear regression line was fitted between $\log(d_c)$ and $\log(I_c)$. The slope of the fitted regression line was used to adjust the intensities for all durations within the interval to the representative duration.
- An exponential distribution was fitted to the partial series of the adjusted intensities within the class interval, and design intensity values I_c(ARI) were computed for ARIs of 2, 5, 10, 20, 50 and 100 years.
- For a selected ARI, the computed I_c(ARI) values for each duration range were used to fit a second-degree polynomial between $\log(d_c)$ and $\log(I_c)$. For the 15 pluviograph stations analysed in Victoria and for all the selected ARIs, the observed R^2 values were greater than 98%. These polynomials can be used to obtain for each selected ARI a value of rainfall intensity I_c for a duration d_c (1 h $\leq d_c \leq$ 100 h).

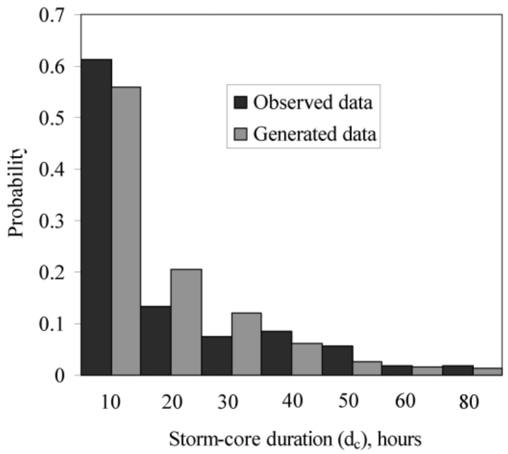

Fig. 4. Distribution of storm-core durations (d_c) at Station 83031. Observed and generated data from the fitted exponential distribution are compared.

The set of storm-core IFD curves obtained for Station 83031 is presented in Fig. 5. The separately fitted curves for different ARIs (up to 100 years) show a high degree of consistency. The application of this procedure to the 15 pluviograph stations produced a generally consistent set of IFD curves.

The adopted Monte Carlo simulation scheme starts with the generation of a d_c value from its marginal distribution. Given this d_c and a randomly generated ARI value (see Section 7 for generation of ARI value), the rainfall intensity value I_c is drawn from the conditional distribution of I_c, expressed in the form of IFD curves. This requires the definition of a continuous distribution function, ideally in the form of a functional relationship between d_c, I_c and ARI. However, as it is difficult to derive a functional relationship that suits different conditions, an IFD table is used here with an interpolation procedure to generate I_c values, for any given combination of d_c and ARI. In an IFD table, I_c values are tabulated for d_c values of 1, 2, 6, 24, 48, 72 and 100 h, and ARIs of 0.1, 1, 1.11, 1.25, 2, 5, 10, 20, 50, 100, 500, 1000 and 1,000,000 years. A linear interpolation in the log domain is used between the tabulated values of d_c and ARI.

It should be noted here that I_c values for ARIs less than 1 year and greater than 100 years are of less

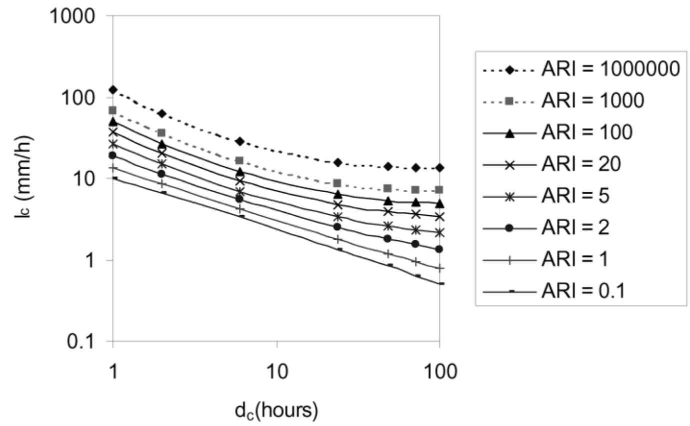

Fig. 5. Storm-core rainfall intensity-frequency-duration curves (Station 83031). Dotted lines indicate extrapolated data.

direct interest in the development of derived flood frequency curves for design flood estimation up to the limit of the 100-year flood. However, these extrapolated values are required to cover the range of events that might produce floods in the range of interest in the Monte Carlo simulation. The part of the developed IFD curves for ARIs of 100–1,000,000 years is subject to very large estimation errors from pluviograph data of limited lengths (normally less than 50 years). Where the interest is on rare to extreme floods (ARI greater than 100 years), this part of the curves needs to be adjusted using design rainfall data obtained from an appropriate technique (e.g. regional methods of estimating rare to extreme design rainfalls or probable maximum precipitation (PMP) methods).

6.3. Rainfall temporal pattern

A rainfall temporal pattern is a dimensionless representation of the variation of rainfall intensity over the duration of the rainfall event. In this study, the time distribution of hourly rainfall during a storm has been characterised by a dimensionless mass curve, i.e. a graph of dimensionless cumulative rainfall depth versus dimensionless storm time with 10 equal time increments.

Using contingency tables and the chi-squared test (similar to Garcia-Guzman and Aranda-Oliver, 1993), the storm-core temporal patterns for 19 pluviograph stations of south-east Victoria were analysed. The results of the analysis indicated that the temporal patterns of rainfall depth for storm-cores (TP_c) are not dependent on season and total storm depth. This means that dimensionless temporal patterns from different seasons and for different rainfall depths could be pooled. However, the patterns were found to be dependent on storm duration, yielding two groups: (1) up to 12 h duration, and (2) greater than 12 h duration. As the rainfall data used in the analysis was only defined at hourly intervals, the minimum storm-core duration used in the temporal pattern analysis was 4 h. Storms with less than 4 h durations are assumed to have the same temporal patterns as the observed 4–12 h storms. The typical variability in dimensionless temporal patterns for Victorian storm rainfall data for 4–12 h duration is presented in Fig. 6. The high degree of variability in the observed temporal patterns highlights the difficulty of defining a single 'representative' pattern and the desirability of modelling the temporal distribution of storm rainfall stochastically.

Design temporal patterns for storm-cores (TP_c) could be generated by the 'multiplicative cascade model' applied by Hoang (2001). However, in the present Monte Carlo simulation technique, historic temporal patterns were used directly instead of generated temporal patterns. That is, observed temporal patterns (in dimensionless form) were drawn randomly from the sample corresponding to the generated d_c value.

6.4. Initial loss

The initial loss for a complete storm (IL_s) is

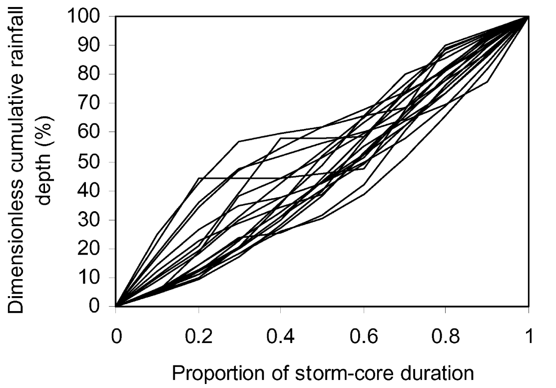

Fig. 6. Typical temporal patterns for 4–12 h durations (Station 83031).

estimated to be the rainfall that occurs prior to the commencement of surface runoff. The storm-core initial loss (IL_c) is the portion of IL_s that occurs within the storm-core. The value of IL_c can range from zero (when surface runoff commences before the start of the storm-core) to IL_s (when the start of the storm-core coincides with the start of the complete storm event).

Initial losses for 11 Victorian catchments were analysed. By definition, $IL_c \leq IL_s$ and the ratio IL_c/IL_s varies between 0 and 1.0. The plots of IL_c/IL_s versus d_c for the study catchments, in general, indicated that as d_c increases IL_c/IL_s approaches unity. Based on these results and findings of other studies in the same region (e.g. Hill et al., 1996), the relationship between IL_c, IL_s and d_c was expressed by the following equation:

$$IL_c = IL_s[0.5 + 0.25 \log_{10}(d_c)] \quad (2)$$

This relationship gives $IL_c = 0.5IL_s$ at $d_c = 1$ h, and $IL_c = IL_s$ at $d_c = 100$ h. It might be noted here that the use of IL_s distribution (with a adjustment factor) as proposed in Eq. (2) is preferable to the use of IL_c directly as IL_s is more readily determined from data and can probably be derived using existing design loss data (e.g. Hill et al., 1996). The distributions of IL_s for the study catchments were found to be positively skewed (Fig. 7), and a Beta distribution was used to approximate the empirical distribution of IL_s. The observed and generated IL_s data for the Boggy Creek catchment are compared in Fig. 7. A similar degree of agreement between the observed and generated IL_s data was observed for the other study catchments.

7. Monte Carlo simulation

'Monte Carlo simulation' refers to a mathematical technique that is used to determine the outputs from a model represented by a complex set of equations that cannot be readily solved analytically. In this study, the Monte Carlo simulation approach is used to generate a sample of *NG* different runoff events from *NG* different combinations of rainfall and loss inputs. For each event, a set of values of d_c, I_c, TP_c and IL_c is generated to define the rainfall excess hyetograph, which is then routed through the calibrated runoff routing model to produce a corresponding streamflow hydrograph. A large number of hydrographs (in the order of thousands) is generated and the resulting flood peaks are extracted and subjected to a frequency analysis to obtain the derived flood frequency curve.

7.1. Number of runoff events generated

The number of separate events to be generated depends on the range of ARIs of interest, the degree of accuracy required, the number of probability-distributed variables involved and the degree of correlation between them. For the study catchments, it was found that at least 2000 years of data (with an average of 3–6 events per year) have to be generated to produce relatively stable estimates of the derived flood frequency curve in the ARI range from 1 to 100 years.

If the purpose of the Monte Carlo simulation was to estimate flood events in the extreme range, or if more independent random variables were involved, the required number of generated events would increase by orders of magnitude. It would then be desirable to apply more efficient Monte Carlo simulation methods, such as importance sampling (e.g. Thompson et al., 1997).

The number of partial series flood events to be generated (*NG*) is obtained from the following equation:

$$NG = \lambda \times NY \quad (3)$$

where λ is the average number of storm-core events per year, and *NY* is the number of years of data to be

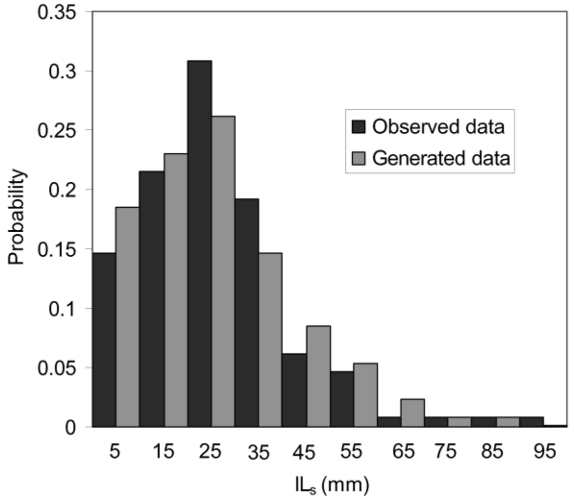

Fig. 7. Distribution of complete storm initial losses (IL_s) at Boggy Creek catchment. Observed and generated data from the fitted Beta distribution are compared.

generated. As an example, for λ equal to 5, a total of 10,000 data points have to be generated to simulate 2000 years of data. The selected value of λ should be similar to the average number of observed storm-cores per year at the catchment of interest using the adopted storm-core definition.

7.2. Steps in simulation

To simplify the Monte Carlo simulation, a total of *NG* runoff events are first generated and stored as a data file for use in the simulation. Each event is defined by random values of rainfall duration and ARI that define the average rainfall intensity, a random temporal pattern, and a random value of initial loss. These values are generated from the distributions identified in Section 6.

As a first step, values of storm-core duration d_c are generated from the fitted exponential distribution. This has one parameter estimated as the observed mean d_c value, obtained from the pluviograph data at the catchment of interest.

In the second step, a random value of storm-core rainfall intensity (I_c) for each given value of d_c is generated, using the IFD table described in Section 6.2. First a random ARI value is selected from the following equation (after Stedinger et al., 1993, Eq. (18.6.3b)):

$$\mathrm{ARI} = \frac{-1}{\ln(1 - \mathrm{AEP})} \tag{4}$$

where AEP is the annual exceedance probability, obtained from a uniform distribution $U(0,1)$. Since the primary aim is to develop derived flood frequency curves in the range of annual exceedance probabilities of say 1 in 100 to 1 in 2, the interval $U(0,1)$ is too wide. However, to cover a sufficiently wide range of rainfall intensities that might be of interest in the simulation, U was limited to the range $10^{-6} \leq U \leq 1 - e^{-\lambda}$. As an example, for an average annual number of storm-core events λ equal to 5, this results in $10^{-6} \leq U \leq 0.993$; in terms of ARI (years) this is equivalent to $10^6 \geq \mathrm{ARI} \geq 0.2$. For the given d_c and ARI values, an I_c value is then read from the IFD table for the site of interest, using linear interpolation with respect to both $\log(d_c)$ and $\log(\mathrm{ARI})$.

For the generation of a temporal pattern in the third step, the adopted simulation method randomly selects a historic temporal pattern recorded at the site of interest depending on the previously generated d_c value. The procedure is repeated *NG* times to sample *NG* temporal patterns.

In the fourth step, storm-core initial loss values are derived by first generating a storm initial loss value from the Beta distribution fitted to the IL_s data from observed events at the site of interest. The generated IL_s value is then converted to a storm-core loss IL_c, using Eq. (2). The procedure is repeated *NG* times to generate *NG* values of IL_c.

In addition to these four stochastic inputs, the simulation of streamflow hydrographs requires the following fixed inputs: (a) catchment area in km^2; (b) an estimate of continuing loss (CL) in mm/h; (c) the value of the runoff routing model parameter (k); and (d) an estimate of baseflow in m^3/s.

Finally, with the above fixed and stochastic inputs, each generated rainfall event can be converted to an input runoff hydrograph for the catchment and then routed through the single storage model to obtain a simulated flood hydrograph at the catchment outlet. The peak of each of the *NG* simulated hydrographs is stored for later analysis to determine a derived flood frequency curve. Given the parameters of the distributions of d_c, I_c, TP_c and IL_c, the generation of data files from these distributions takes about 30 min (for

20,000 events), and the simulation of streamflow hydrographs takes about 5 min on a Pentium 500 computer.

8. Flood frequency analysis

A non-parametric frequency analysis method is used to construct a derived flood frequency curve from the set of *NG* simulated flood peaks. As these flood peaks are obtained from a partial series of storm-core rainfall events, they also form a partial series. Construction of the derived flood frequency curve from the generated partial series of flood peaks involves the following steps:

1. Arrange the *NG* simulated peaks in decreasing order of magnitude and assign rank 1 to the highest value, 2 to the next one and so on.
2. For each of the ranked floods, compute an ARI from the following equation:

$$\mathrm{ARI} = \frac{NG + 0.2}{m - 0.4}\frac{1}{\lambda} \cong \frac{NY + 0.2}{m - 0.4} \tag{5}$$

   where *NG* is the number of simulated peaks, *m* the rank, λ the average number of storm-core events per year at the catchment of interest, and *NY* is the number of years of simulated flood data.
3. Prepare a plot of ARI versus flood peaks, i.e. a plot of the empirical flood frequency curve defined by the simulated flood peaks.
4. Compute flood quantiles for selected ARIs by fitting a smooth curve through neighbouring points. (Given the large number of data points, logarithmic interpolation between the two neighbouring data points, without any smoothing, has been adopted in this study.)

9. Application results

The Monte Carlo simulation technique was applied to three catchments in Victoria: Boggy Creek, Tarwin River and Avoca River (see details in Section 4). The resulting derived flood frequency curves for the three catchments are compared in Figs. 8–10 with the empirical flood frequency curve defined by the observed partial series data, plotted using Cunnane's plotting position formula. These show that the Monte Carlo simulation technique can reproduce observed flood frequency curves with reasonable accuracy over a wide range of frequency and can cope well with the non-linearity of the rainfall and runoff process. The underestimation of frequent floods in the case of Tarwin River (Fig. 9) is associated with inappropriate modelling of the substantial baseflow contribution in this catchment.

The Design Event Approach, as recommended in Australian Rainfall and Runoff (IEAUST, 1998) involves the definition of a 'design event' i.e. a single combination of design rainfall and loss parameters for a given rainfall duration and ARI (see Section 2). This approach was applied to the three study catchments to obtain another basis for comparison. The design rainfall intensities and temporal patterns from Australian Rainfall and Runoff (IEAUST, 1998) were adopted. The storm burst initial losses were obtained based on the Hill et al. (1996) recommendations (Eq. (11.2)). The same values of continuing loss and average baseflow as used to determine the derived flood frequency curves for the three study catchments were adopted here. Also, the same non-linear storage model with identical values of *m* and *k* was used for runoff routing. The results are compared in Figs. 8–10 and show that the Monte Carlo simulation technique can well reproduce the slope of the observed flood frequency curve over a wide range of frequency relative to the Design Event Approach.

The analyses indicate that, while the Design Event Approach can easily be tweaked to provide design flood estimates for a specific range of ARIs, the further one moves away from the range of events used in calibration, the greater the error in the results introduced by the effects of non-linearities in the rainfall–runoff process. By contrast, the proposed Monte Carlo simulation approach implicitly allows for the non-linearities involved and is able to faithfully reproduce the skewed nature of the frequency curve. In addition, the Design Event Approach involves the statistically flawed concept of a critical duration, and subjective smoothing of the final design flood frequency curve to overcome any inconsistencies introduced by the arbitrary selection of representative design values. In contrast, the Monte Carlo simulation technique considers rainfall duration as a random variable, and smoothing of the final flood frequency curve is not generally necessary.

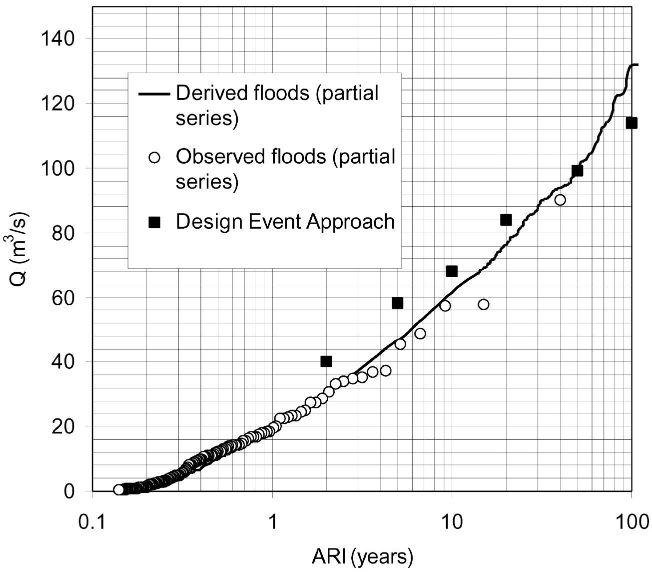

Fig. 8. Comparison of derived flood frequency curve with observed partial series and Design Event Approach (Boggy Creek catchment). Q is flood discharge.

Sensitivity analyses were carried out to examine the effect of a change in any one of the input distributions on the final derived flood frequency curve. The results indicated that a change in the parameters of the distribution of d_c and IL_c of about 10–15% would not have any significant effect on the derived flood frequency curve. It was also found that the derived flood frequency curve is moderately sensitive to the correct representation of temporal pattern variability. The input distribution of rainfall intensity had a dominant effect on the final derived flood frequency curve, as is the case for the Design Event Approach.

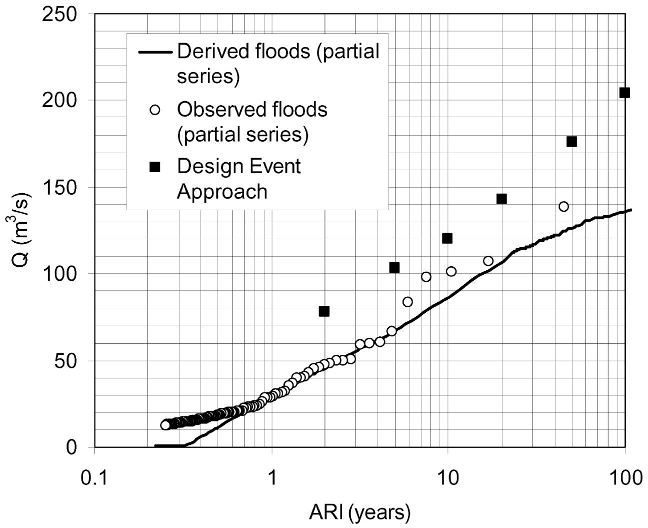

Fig. 9. Comparison of derived flood frequency curve with observed partial series and Design Event Approach (Tarwin River catchment). Q is flood discharge.

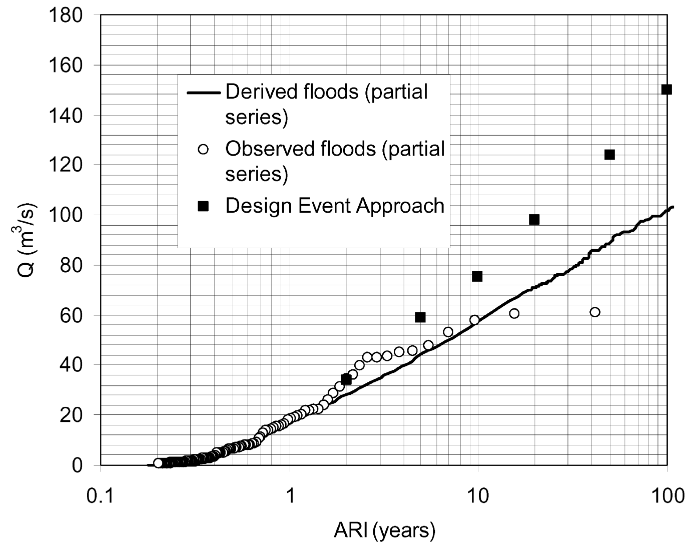

Fig. 10. Comparison of derived flood frequency curve with observed partial series and Design Event Approach (Avoca River catchment). Q is flood discharge.

The generated distribution values for d_c, I_c, TP_c and IL_c used in the initial series of simulation runs were stored in a data file to allow selective changes of individual inputs without changes in others. However, if a new set of random deviates is used for a subsequent set of simulation runs, a different set of distribution values for d_c, I_c, TP_c and IL_c will be generated, even if the parameters of the distributions remain unchanged. Thus each new series of simulation runs will result in a different derived flood frequency curve. To examine the degree of variation in the derived flood frequency curve from one series of simulation runs to another, five series of runs of 15,000 events for Boggy Creek catchment were implemented without changing the parameters of the component distributions. The results are compared in Fig. 11. They show that the simulated derived flood frequency curves are quite stable, up to an ARI of say 50–100 years.

10. Discussion and conclusion

The derived flood frequency curves for the three study catchments compare quite well with the results of flood frequency analyses of the flood series available at the sites. A relatively small degree of calibration, e.g. by adjustment of the continuing loss parameter, would produce a closer match of the derived distributions and the observed ones. The particular strength of the described Monte Carlo simulation technique is that it allows reproduction of frequency curves over a large range of ARIs, while the Design Event Approach generally focuses on reproducing frequency curves over a narrower ARI range. Given appropriate data for the catchment of interest and a calibrated runoff routing model, the new approach takes less than an hour to produce a derived flood frequency curve.

The derived flood frequency curves reflect the variability in the key input variables considered here, but not the uncertainty in parameter selection. While, the Monte Carlo simulation technique could also be used to derive confidence limits reflecting uncertainties in all design parameters, this has been left to future research efforts.

The following conclusions can be drawn from this study:

- The new Monte Carlo simulation technique based on the Joint Probability Approach offers a theoretically superior method of design flood estimation as it allows explicitly for the effects of inherent variability in the flood producing factors and correlations between them. It is readily applicable to gauged

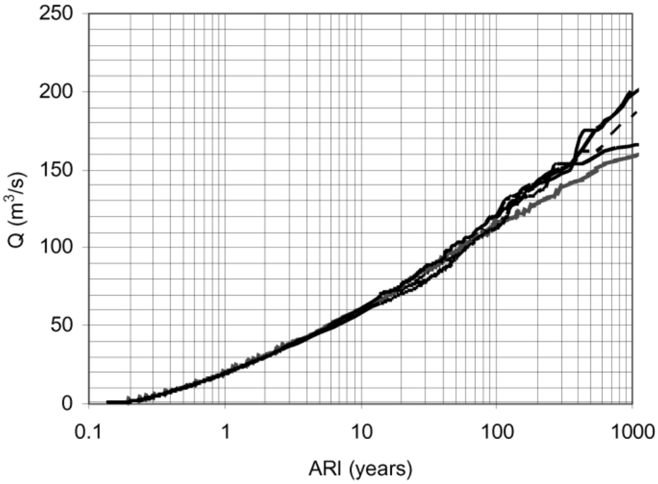

Fig. 11. Derived flood frequency curves for five series of simulation runs (Boggy Creek catchment). Q is flood discharge.

catchments with good pluviograph data and limited streamflow data.

- In contrast to the Design Event Approach, the simulation technique does not involve the unrealistic concept of a critical rainfall duration, nor any subjective smoothing of the final design flood frequency curve to compensate for any inconsistencies in the selection of representative design values.
- The new technique can reproduce observed flood frequency curves with reasonable accuracy over a wide range of frequency and this can cope well with non-linearity of the rainfall and runoff process.
- Derived flood frequency curves can be produced for a site reasonably quickly (in less than an hour provided the parameters of the distributions of the input variables are known).
- While the current research has established the strong potential of the new technique to become a practical design tool, further development work is needed before the technique can be handed over to practitioners for routine application.
- The new technique is appropriate for the derivation of flood frequency curves in the ARI range from 1 to 100 years. However, use of more reliable estimates of rare rainfall intensities (e.g. based on regional frequency analysis techniques), appropriate adjustment to other design parameters and the use of more efficient Monte Carlo sampling schemes should allow more reliable estimation of design floods beyond 100 years ARI.
- The Monte Carlo simulation technique presented here could be used to determine the derived frequency curve of other streamflow hydrograph characteristics, e.g. flood volume, time to peak. With the inclusion of a distribution of initial reservoir storage contents, the approach could easily be extended to derive the distribution of reservoir outflow characteristics.
- The technique also lends itself to estimating design floods on a seasonal basis, taking account of distinct seasonal variations of design inputs such as rainfall and loss characteristics, and the likelihood of the joint occurrence of critical values in any season. This is seen as an important direction of research to allow future improvement in design flood estimation, particularly in relation to extreme events.
- This study has been limited to situations where floods are predominantly caused by heavy storm rainfalls. Where other flood causing mechanisms, such as snowmelt, play a significant role, their contribution to flood probabilities needs to be evaluated separately and added.

Acknowledgements

The authors would like to thank the Cooperative Research Centre for Catchment Hydrology for funding the research; the Bureau of Meteorology and Thiess Environmental Services for supplying the rainfall and streamflow data, respectively; and Professor Russell Mein, Dr Rory Nathan and Dr Sri Srikanthan for support and constructive suggestions throughout the work.

References

Beran, M.A., 1973. Estimation of design floods and the problem of equating the probability of rainfall and runoff. Symposium on the Design of Water Resources Projects with Inadequate Data, Madrid, Spain, pp. 33–50.

Blöschl, G., Sivapalan, M., 1997. Process controls on flood frequency. 2. Runoff generation, storm properties and return period. Centre for Water Research Environmental Dynamics Report, ED 1159 MS, Department of Civil Engineering, The University of Western Australia.

Diaz-Granados, M.A., Valdes, J.B., Bras, R.L., 1984. A physically based flood frequency distribution. Water Resources Research 20 (7), 995–1002.

Durrans, S.R., 1995. Total probability methods for problems in flood frequency estimation, in statistical and Bayesian methods in Hydrological Sciences. Proceedings of the International Conference in Honour of Jacques Bernier, September 11–13, Chapter III, UNESCO, Paris.

Eagleson, P.S., 1972. Dynamics of flood frequency. Water Resources Research 8 (4), 878–898.

Garcia-Guzman, A., Aranda-Oliver, E., 1993. A stochastic model of dimensionless hyetograph. Water Resources Research 29 (7), 2363–2370.

Hebson, C., Wood, E.F., 1982. A derived flood frequency distribution using Horton order ratios. Water Resources Research 18 (5), 1509–1518.

Hill, P.I., Maheepala, U.K., Mein, R.G., Weinmann, P.E., 1996. Empirical analysis of data to derive losses for design flood estimation in south-eastern Australia. Report 96/5, Cooperative Research Centre for Catchment Hydrology, Department of Civil Engineering, Monash University, pp. 98.

Hoang, T.M.T., 2001. Joint probability approach to design flood estimation. Unpublished PhD Thesis, Department of Civil Engineering, Monash University.

IEAUST, 1998. Australian Rainfall and Runoff — A Guide to Flood Estimation, vols.1 and 2. Institution of Engineers, Australia, Canberra, Australia.

Laurenson, E.M., 1974. Modelling of stochastic–deterministic hydrologic systems. Water Resources Research 10 (5), 955–961.

Laurenson, E.M., Mein, R.G., 1997. RORB version 4 runoff routing program user manual. Department of Civil Engineering, Monash University, pp. 186.

Lyne, V., Hollick, M., 1979. Stochastic time-variable rainfall–runoff modelling. Hydrology and Water Resources Symposium, I. E. Australia Perth, Australia.

Muzik, I., 1993. Derived physically based distribution of flood probabilities. Extreme hydrological events: precipitation, floods and droughts. Proceedings of the Yokohama Symposium, July 1993, IAHS Publ. No. 213, pp. 183–188.

Pilgrim, D.H., Cordery, I., 1993. Flood runoff. In: Maidment, D.R. (Ed.). Handbook of Hydrology. McGraw-Hill, New York Chapter 9.

Russell, S.O., Kenning, B.F.I., Sunnell, G.J., 1979. Estimating design flows for urban drainage. Journal of the Hydraulics Division 105, 43–52.

Siriwardena, L., Weinmann, P.E., 1996. Derivation of areal reduction factors for design rainfalls in Victoria. Report 96/4, Cooperative Research Centre for Catchment Hydrology, Department of Civil Engineering, Monash University, p. 60.

Stedinger, J.R., Vogel, R.M., Foufoula-Georgiou, E., 1993. Frequency analysis of extreme events. In: Maidment, D.R. (Ed.). Handbook of Hydrology. McGraw-Hill, New York Chapter 18.

Thompson, D.K., Stedinger, J.R., Heath, C.D., 1997. Evaluation and presentation of dam failure and flood risks. Journal of Water Resources Planning and Management 123 (4), 199.

Wood, E.F., 1976. An analysis of the effects of parameter uncertainty in deterministic hydrologic models. Water Resources Research 12 (5), 925–932.