Journal of Hydrology 510 (2014) 412–423

Contents lists available at ScienceDirect

Journal of Hydrology

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

An efficient causative event-based approach for deriving the annual flood frequency distribution

Jing Li [a,*], Mark Thyer [a], Martin Lambert [a], George Kuczera [b], Andrew Metcalfe [c]

[a] School of Civil, Environmental & Mining Engineering, The University of Adelaide, Engineering North N136, North Terrace Campus, Adelaide, SA 5005, Australia
[b] School of Engineering, University of Newcastle, University Drive, Callaghan, NSW 2308, Australia
[c] School of Mathematical Sciences, The University of Adelaide, School of Mathematical Sciences, North Terrace Campus, Adelaide, SA 5005, Australia

ARTICLE INFO

Article history:
Received 24 May 2013
Received in revised form 19 December 2013
Accepted 23 December 2013
Available online 3 January 2014
This manuscript was handled by Andras Bardossy, Editor-in-Chief, with the assistance of Niko Verhoest, Associate Editor

Keywords:
Flood distribution estimation
Design storm
Rainfall-runoff process
Continuous simulation
Peak over threshold method
Derived flood frequency methods

SUMMARY

In ungauged catchments or catchments without sufficient streamflow data, derived flood frequency methods are often applied to provide the basis for flood risk assessment. The most commonly used event-based methods, such as design storm and joint probability approaches are able to give fast estimation, but can also lead to prediction bias and uncertainties due to the limitations of inherent assumptions and difficulties in obtaining input information (rainfall and catchment wetness) related to events that cause extreme floods. An alternative method is a long continuous simulation which produces more accurate predictions, but at the cost of massive computational time. In this study a hybrid method was developed to make the best use of both event-based and continuous approaches. The method uses a short continuous simulation to provide inputs for a rainfall-runoff model running in an event-based fashion. The total probability theorem is then combined with the peak over threshold method to estimate annual flood distribution. A synthetic case study demonstrates the efficacy of this procedure compared with existing methods of estimating annual flood distribution. The main advantage of the hybrid method is that it provides estimates of the flood frequency distribution with an accuracy similar to the continuous simulation approach, but with dramatically reduced computation time. This paper presents the method at the proof-of-concept stage of development and future work is required to extend the method to more realistic catchments.

1. Introduction

Flooding is one of the most frequently occurring natural hazards worldwide, and often causes major damage to our society. For example, every year in Australia, floods incur millions of dollars damage to critical infrastructure and threaten humans lives. Appropriate designs of flow regulation structures, such as dam spillways, bridges, pipelines and flood detention basins are vital for flood mitigation and the protection of important domestic and commercial resources. These designs rely on the estimation of both the frequency and the magnitude of extreme flow events. However, due to the highly variable and complex climatic and hydrological processes that drive flood extremes, it is a major challenge to provide reliable predictions.

Existing flood estimation methods can be broken down into two major groups: flood frequency analysis and derived flood frequency methods (Moughamian et al., 1987).

1.1. Flood frequency analysis

Flood frequency analysis involves fitting a distribution model to streamflow data so that the flow magnitude associated with a certain occurrence probability can be calculated using the mathematical equation of the fitted distribution. The success of the analysis depends on achieving a reliable fit for the distribution, which requires a sufficiently long and high quality streamflow record. Unfortunately it is not available in the vast majority of catchments. Furthermore if the catchment has undergone significant land-use or climate changes in the past, the historical record cannot support an accurate estimation of the flood frequency distribution.

1.2. Derived flood frequency methods

Derived flood frequency methods have been developed to overcome the limitations of flood frequency analysis. These approaches use meteorological data (rainfall, potential evapotranspiration) as inputs for a rainfall-runoff (RR) model to generate streamflow data. In general, historical rainfall data are longer and have more reliable records than streamflow data and only a relatively short streamflow record is required to calibrate the RR model. Furthermore,

* Corresponding author. Tel.: +61 8 8313 1113.
E-mail address: jing.li06@adelaide.edu.au (J. Li).

0022-1694/$ - see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jhydrol.2013.12.035

to provide projections of the impact of climate change, a weather generator can be used to simulate the meteorological data for a certain climate scenario. The simulated meteorological data is then input into the RR model to generate streamflow data, from which the flood frequency distribution (FFD) under the projected climate condition can be derived. Derived flood frequency methods are, therefore, generally preferred over flood frequency analysis, and have been developed as both analytical and simulation approaches.

Analytical methods were initiated in the early 70 s by Eagleson (1972). The author derived the peak streamflow distribution from the distributions of catchment and climate characteristics using a kinematic runoff model in an idealised V-shaped flow plane. Further development of the analytical methods was achieved by other researches, e.g., Hebson and Wood (1982); James et al. (1986) and Raines and Valdes (1993).

Recently, numerical simulation methods for deriving flood frequency distribution have undergone considerable development. These simulation techniques can be classified into two groups: continuous simulation (CS) (Calver et al., 2000) and event-based (EB) approaches (e.g. Rahman et al., 2002). CS runs a weather generator and a RR model in parallel continuously to produce a time series of streamflow data from which the flood frequency curve can be derived, while EB approaches focus on the events of interest. These usually include rainfall events and catchment wetness conditions that drive extreme flood events and are sampled from their distributions to serve as inputs for the RR model that runs in an event-based fashion. The average return intervals (ARI) of the generated flood events are associated with the ARI of the input events based on certain assumptions.

In the following, two mainstream event-based (EB) approaches, i.e., the design storm and the joint probability approaches will be reviewed, followed by a brief discussion of continuous simulation (CS).

1.2.1. Design storm approach

Among the EB methods, the most widely adopted one in the guidelines of the world practicing water resource institutions (for example, Australian Rainfall and Runoff AR&R Pilgrim, 1987) can be attributed to the design storm (DS) approach, mainly because of its simplicity. This approach involves design event rainfall generation, runoff production and hydrograph formation. It assumes that a design rainfall event of a given ARI can be converted to a design flood of the same ARI and it relies on the specification of a rainfall loss (aka antecedent soil moisture deficit) as an indicator of the catchment wetness condition. A fixed value, typically the median, is taken to represent the rainfall loss/soil moisture deficit (AR&R Pilgrim, 1987), which ignores its variability. This assumption (also referred to as the ARI neutrality assumption) can lead to significant prediction errors, as the rainfall-runoff process is basically a joint probability problem (Kuczera et al., 2003). For example, a 1 in 100 year flood can be caused by a 1 in 50 year rainfall event falling on a wet catchment or by a 1 in 200 year rainfall event falling on a dry catchment (Michele and Salvadori, 2002). Thus it is important to capture the interactions of antecedent soil moisture conditions and extreme rainfall events.

In order to overcome the problems of the ARI neutrality assumption, Camici et al. (2011) proposed to calibrate the antecedent soil moisture to the value that produces a flood with the same ARI as that of the input rainfall event. For each return period of the flood, a design soil moisture value is calibrated using the result of a long-term CS as a reference. The design soil moisture values are then regionalised as a function of the geo-morphological characteristics of the catchment so that they can be applied to ungauged catchments with similar characteristics. Given the popularity of the DS approach and its major problem of defining the antecedent soil moisture condition, the attempt to find the critical soil moisture value that maintains ARI neutrality during the transformation from rainfall to runoff seems to be practical. Walsh et al. (1991) undertook a similar study for New South Wales in Australia. However the regionalisation showed huge variability. This indicates the success of this method strongly depends on the strength of regionalisation and the quality of the data. The other significant limitation of this approach is that the design soil moisture is likely to undergo significant change under climate change conditions. The regionalised design soil moisture inputs are therefore likely to produce unreliable estimates of the FFD.

1.2.2. Joint probability approaches

To account for the joint probability nature of the estimation of extreme flood events, event-based Monte Carlo simulation techniques have been developed (Rahman et al., 2002), in which the values of the input variables, e.g., rainfall depth and antecedent soil moisture amount are sampled from either their joint or independent distribution and input into the RR model to generate a range of streamflow events. Using the *total probability theorem* the exceedance probability of these events can be estimated (Rahman et al., 2002). To reduce the computational time, stratified Monte-Carlo (SMC) techniques are used in Nathan et al. (2003), where the sampling procedure of the input variables focuses selectively on the probabilistic range of interest.

The major challenge of these techniques is to obtain the correct input distributions from the causative events of the annual maximum extreme flows that are of interest. These are very difficult to obtain because long-term historical records with many extreme events are not readily available. Moreover, catchment soil moisture conditions are not routinely measured, which requires calibrating a RR model to flood events. Currently, practical guidelines (e.g., RORB by Laurenson et al., 2010) recommend using the distribution of annual maximum rainfall and some documented rainfall loss distribution (e.g. Hill et al., 1997) estimated from short historical data to derive the annual FFD. Part of this study will evaluate the use of these practical guidelines in the EB approaches for estimating the annual FFD.

As these procedures use the annual maximum rainfall as input and take into account the joint probability of rainfall and catchment antecedent soil moisture condition, we will collectively name these methods as AMXJP methods hereafter, where *AMX* stands for annual maximum rainfall and *JP* stands for joint probability.

1.2.3. Continuous simulation

In contrast to event-based approaches, continuous simulation (CS) (Calver et al., 2000; Heneker et al., 2003) seems to solve all the problems mentioned above, under the assumption that the applied weather generator and RR model adequately simulate the rainfall-runoff process. It does not postulate ARI neutrality between rainfall and runoff, nor does it require estimation of the input distributions for an EB procedure. It simply runs a weather generator coupled with a RR model in a continuous manner to simulate a long time series of streamflow data, from which the annual maximum flows can be extracted and in turn the annual FFD can be derived.

The major limitation of the CS approach is that it is computationally demanding. For instance, as will be shown in Section 4.4.2, to get an estimate of the exceedance probability of 1 in 100 year flood with a prediction error less than 20%, the minimum length of the simulated streamflow data needs to be more than 9500 years at a daily time step. If a complicated RR model, such as a distributed and/or physically based model is required, the computational time can be prohibitive.

1.3. Contribution of this work

The main contribution of this paper is to develop a hybrid event-based approach which overcomes the limitations of current

EB approaches with a significantly reduced computational time compared with a long-term CS. This hybrid method uses a short CS run (e.g. 30–100 years) to provide input distributions into an EB approach. As this method explicitly uses concurrent input events that are the true causative events of the output flows, it is named as the *hybrid-causative events* approach (hybrid-CE). A key innovation is that the EB approach is combined with the *total probability theorem* to produce a so-called *event streamflow distribution*, which is converted to the annual FFD using the peak over threshold (POT) method. This enables improvement in the accuracy of the predictions of the annual FFD compared with the existing EB approaches, and a remarkable enhancement in computation efficiency compared with a long-term CS.

The paper is organised as follows: Section 2 outlines the hybrid-CE methodology. Section 3 presents a synthetic case study to demonstrate the advantages of the hybrid technique over the existing EB approaches mentioned above, i.e., the DS and AMXJP methods. Section 4 presents the results, which illustrate how the limitations of the DS and AMXJP methods produce significant errors in the estimation of annual FFD and then demonstrate the accuracy of the hybrid-CE method. The final part of Section 4 compares the three different approaches. Section 5 provides some discussion of relevant issues, including future research topics. Section 6 provides the summary and conclusions.

2. Development of the Hybrid-CE approach

The hybrid-CE approach combines continuous simulation and event-based approaches. A long CS of rainfall provides the rainfall distribution and a short CS of the rainfall-runoff process provides the soil moisture distribution. Together, they drive an EB simulation of the rainfall-runoff process to produce the streamflow distribution. Unlike the AMXJP method, for the hybrid-CE method the input rainfall and soil moisture values are drawn from the distributions that are estimated from causative events to produce an *event streamflow distribution*. The POT method is then applied to convert this distribution to the annual FFD.

A schematic diagram shown in Fig. 1 illustrates the interactions between different components of the hybrid-CE method. The following sections describe the three major components (continuous part, event-based part and FFD conversion part) in details. This method is generic and can be adapted to provide estimates of the distribution of extremes for the events of interest, e.g., either instantaneous peak flow rates or event volumes. For the purposes of demonstrating the value of the hybrid-CE method, we chose the simplest case study, which is to estimate daily streamflow volume extremes using daily rainfall depth and antecedent soil moisture. Section 5.2 discusses future extensions to the hybrid-CE method to estimate the more practically relevant distribution of extremes of the instantaneous peak flow rate.

In the following discussions, the capital letters R, S and Q denote the random variables representing rainfall, soil moisture and streamflow, respectively and small letters r, s and q the corresponding variates. $F()$ is used to denote the cumulative distribution function, while $f()$ is used to denote the probability density function.

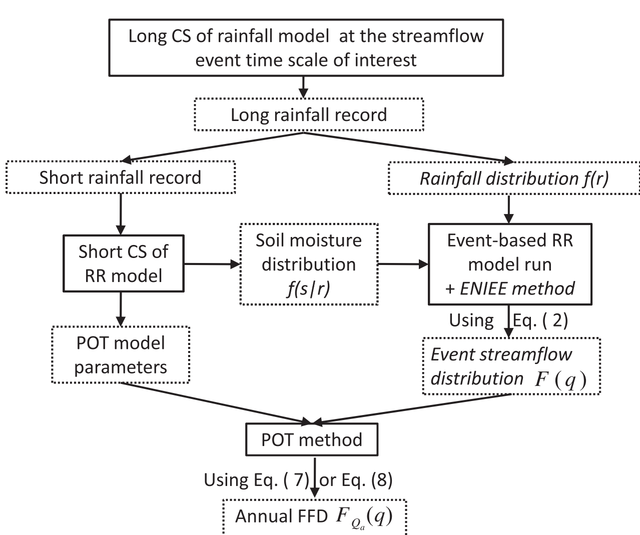

Fig. 1. Flow chart showing the procedure of the hybrid-CE method.

2.1. The continuous part

Although rainfall records are more numerous than streamflow records, they may not be available at the time scale or location of interest. In general, stochastic rainfall models (e.g. Heneker et al., 2001; Cowpertwait, 2006) can be used to circumvent limitations of rainfall records and provide the required long-term rainfall simulations.

As in the event-based part of the hybrid-CE method the rainfall distribution is needed, the continuous part of the hybrid-CE approach first runs the rainfall simulation to generate a long-term rainfall record based on the assumption that the rainfall simulation runs much faster than the RR model. The grounds for this assumption will be addressed in Section 3.5. Thus the rainfall distribution can be estimated from this long-term record which covers more extreme events than the observed data, or under climate change conditions, predicts the rainfall in the future in a probabilistic sense.

After that a short-term continuous simulation of the RR model is run using part of the generated long rainfall record as input. From this short term CS of the RR process, a short time series of soil moisture values as well as streamflow values are obtained. Given that soil moisture is less variable than rainfall, this short record of the soil moisture is sufficient for the estimation of its distribution. The short streamflow record will be used to assess the POT model parameters, as will be discussed in Section 2.3.

2.2. The event-based part

After obtaining the rainfall and soil moisture distributions, their values (r and s) can be sampled to be input into the RR model. For each EB run of the RR model, a streamflow value ($\hat{q}$) is generated. This value is compared to the streamflow value of interest (q). Note that, in general, q can be either an instantaneous flow rate at a given point in time or the volume over a given time period during which the amount of rainfall and soil moisture are accumulated. As noted earlier, we chose to adopt the simpler case of the daily flow volume to exemplify the method. A follow-up discussion on the extension of the method to estimate the more complicated case, i.e., the instantaneous flow rate, is provided in Section 5.2.

Assuming that the RR model is deterministic, with no prediction error, the conditional exceedance probability of the streamflow conditioned on the rainfall and soil moisture values, $P(Q > q|r,s)$, can be evaluated:

$$P(Q > q|r,s) = \begin{cases} 1 & \text{if } \hat{q} > q \\ 0 & \text{if } \hat{q} \leqslant q \end{cases} \tag{1}$$

In reality, RR models can have significant predictive errors due to data and model structural errors (Thyer et al., 2009; Renard et al., 2010). If a prediction error is introduced into the RR model, the value of $P(Q > q|r, s)$ will range between 0 and 1. For the current study, we assume the RR model is deterministic.

Based on the total probability theorem, the unconditional exceedance distribution $1 - F(q)$ of the streamflow can be calculated by:

$$\begin{aligned} 1 - F(q) &= \int_{\Omega_R}\int_{\Omega_S}(1 - F(q|r,s))f(r,s)dsdr \\ &= \int_{\Omega_R}\int_{\Omega_S}P(Q > q|r,s)f(r|s)f(s)dsdr \end{aligned} \quad (2)$$

where $f(r, s)$ denotes the joint probability density of rainfall and soil moisture, while $f(s)$ stands for the rainfall probability density obtained from the long-term rainfall simulation and $f(r|s)$ denotes the conditional probability density of soil moisture conditioned on rainfall, which is obtained through the short-term CS of the RR model. It is worth mentioning that if r and s are independent, $f(r, s)$ can be broken down into $f(r) \cdot f(s)$. $F(q|r, s)$ denotes the cumulative conditional distribution of streamflow conditioned on the input r and s values. $P(Q > q|r, s)$ is evaluated in Eq. (1).

The double integral in Eq. (2) can be computed through Monte Carlo integration (Davis and Rabinowitz, 1975). Nathan et al. (2003) developed a stratified Monte-Carlo (SMC) method which improves the calculation efficiency by using stratified sampling of the input values on the probabilistc range of interest.

In the hybrid-CE method, we developed an efficient numerical integration for extreme events (ENIEE) to solve Eq. (2), where the pairing of r and s is done on a grid of the domain $Dom = R \times S$. Using ENIEE Eq. (2) becomes:

$$1 - F(q_k) = \sum_i^n\sum_j^n P(q_{ij} > q_k|r_i, s_j)f(r_i, s_j)\Delta s\Delta r \quad (3)$$

Compared to the SMC technique, the ENIEE is more efficient, as the input r and s values are checked in an ordered manner so that it is easy to terminate further evaluations of the RR model at any point of (r_i, s_j) that does not contribute to the q_k value under investigation. For the SMC method, on the other hand, the program has to wait until all the random samplings within the specific intervals are finished. A detailed description of the ENIEE is provided in the Appendix.

Like the AMXJP methods, the mathematical theory underpinning the event-based part of the hybrid-CE method is also the *total probability theorem*. However the major difference lies in the fact that the AMXJP methods use the annual maximum rainfall and user-defined soil moisture events (see Section 1.2.2) to assess the input distributions for the calculation of the annual FFD. In contrast, the hybrid-CE method uses the rainfall and soil moisture events that are truly concurrent/causative to the streamflow events at the event temporal scale of interest. For example, if the event temporal scale of the streamflow is daily/hourly, then the input rainfall and soil moisture distributions will be evaluated through the daily/hourly rainfall and soil moisture events, respectively.

Hence the term $F(q_k)$ in Eq. 3 becomes the distribution of streamflows at the event temporal scale of interest (referred to as *event streamflow distribution* hereafter). Then the POT method is incorporated to convert this distribution to the annual FFD, which will be introduced in the next section.

One may argue that the *event streamflow distribution* can be directly estimated from the output streamflow data of the short CS run of the RR model and that is therefore unnecessary to use the EB simulation of the RR process and the ENIEE method. However, the short time series of the rainfall data that drive the RR model for a short CS run may not contain enough extreme events of major interest. Therefore, the short series of streamflow data generated by the short CS can lead to enormous uncertainties in the subsequent estimation of the extreme events in the annual maximum flow series, whereas in the EB component of the hybrid-CE method, the input rainfall events are drawn from the distribution which is estimated from the long-term rainfall record where more extreme events are present. Therefore the resultant *event streamflow distribution* is more reliable for use in the subsequent derivation of the annual FFD.

2.3. Derivation of the annual FFD using the POT method

The POT method (Shane and Lynn, 1964; Todorovic and Zelenhasic, 1970) is often applied in flood frequency studies as an alternative to the annual maximum series (AMS) method. A comprehensive discussion on the POT method can be found in Rosbjerg (1993). As the current study was focused on the estimation of annual FFD, we continued seeking the distribution of annual maximum flows. The POT method was adopted as a tool to derive the annual FFD from the *event streamflow distribution*.

In the POT method, the number of peaks over the selected flow threshold q_0 per year is considered as a random variable, the probability of which is denoted by:

$$P(\text{w peaks } > q_0 \text{ in a year}) = P_w \quad (4)$$

Under the assumption that the peak magnitudes are independent and identically distributed (i.i.d) with function $F(Q \leqslant q|q \geqslant q_0)$, the distribution of the annual maximum flows (Q_a) can be calculated by Todorovic and Zelenhasic (1970):

$$F_{Q_a}(Q_a \leqslant q) = \sum_{w=0}^{W}P_w(F(Q \leqslant q|q \geqslant q_0))^w \quad (5)$$

where W denotes the number of basic time steps (e.g., daily or hourly) in a year, depending on the measurement temporal resolution or the event time scale of interest. The probability distribution of the number of peaks exceeding the threshold per year (P_w) is often modelled by the Poisson distribution (Rosbjerg, 1993). However Cunnane (1979) suggests that the negative binomial distribution is more suitable for a POT series which exhibits great variability. In the current study (Section 3.5.3), it was found that a negative binomial distribution fits better to the data, hence it was adopted to the model the P_w and thus Eq. (5) becomes:

$$\begin{aligned} F_{Q_a}(q) &= \sum_{w=0}^{W}\frac{\Gamma(\gamma + w)}{w!\Gamma(\gamma)}(1-p)^{\gamma}p^w(F(Q \leqslant q|q \geqslant q_0))^w \\ &= (1-p)^{\gamma}(1 - F(q|q \geqslant q_0)p)^{-\gamma} \end{aligned} \quad (6)$$

where p and γ are parameters of the negative binomial distribution and $F(q|q \geqslant q_0)$ is the truncated distribution:

$$F(q|q \geqslant q_0) = \frac{F(q)}{1 - F(q_0)} \quad (7)$$

where $F(q)$ is the *event streamflow distribution* which was defined in Section 2.2. The denominator $1 - F(q_0)$ is a normalising factor. The problem of estimating the input distribution of annual concurrent events is therefore reduced to estimating the distribution of the input variables in accordance with the event time scale of interest. In other words, the extraction of the annual causative events from a long data series is no longer necessary and the distribution of the input variables can be much more easily obtained either through measurements or a short CS run.

2.4. Summary of the hybrid-CE approach

In summary, the hybrid-CE approach requires the following steps:

1. A long-term CS is run for the rainfall simulation at the streamflow event time scale of interest to generate a long time series of rainfall data. The rainfall distribution is estimated from this record.
2. A short rainfall record sampled from the simulated data is put into the RR model for a short-term CS run at the same event time scale to generate a series of soil moisture values for the estimation of the soil moisture distribution. The streamflow record generated by the short CS is used to estimate the POT model parameters (q_0, p and γ in Eq. (6)).
3. The RR model is run in an event-based manner using the rainfall and soil moisture values sampled from the estimated distributions and the ENIEE method is implemented to evaluate the *event streamflow distribution* using Eq. (3).
4. The POT method is applied to convert the *event streamflow distribution* to the annual FFD using Eq. (5).

The flow chart of the above steps is illustrated in Fig. 1.

3. Case study

A synthetic case study is presented to demonstrate how the assumptions underpinning the DS and AMXJP approaches impact on the estimation of the annual FFD. It also shows that the hybrid-CE approach can avoid this bias and provide more reliable estimates of the annual FFD in an efficient manner.

The rainfall data of the synthetic catchment were generated through a 1-D continuous rainfall simulation model. The simulated rainfall data were input into a lumped RR model to generate a long-term (10,000 years) sequence of daily streamflow values in order to derive the *virtual truth* annual FFD.

Simple lumped rainfall and RR models were applied in this case study, because the aim was to demonstrate the problems of the existing approaches and the relative efficacy of the hybrid-CE method. Extensions of the hybrid-CE method to a more complicated RR model using realistic catchment data will be undertaken in future research (see Section 5.2).

3.1. Rainfall simulation model

The daily rainfall simulation model consists of two parts: an occurrence model for the generation of the dry-and-wet-day sequence and a model for the generation of the rainfall amount on wet days (Srikanthan and McMahon, 2001).

The dry/wet day sequence is modelled by a first order stationary Markov chain (Weiss, 1964), the parameters of which are the initial wet-day probability P_{W0} and two conditional probabilities P_{WW} (the probability of a wet day given that the previous day was wet) and P_{DW} (the probability of a wet day given the previous day was dry).

The rainfall amount on wet days in the case study was drawn from a log-normal distribution with parameter values $\mu = 1.5, \sigma = 1.0$.

3.2. Rainfall-runoff model

The applied RR model is a simplified HBV model (Bergström, 1995) with the snow and the dual-reservoir modules omitted. The snow module was eliminated in order to illustrate a technique that focuses on extreme rainfall driven peak flow events, rather than snow-melt driven (or rain-on-snow) peak flow events, as these types of events are rare in Australia. The reservoir module was removed because this study was focussed on the frequency distribution of extreme flows. The recession part of the hydrograph which is emulated by the reservoir module is not essential to the problem.

Table 1
Summary of the annual statistics of the two climate scenarios. *CV* stands for the coefficient of variation for the annual sums.

	Annual max (mm)	Annual min (mm)	Annual mean (mm)	CV (–)	Annual POE (mm)
Dry					
Rainfall	1468.18	258.98	674.67	0.20	1277
Discharge	249.69	2.61	32.91	0.57	
Wet					
Rainfall	2452.59	928.26	1540.76	0.11	1387
Streamflow	998.98	103.62	321.06	0.27	

3.3. Climate scenarios

To test the performance of different EB approaches under different climate conditions, wet and dry climate scenarios were generated using different parameter settings for the rainfall simulator and HBV model. The selection of the parameters for the two climate scenarios was based on a comparison of the annual rainfall and runoff statistics from a database of 330 Australia catchments (Peel et al., 2000). The wet/dry climate scenario was assigned an annual mean rainfall in the upper/lower 1% of the Peel et al. (2000) dataset. Table (1) summarizes the annual statistics of the two climate scenarios.

3.4. Virtual truth reference for the annual FFD

After the model setup, a 10,000-year continuous simulation of the rainfall and rainfall-runoff models was carried out at a daily time step for both climate scenarios. As mentioned at the beginning of the case study, the output streamflow data were used to derive the *virtual truth* annual FFD, which was used to evaluate the results of the different methods tested in the following.

3.5. Input information

In EB joint probability approaches, the distribution of the input rainfall is required. In this case study, access to the long-term synthetic rainfall record (10,000 years of daily values) and a short streamflow record (e.g., 30–100 years of daily values) was assumed. The difference in the accessible record lengths was based on the assumption that the rainfall simulation would be much faster than the simulation of the rainfall-runoff process. A space–time rainfall model using the circulant embedding method and fast Fourier transformation needs just one second to simulate a 512×512 image (Qin, 2010). In contrast, it can take hours to run a 2D hydrodynamic model at a smaller spatial resolution (Neal et al., 2009).

3.5.1. DS approach

Given the ARI neutrality assumption of the DS approach, annual maximum rainfalls should be used as inputs into the RR model to derive the annual FFD. In this case study, the annual maximum rainfalls were extracted from the simulated 10,000-year daily rainfall series.

Regarding the antecedent catchment wetness condition, the primary assumption of the DS approach is that it uses a single fixed representative loss value. Typically, a rainfall loss model (e.g, proportional, initial/continuing) and a runoff routing procedure are used to convert rainfall to runoff (e.g. Laurenson et al., 2010). Traditionally the representative value of the initial loss is taken as the median of some documented distribution assessed from historical data. In Hill et al. (1997), the distribution of the initial loss is calibrated based on the rainfall events from a POT series (events with ARI greater than one year) and their concurrent flow events. The continuing loss value is determined through mass balance.

For this case study we used the simplified HBV model to convert rainfall to runoff in the DS approach because it was exactly the same RR model used to generate the *virtual truth* FFD. This enabled us to specifically test the impact of assessing a single representative antecedent catchment wetness value, without introducing errors due to the ability of the RR model to represent the *virtual truth*. Thus a single representative antecedent soil moisture (SM) value was used, as it plays the same role in the HBV model as the rainfall loss in a routing model. The rainfall threshold was evaluated based on the 10,000-year daily rainfall record. Then, with the short daily records (100 years), the soil moisture values prior to the rainfall events that are above the threshold were selected to estimate the SM distribution. Finally the median SM value was calculated from this distribution as the representative value.

3.5.2. AMXJP method

For AMXJP methods such as Nathan et al. (2003), the design guidelines, e.g., RORB by Laurenson et al. (2010), recommend that the input variables (rainfall and soil moisture) are treated as independent variables. Thus the term $f(r,s)$ in Eq. (3) becomes $f(r)\cdot f(s)$. For the rainfall distribution, the distribution of annual maxima is used (Nathan et al., 2003). In this case study, this distribution is estimated from annual maximum rainfalls extracted from the 10,000-year daily rainfall data.

Nathan et al. (2003) and Laurenson et al. (2010) recommend that the loss distribution is taken from the documented distribution as described in Hill et al. (1997). This is the same as used by the DS approach to obtain the representative value. Therefore the soil moisture distribution estimated for the DS approach in the previous section was used to test the AMXJP method in this study.

3.5.3. Hybrid-CE method

In the application of the hybrid-CE method, first the dependence between the daily rainfall depth and soil moisture amount was investigated using Pearson's ρ, Spearman's ρ and empirical copulas (Nelsen, 2006) as measures of dependence. No significant dependence was found. Therefore as in the AMXJP method, the individual distributions of rainfall and soil moisture were used. The distribution of the daily rainfall depth was directly assessed from the entire 10,000-year daily rainfall record. The distribution of the daily soil moisture conditions was estimated using the 100-year daily SM record sampled from the long-term daily SM record (10,000 years).

In addition to the input distributions, the occurrence model of the peaks over threshold and its parameters have to be specified for the POT method to convert the daily flow distribution to the annual FFD.

First of all, the peak threshold should be chosen. Rosbjerg (1987) pointed out that a flow threshold that corresponds to a yearly occurrence number exceeding 5 leads to a significant positive correlation between the peak magnitudes which violates the basic assumption of the POT method. On the other hand, too small value of the occurrence rate limits the number of events in a short record for statistical analysis. Therefore in this study, a flow threshold was chosen such that its average yearly occurrence number was 3.

Two different models of the occurrence rate of peaks were considered, the Poisson and the negative binomial distributions. Visual inspection of the frequency curves of the number of peaks per year from the 10,000 year streamflow record (not shown) showed that the negative binomial distribution provided a better fit than the Poisson to the observed data. Table 2 reports the results of the chi-squared test and confirms the above findings.

Table 2
Result of the chi-squared test for the goodness of fit of the occurrence models. 'df' denotes the degree of freedom, 'Chi-S' denotes the chi-squared test statistics.

	Poisson			Negative binomial		
	df	Chi-S	*P*-value	df	Chi-S	*P*-value
Dry	10	170270.3	0	24	22.1	0.57
Wet	10	26792.4	0	17	19.3	0.31

Therefore the negative binomial distribution was adopted. The model parameters γ and p in Eq. (6) were estimated using the method of moments (Cunnane, 1979).

In the following application of the hybrid-CE approach, the POT model parameters (q_0, p and γ) were assessed using the 100-year daily records of streamflow randomly sampled from the 10,000-year record generated by the long CS. This means for different random samples, different sets of POT model parameters were estimated.

4. Results

4.1. DS approach

Fig. 2 shows the predicted annual FFD from the DS approach for both the wet and the dry climate scenarios compared to the *virtual truth* annual FFD. Black curves indicate the *virtual truth* distributions. The light blue curves *DS-100* indicate the results using randomly sampled 100-year records (in total 100 independent records) to assess the representative SM value. In addition, in order to check the model performance in a condition free from sampling error, the entire 10,000-year record was used to derive the representative SM value, and the results *DS-10000* are shown by the dark blue curves.

The results highlight an overall under-estimation. For the very small flood values, however, the DS approach produces a slight over-estimation.

4.2. AMXJP approach

Fig. 3 shows the results of the AMXJP approach. The curves representing *AMXJP-10000* and *AMXJP-100* have similar meanings as *DS-10000* and *DS-100* described in Section 4.1. The results show an averaged good agreement with the *virtual truth*, but with relatively large estimation uncertainties.

The purple dashed lines representing the results of *JPCE-10000* indicate the outcome of the joint probability (indicated by 'JP') method using input distributions estimated from the causative events (indicated by 'CE'), i.e., rainfall and SM events that are concurrent with/prior to the annual maximum flow events. They are also in line with the *virtual truth*. The slight discrepancies are due to the fact that the JPCE approach ignores the dependence between the causative rainfall and SM events.

4.3. Hybrid-CE method

Fig. 4 shows the results of the hybrid-CE method. A relatively good agreement between the average behaviour of the predictions using the short records (*HCE-100*) and the *virtual truth* can be observed. The same applies to the predictions resulted from the use of the entire 10,000-year record (*HCE-10000*).

4.3.1. Optimal short record for the hybrid-CE method

As shown above, the predictions of *HCE-10000* by the hybrid-CE method are in line with the *virtual truth* distribution. But it relies

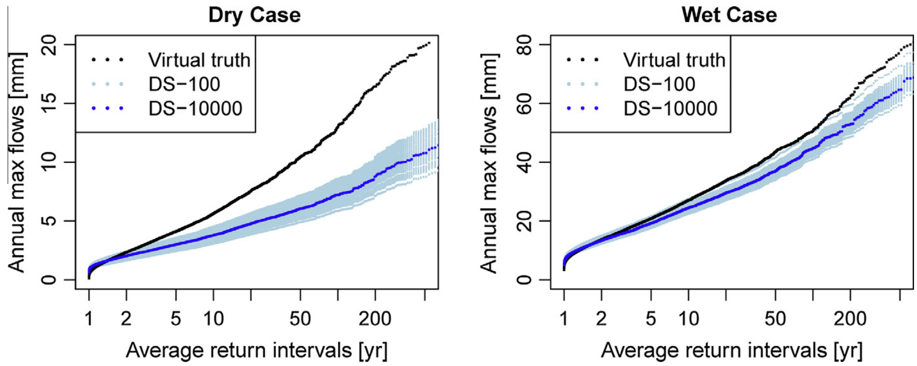

Fig. 2. Results of the DS approach. Black curves indicate the virtual truth distributions, while the light blue curves indicate the predicted distributions using randomly sampled 100-year synthetic records to derive 100 distributions of SM, to illustrate the impact of sampling error. The dark blue curves show the predicted distribution based on the SM values from the 10,000-year synthetic records to illustrate the results free from sampling error. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

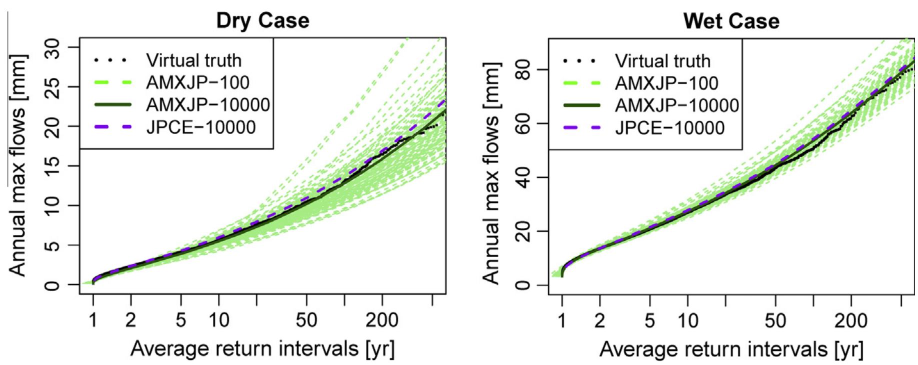

Fig. 3. Results of the AMXJP approach. Black curves indicate the virtual truth distribution, light green curves the results of using randomly sampled 100-year synthetic records to derive the SM distributions, dark green curves the results of using 10,000-year synthetic records, purple curves the results of using causative input events. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

on obtaining the input daily SM distribution and the POT model parameters from the entire 10,000-year data records. That requires a long CS of the RR model. However, as noted before, the aim of the hybrid-CE method is to avoid running a long CS of the RR model, as it can be very computationally expensive. On the other hand, when using short records generated by a short CS of the RR model for the estimation, the predicted distribution can have large or small errors compared with the *virtual truth* distribution. Therefore the question is whether certain statistics of the short record can be found which select a short record among the different random samples so that the error in predicting the annual FFD due to random sampling is minimised.

As stated in Section 3.5, it was assumed that a long-term rainfall record can be simulated. The goal here was to choose a short (30–100 years) rainfall record from the long rainfall record in order to produce a short CS of the RR model from which the best estimates of the SM distribution and POT parameters can be obtained. The selection of the short rainfall record was determined by the match between the statistical properties of the short rainfall record and those of the long record.

Several statistics (daily mean, median, standard deviation and skewness) and different record lengths were tested (30–100 years with an increment of 10 years). It was found that mean daily rainfall provided the best statistics for selecting the short rainfall record.

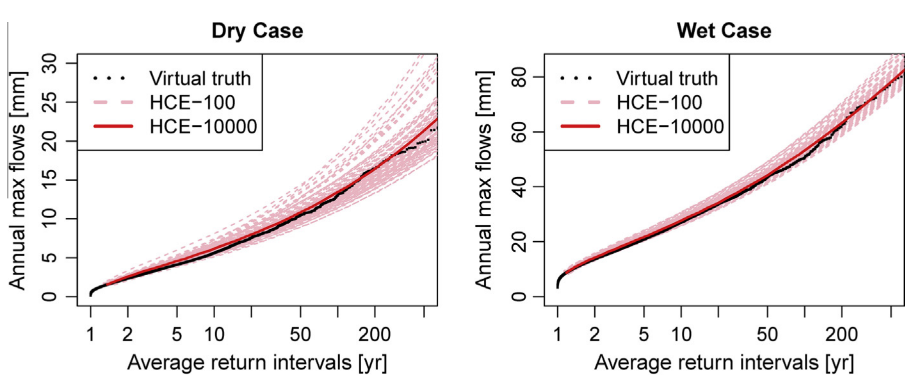

Fig. 4. Results of the hybrid-CE method. Black curves indicate the virtual truth distribution. Pink curves show the results of using randomly sampled 100-year synthetic records to assess the daily SM distributions and the POT model parameters, while red curves the results of using 10,000-year CS results. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Fig. 5 shows the results of using this approach for choosing the optimal short record for the 30, 40 and 50-year record lengths. The values of RP in Fig. 5 indicate the percentage of the random samples of short records outperform the optimal short record. These figures and the low RP values illustrate the fact that this method for choosing the optimal short record provides a good match to the *virtual truth* distribution, even for record lengths of 30 years.

4.4. Comparison of methods

4.4.1. Predictive ability

Figs. 6 and 7 compare the 95% confidence limits and the averaged results of the three methods for the dry and wet cases, respectively. They show that the DS approach produced the worst performance. There are significant under-estimations especially for the high annual maximum extreme flows. This outcome demonstrates that using a fixed representative antecedent SM value produces poor performance and highlights the importance of considering the variabilities of key input variables other than rainfall.

The AMXJP approach provided good predictive performance on average, however, it produced the largest prediction uncertainties among the three methods. This good performance was despite using arbitrarily chosen SM distributions, that were not based on the causative events. Fig. 8 compares the rainfall and SM distribution from the causative events with the distributions used by AMXJP for the dry case. This shows that the good predictive performance is due to a compensation of errors. The annual max rain distribution over-estimates the causative event rain distribution, while the AMXJP SM distribution under-estimates the causative event SM distribution (similar effect is observed for the wet case which was not shown here). The AMXJP approach relies on this compensation of errors to produce reliable predictions of the annual FFD. A relevant question is whether this compensation of errors applies only to this simplified case study and if it can be relied upon over a large range of climate and catchment conditions.

The hybrid-CE method provided good predictive performance on average except for a slight overestimation for the low flows in the dry case. The resultant estimation uncertainty was smaller

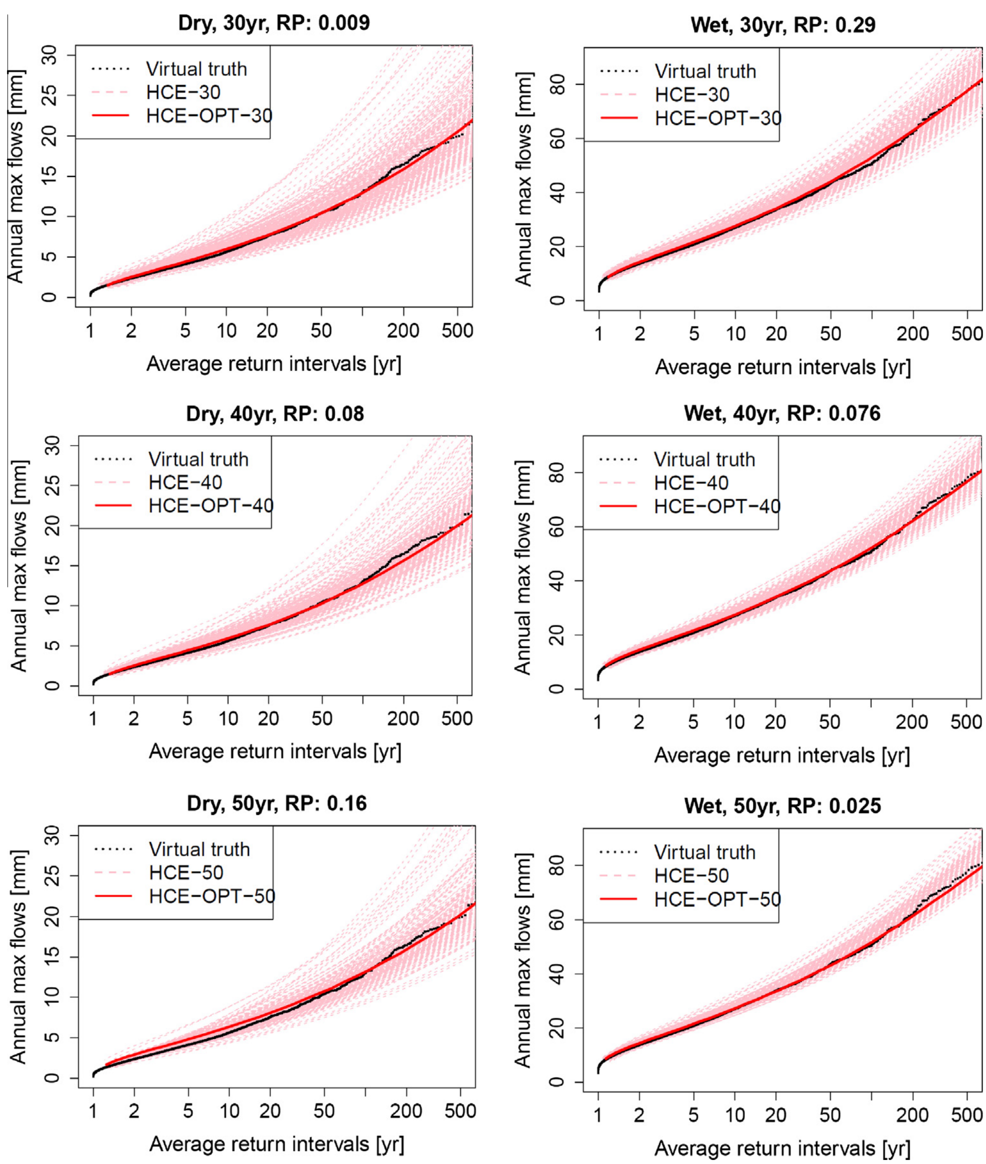

Fig. 5. Prediction of results using the selected optimal short records. Black curves indicate the virtual truth distributions, pink curves the results using randomly sampled short records, red curves the results using the optimal short records. The values of RP indicate the % of the random samples that outperform the optimal record. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

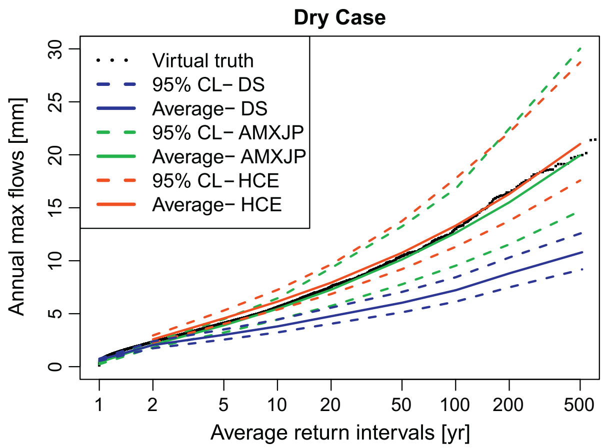

Fig. 6. Comparison of the 95% confidence limits and averaged predictions of different methods for the dry case.

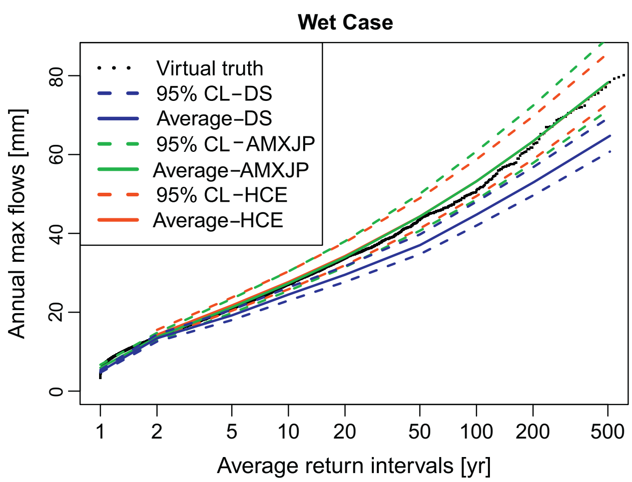

Fig. 7. Comparison of the 95% confidence limits and averaged predictions of different methods for the wet case.

than for the AMXJP approach, but higher than for the DS approach. The DS approach produces the narrowest prediction band simply because it does not take into account the variability of SM conditions unlike the other two methods. Despite the additional complexities in the hybrid-CE method (estimating input distributions and POT model parameters) when compared with the AMXJP approach, the hybrid-CE produces smaller prediction uncertainties. This demonstrates the relative robustness of the hybrid-CE method. Note that the relative uncertainty due to sampling variability is greater in the dry case than in the wet case for all three methods. This is likely to be because of the larger coefficients of variation of the rainfall and runoff data (Table 1).

A comparison of the relative prediction errors of the three different methods for different record lengths is presented in Figs. 9 and 10 which show the probability distribution of the difference in the normalised root mean square errors (NRMSE) normalised to the range of the true values from the *virtual truth* distribution. The differences in the NRMSE were calculated between the results of different methods. For example, to compare the performance of the DS and the hybrid-CE approaches, the NRMSE_DS minus NRMSE_HCE was calculated, while to compare the AMXJP and hybrid-CE approaches, the NRMSE_AMXJP minus NRMSE_HCE was calculated. A positive NRMSE difference indicates that the hybrid-CE outperforms either the DS or AMXJP. The probability distribution was based on 400, 200 and 100 independent replicates (from the 10,000-year record) for the different record lengths of 25, 50 and 100 years, respectively. The percentage of replicates with a positive NRMSE differences indicates the probability that hybrid-CE outperforms either DS or AMXJP. Figs. 9 and 10 show that hybrid-CE clearly outperforms DS (greater than 90% positive NRMSE difference for the dry case and 85% to 95% for the wet case), and also outperforms the AMXJP approach for the dry case (60–70% positive NRMSE difference), while there is only a marginal improvement in performance compared to the AMXJP for the wet case (55–60% positive NRMSE difference).

These results indicate that if a single short record is randomly selected it is likely that the hybrid-CE method will produce more accurate estimates of the annual FFD than the DS and AMXJP approaches, particularly for the dry case. In addition, Section 4.3.1 has shown that by selecting the optimal short record for the hybrid-CE method the prediction error due to random sampling of the short records is significantly reduced and the result is very close to that of using the entire 10,000-year records. Overall, these results clearly illustrate that the hybrid-CE method provides more reliable predictions than both the DS and AMXJP approaches.

4.4.2. Computational efficiency

The previous section showed that the hybrid-CE method provides more reliable predictions of the annual FFD than the DS and AMXJP methods. The main advantage of the hybrid-CE approach over the long-term CS approach is its computational efficiency. For example, to achieve a prediction error less than 20% for the exceedance probability of the 1 in 100 year flood the required number of years n to be simulated in the CS at a daily time step can be calculated according to the principle of Binomial proportion confidence interval (Brown et al., 2001):

$$1.96 \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < 0.20\hat{p} \qquad (8)$$

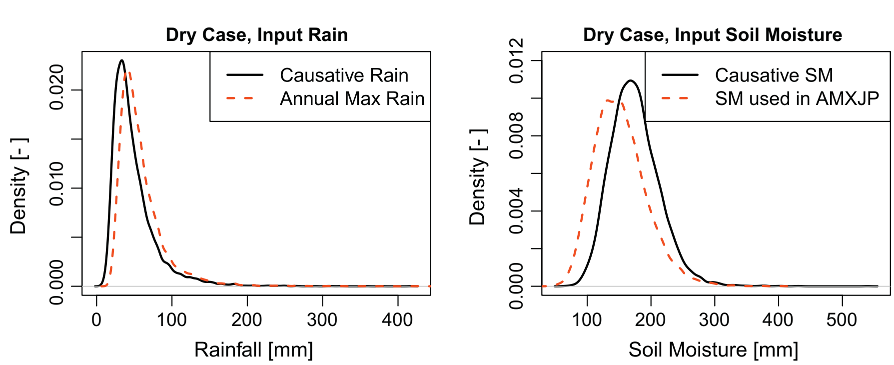

Fig. 8. Comparison of the input distributions used in the AMXJP approach and the distributions of the causative events of annual maximum flows for the dry case.

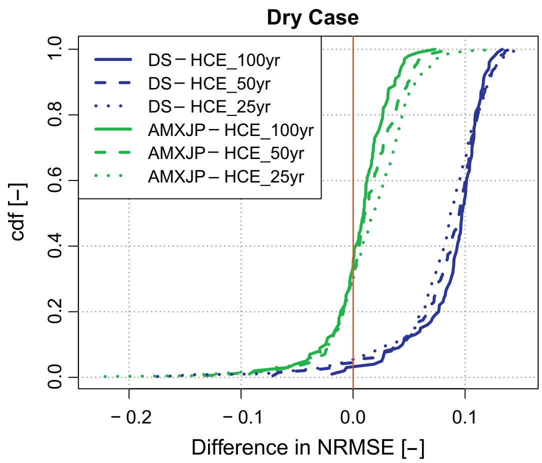

Fig. 9. Comparison of NRMSE between DS and hybrid-CE method (DS-HCE), AMXJP and hybrid-CE methods (AMXJP-HCE) for different record lengths of the dry case.

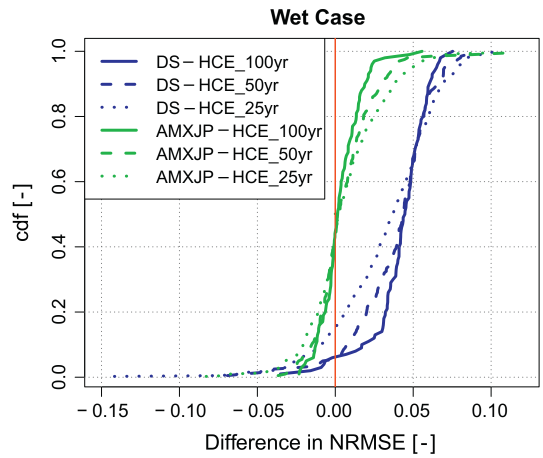

Fig. 10. Comparison of NRMSE between DS and hybrid-CE methods (DS-HCE), AMXJP and hybrid-CE methods (AMXJP-HCE) for different record lengths of the wet case.

Eq. (8) shows that to achieve the desired prediction accuracy, the number of events which the RR model must simulate in the CS at a daily time step is 3,470,420. In comparison, the hybrid-CE method needs only 1.4% of the number of events to be simulated by the RR model, if 100 years of daily CS run is required for estimating the SM distribution and the POT parameters. If only 30 years of a CS run is sufficient to get this information, the number of events which the RR model simulates can be further reduced to 0.68% of that of the long CS run. This reduction in computational time offers major advantages when a complicated distributed RR model such as TOPKAPI (Vischel et al., 2008) or HydroGeoSphere (Therrien et al., 2010) is required to estimate the annual FFD. Table 3 compares the prediction accuracy and computational efficiency of the different methods.

5. Discussions

5.1. Comparing the hybrid CE approach against existing approaches

The synthetic case study demonstrated that the hybrid-CE method outperforms the traditional DS and AMXJP event-based methods in terms of prediction accuracy. For the DS method, the ARI neutrality assumption and the use of a fixed representative SM value lead to significant under-prediction (13–46% for the dry case and 2.3–17% for the wet case on average) of the annual FFD. This under-estimation is due to a combination of assuming a fixed value of the SM and the non-linear increase in event runoff response when the SM increases. Although in practical applications it is likely that the negative bias is compensated by low biased design values of losses (high soil moisture) and possibly high biased temporal patterns of rainfall, these results should sound a warning for flood engineers who use DS approaches.

For the AMXJP method, the use of the SM distribution instead of a single value resulted in improved performance relative to the DS approach, but with a lower predictive accuracy and higher predictive uncertainty than provided by the hybrid-CE method (see Figs. 6, 7, 9 and 10). Another significant concern with the AMXJP method is that it relies on the compensation of errors arising from the use of an arbitrarily assumed SM distribution combined with the annual maximum rainfall distribution to provide good predictive performance. In the simplified synthetic case study this produced reasonable performance. However, whether this is true, in a more realistic case study, using a more realistic rainfall and RR model is an open question. A more realistic rainfall model would produce subdaily rainfall predictions, taking into account seasonally varying wet and dry spell durations and rainfall intensities (e.g. the DRIP model of Heneker et al. (2001)) and also inter-annual and multi-decadal variability (e.g. CIMSS approach of Henley et al. (2011)). A more realistic RR model would provide predictions of the subdaily flow, taking into account the non-linear spatially varying catchment processes of infiltration and soil moisture to generate baseflow, interflow and surface flow, which at any time can contribute to the flood peak (e.g. TOPKAPI, Vischel et al., 2008). Given these complexities it is unclear that assuming an arbitrary SM distribution based on a POT series of the rainfall (see Section 3.5.1) would provide reliable predictive performance across a large range of catchment and climate conditions. In contrast, the hybrid-CE is conceptually sounder because it uses the rainfall and SM distributions of the causative events that produce the streamflow events to provide efficient and reliable estimates of the annual FFD.

As mentioned in the introduction, CS has the greatest potential to provide reliable estimates of the FFD for both current and changed climate scenarios, but is the most computationally expensive method, particularly as RR models are likely to become complex in the future (e.g. TOPKAPI, Hydrogeosphere). The hybrid-CE approach is approximately 100–1000 times faster than the CS approach. Though the hybrid-CE approach does require some additional calculations related to the EINEE and POT methods, the additional computational time of these is minor compared to the computational efficiency resulting from a 100 to 1000 times reduction in the runtime of a complicated distributed rainfall-runoff model. This would further improve if parallel computing was utilised, since event based approaches are far easier to parallelise than a single long run of CS.

Given the conceptually sounder approach of using causative events and the improved predictive accuracy compared with existing EB approaches, and the vastly increased computational efficiency compared with the CS approach, the hybrid-CE approach ranks ahead of the other approaches for estimating the annual FFD. However, there is still significant work required to further develop the hybrid-CE approach in order to provide the practically relevant estimates of floods in more realistic case-study catchments.

5.2. Future development of the hybrid-CE method

The advantages of the hybrid-CE method were demonstrated in this paper using a simplified synthetic case study where the extreme daily flow volumes were estimated. Future research will

Table 3
Comparison of the performance of different methods.

Method	Relative computation time	Predictive performance
DS	1	Significant bias, accuracy worse than HCE 85–90% of the time
AMXJP	100*	Large prediction uncertainties, accuracy worse than HCE 55–70% of the time reliability based on arbitrary assumptions
CS	10^4–10^6	Minimal bias, least uncertainty
HCE	100*	Small bias

* Long enough to manage the sampling error to acceptable level.

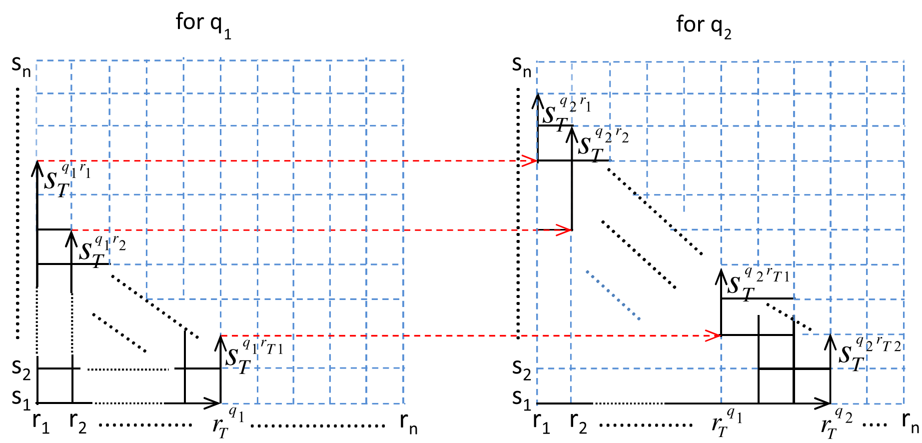

Fig. 11. Illustration of the gridding procedure for the first two flow values. Black area indicates the grid points at which the RR model is evaluated and blue area the grid points that are unnecessary to be checked. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

extend the hybrid-CE method to provide flood predictions for more realistic practical applications. Of primary interest is estimating the instantaneous peak flood rate instead of the daily flow volume. As mentioned in the previous section, this will require using a more realistic subdaily rainfall model, that takes into account spatially and temporally varying rainfall characteristics and a RR model that captures spatial variability of catchment properties and runoff-routing at the subdaily time steps. The EB component of the hybrid-CE model must be run for the entire event duration, as opposed to a single time step. As the current AMXJP method (Nathan et al., 2003), already takes into account several of these factors (seasonality, event duration modelling, temporal rainfall patterns) these existing techniques will be incorporated into the hybrid-CE approach, tested and refined as necessary. These future extensions will enable the hybrid-CE approach to provide more realistic predictions for practical applications.

One of the major assumptions of all the derived flood frequency methods is the ability of the rainfall model and RR model to properly capture the dominant physical processes which produce extreme flood events. Inherent in the development of any environmental model is the predictive uncertainty produced by data errors and model structural uncertainty (refer to Thyer et al., 2009 and Renard et al., 2010 for further discussions). These prediction errors can be incorporated into the hybrid-CE approach, by modifying Eq. (1) to be probabilistic rather than deterministic. Note the challenge is how to specify this probabilistic description given the complex, heteroscedastic and autocorrelated errors in hydrological model predictions. Research is ongoing on developing robust approaches to handle these errors, see for example Schoups and Vrugt (2010) and enhancements proposed by Evin et al. (2013).

6. Conclusions

This paper has introduced a new hybrid causative event method for providing an efficient and robust estimation of annual flood frequency distribution. The method uses a short continuous simulation of the rainfall-runoff process to provide inputs to an event-based approach for estimating the distribution of streamflow events at the time scale of interest. The peak over threshold method is used to convert this distribution to the annual frequency distribution. It successfully combines the accuracy of continuous simulation method with the efficiency of event-based methods. It takes into account the joint probability nature of the rainfall-runoff process, which avoids the potential for predictive bias in the widely adopted design storm approach. The use of causative events provides a conceptually sounder approach than the AMXJP method by avoiding reliance on arbitrary assumptions about relevant soil moisture distribution and compensatory errors. Significantly, it reduces computational demand compared with a long continuous simulation run of the rainfall-runoff model. The study reported here demonstrated the advantages (more efficient and reliable predictions) of the hybrid causative event approach over existing approaches using a simplified case study which estimated extreme daily flow volumes. Future work will extend the hybrid causative event approach to more realistic practical applications which estimate extreme instantaneous peak flows, taking into account the spatially and temporally varying characteristics of the rainfall and rainfall-runoff processes.

Appendix A. Efficient numerical integration for extreme events

The procedure of the ENIEE method is outlined as follows:

1. The range of the streamflow Q values of interest is discretised into m number of intervals. The mid points q_k of these intervals are extracted.
2. The ranges of the rainfall depth R and soil moisture amount S that are causative to the streamflows of interest are discretised into n intervals with increments of Δr and Δs, respectively. The mid points r_i and s_i are extracted.

3. The outmost loop starts from the highest value of Q, namely, q_1. For q_1, the inner loop also starts from the biggest value of R, i.e., r_1. r_1 is combined with every possible S value s_j in the innermost loop to produce a streamflow using the RR model.
4. The innermost loop also begins by first starting at the highest value s_1 and search along the S values, until the smallest streamflow which is greater than q_1 is found. The innermost loop is terminated at this point and the corresponding s_j value is recorded and denoted as $s_T^{q_1,r_1}$.
5. The R loop continues to the next value r_2 and the terminating $s_T^{q_1,r_2}$ is recorded likewise.
6. Step 5 moves onto the lower end of the R range until the smallest R value which contributes to a streamflow that is greater than q_1. The loop of R is terminated and this R value is recorded and denoted as $r_T^{q_1}$. Any R value that is smaller than $r_T^{q_1}$ will not produce a streamflow that is greater than q_1 even it is combined with the biggest S value s_1.
7. Then a set of the recorded S values $\mathbf{S_T} = (s_T^{q_1,r_1}, s_T^{q_1,r_2}, \ldots, s_T^{q_1,r_T^{q_1}})$ corresponding to all the checked R values, i.e., $r_1, r_2, \ldots, r_T^{q_1}$ is constructed.
8. The exceedance probability of q_1 is calculated using Eq. (3) for every checked pair of (r_i, s_j).
9. The Q loop moves onto q_2. For each r_i value, steps 4–8 are repeated, except that the starting point of the S loop is signified by the previously recorded $s_T^{q_1,r_i}$ and a new ending value $s_T^{q_2,r_i}$ for each r_i is recorded to replace this entry in $\mathbf{S_T}$ set for the next Q value to be checked.
10. As this procedure moves beyond the previously recorded $r_T^{q_1}$, the loop of S starts from the very beginning, i.e., s_1. The R loop continues until the smallest R value $r_T^{q_2}$ that contributes to q_2 as described in step 5. Thus the set $\mathbf{S_T}$ is updated as $(s_T^{q_2,r_1}, s_T^{q_2,r_2}, \ldots, s_T^{q_2,r_T^{q_2}})$.
11. The exceedance probability of q_2 is calculated using Eq. (3) for all the checked combinations of R and S values in this run and added by the exceedance probability of q_1 calculated before. As q_2 is less than q_1, the part of the probability exceeding q_1 does not need to be recalculated for q_2.
12. This procedure repeats for the rest of the Q values under study.

Fig. 11 illustrates this procedure. As one can see, as the evaluation moves onto the lower end of Q range, the computation accelerates as all the calculations done for the previous Q values can be used.

References

Bergström, S., 1995. The HBV model. In: Singh, V.P. (Ed.), Computer Models of Watershed Hydrology. Water Resources Publications, Highlands Ranch, CO, pp. 443–476.
Brown, L.D., Cai, T.T., DasGupta, A., 2001. Interval estimation for a Binomial proportion. Stat. Sci. 16 (2), 101–133.
Calver, A., Lamb, R., April 1–3 2000. Generalised flood frequency estimation using continuous simulation. In: Bronstert, A., Bismuth, C. (Eds.), Proceedings of the European Conference on Advances in Flood Research, Potzdam, Germany, pp. 412–421.
Camici, S., Tarpanelli, A., Brocca, L., Melone, F., Moramarco, T., 2011. Design soil moisture estimation by comparing continuous and storm-based rainfall-runoff modelling. Water Resour. Res. 47 (5) (art. no. W05527).
Cowpertwait, P.S., 2006. A spatial-temporal point process model of rainfall for the Thames catchment, UK. J. Hydrol. 330 (3).
Cunnane, C., 1979. A note on the Poisson assumption in partial duration series models. Water Resour. Res. 15 (2), 489–494.
Davis, P.J., Rabinowitz, P., 1975. Methods of Numerical Integration. Academic Press, New York.
Eagleson, P.S., 1972. Dynamics of flood frequency. Water Resour. Res. 8 (4), 878–898.
Evin, G., Kavetski, D., Thyer, M., Kuczera, G., 2013. Pitfalls and improvements in the joint inference of heteroscedasticity and autocorrelation in hydrological model calibration. Water Resour Res 49 (7), 4518–4524, http://dx.doi.org/10.1002/wrcr.20284.
Hebson, C., Wood, E.F., 1982. A derived flood frequency distribution using Horton order ratios. Water Resour. Res. 18 (5), 1509–1518.
Heneker, T.M., Lambert, M.F., Kuczera, G., 2001. A point rainfall model for risk-based design. J. Hydrol. 247 (1-2), 54–71.
Heneker, T.M., Lambert, M.F., Kuczera, G., 2003. Overcoming the joint problem associated with initial loss estimation in design flood estimation, Aust. J. Water Resour. 7 (2), 101–109.
Henley, B.J., Thyer, M.A., Kuczera, G., Franks, S.W., 2011. Climate-informed stochastic hydrological modelling: Incorporating decadal-scale variability using paleo data. Water Resour. Res. 47 (11), W11509, http://dx.doi.org/10.1029/2010WR010034.
Hill, P., Mein, R.G., Weinmann, E., 1997. Development and testing of new design losses for south-eastern Australia. In: Society, N.Z.H., on Water Engineering, A.N.C. (Eds.), Proceedings of the 24th Hydrology and Water Resources Symposium, Auckland, New Zealand.
James, W., Robinson, M., April 8–11 1986. Continuous deterministic urban runoff modelling. In: Maksimovic, C., Radojkovic, M. (Eds.), Proceedings of the International Symposium on the Comparison of Urban Models with Real Catchment Data, Dubrovnik, Yogoslavia, pp. 347–378.
Kuczera, G., Lambert, M., Heneker, T., Jennings, S., Frost, A., Coombes, P., November 10–13 2003. Joint probability and design storms at the crossroads. In: Boyd, M. (Ed.), About Water: 28th International Hydrology and Water Resources Symposium, Wollongong, Australia.
Laurenson, E.M., Mein, R.G., Nathan, R.J., 2010. RORB Version 6: Runoff Routing Program (User Manual). Monash University, Department of Civil Engineering in conjunction with Sinclair Knight Merz Pty. Ltd.
Michele, D.C., Salvadori, G., 2002. On the derived flood frequency distribution: analytical formulation and the influence of antecedent soil moisture condition. J. Hydrol. 262 (1–4), 245–258.
Moughamian, M.S., Mclauchlin, D.M., Bras, R.L., 1987. Estimation of flood frequency: an evaluation of two derived distribution procedures. Water Resour. Res. 23 (7), 1309–1319.
Nathan, R., Weinmann, E., Hill, P., November 10–13 2003. Use of Monte Carlo simulation to estimate the expected probability of large to extreme floods. In: Boyd, M. (Ed.), About Water: 28th International Hydrology and Water Resources Symposium, Wollongong, Australia.
Neal, J., Fewtrell, T., Trigg, M., 2009. Parallelisation of storage cell flood models using OpenMP. Environ. Modell. Softw. 24, 872–877.
Nelsen, R.B., 2006. An Introduction to Copulas. Springer.
Peel, M., Chiew, F., Weston, A., McMahon, T., 2000. Extension of unimpaired monthly streamflow data and regionalisation of parameter values to estimate streamflow in ungauged catchments. Report 37 pp, National Land and Water Resources Audit.
Pilgrim, D.H., 1987. Australian rainfall and runoff: a guide to flood estimation. Barton, A.C.T.: Institute of Engineers, Australia.
Qin, J., 2010. A High-Resolution Hierarchical Model for Space-Time Rainfall. Ph.D. thesis. The University of Newcastle.
Rahman, A., Weinmann, P.E., Hoang, T.M.T., Laurenson, E.T., 2002. Monte Carlo simulation of flood frequency curves from rainfall. J. Hydrol. 256 (3).
Raines, T.H., Valdes, J.B., 1993. Estimation of flood frequencies for ungauged catchments. J. Hydr. Engrg. ASCE 119 (10), 1138–1154.
Renard, B., Kavetski, D., Kuczera, G., Thyer, M., Franks, S.W., 2010. Understanding predictive uncertainty in hydrologic modelling: the challenge of identifying input and structural errors. Water Resour. Res. 46 (5), W05521, http://dx.doi.org/10.1029/2009WR008328.
Rosbjerg, D., 1987. On the annual maximum distribution in dependent partial duration series. Stoch. Hydrol. Hydraul. 1, 3–16.
Rosbjerg, D., 1993. Partial Duration Series in Water Resources. Habilitation Thesis, Technical University of Denmark.
Shane, R.M., Lynn, W.R., 1964. Mathematical model for flood risk evaluation. J. Hydraul. Div. ASCE (90), 1–20.
Schoups, G., Vrugt, J.A., 2010. A formal likelihood function for parameter and predictive inference of hydrologic models with correlated, heteroscedastic, and non-Gaussian errors. Water Resour. Res. 46 (10), W10531, http://dx.doi.org/10.1029/2009WR008933.
Srikanthan, R., McMahon, T.A., 2001. Stochastic generation of annual, monthly and daily climate data. Hydrol. Earth Syst. Sc. 5 (4), 653–670.
Therrien, R., McLaren, R.G., Sudicky, E.A., 2010. HydroGeoSphere: a three-dimensional numerical model describing fully-integrated subsurface and surface flow and solute transport. Groundwater Simulation Group.
Thyer, M., Renard, B., Kavetski, D., Kuczera, G., Franks, S.W., Srikanthan, S., 2009. Critical evaluation of parameter consistency and predictive uncertainty in hydrological modelling: a case study using Bayesian total error analysis. Water Resour. Res. 45. http://dx.doi.org/10.1029/2008WR006825.
Todorovic, P., Zelenhasic, E., 1970. A stochastic model for flood analysis. Water Resour. Res. 6, 1641–1648.
Vischel, T., Pegram, G., Sinclair, S., Parak, M., 2008. Implementation of the TOPKAPI model in south Africa: initial results from the Liebenbergsvlei catchment. Water SA 34 (3), 331–342.
Walsh, M., Pilgrim, D., Cordery, I., October 2–4 1991. Initial losses for design flood estimation in New South Wales. In: 1991 International Hydrology and Water Resources Symposium: Challenges for Sustainable Development, Perth, Australia.
Weiss, L.L., 1964. Sequence of wet or dry days described by a Markov Chain probability model. Water Resour. Res. 92 (4), 169–176.