Jul 1st, 12:00 AM

Stochastic Generation of Daily Spatial Rainfall for Regional Flood Risk Assessment

K. S. Tan
F. H. S. Chiew
R. B. Grayson

Follow this and additional works at: https://scholarsarchive.byu.edu/iemssconference


This Event is brought to you for free and open access by the Civil and Environmental Engineering at BYU ScholarsArchive. It has been accepted for inclusion in International Congress on Environmental Modelling and Software by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.
Stochastic Generation of Daily Spatial Rainfall for Regional Flood Risk Assessment

K.S. Tan\(^a\), F.H.S. Chiew\(^{a,b}\) and R.B. Grayson\(^a\)

E-Mail: kstan@civenv.unimelb.edu.au

\(^a\) Department of Civil and Environmental Engineering, The University of Melbourne, Australia
\(^b\) CSIRO Land and Water, Canberra, Australia

Abstract: Spatial rainfall is a key input into models that simulate flood behaviour at regional scale. Stochastic rainfall data provide alternative realisations that are equally likely to have occurred, and are often used to drive hydrologic models to quantify uncertainty in environmental systems associated with climatic variability. This paper describes the development and testing of a stochastic daily spatial rainfall generation approach which comprises two components. Daily temporal rainfalls in two meso-scale square regions are first generated using a bi-variate first-order transition probability matrix model. The spatial rainfall in each region is then disaggregated using a modified non-homogeneous random cascade model that utilises scaling invariance features in the historical rain field. The models are parameterised using 100 years of daily grided rainfall data across the Gippsland Lakes region in southeast Australia. The approach is used to generate 20 replicates of 100-year daily concurrent catchment average rainfall time series for the six major catchments in the region. The generated stochastic rainfalls are evaluated by comparing key spatial and temporal statistics with those in the historical data. The results indicate that the approach is suitable for regional flood risk assessment, although the simulated 1-day and 3-day rainfall AEP (annual exceedence probability) are slightly underestimated, while the simulated rainfall correlations between catchments are mostly higher than the observed spatial correlations. The main limitations of the approach are the absence of space-time correlation of rain fields on consecutive days, and problems in simulating the clustering (i.e. spatial correlation) of daily rain field during extreme storm events, both of which would require significant research to overcome.

Keywords: Regional flood risk; Spatial rainfall; Stochastic rainfall; Random cascade.

1. INTRODUCTION

Rainfall is highly variable over space (e.g. point, catchment, regional) and time (e.g. daily, seasonal, inter-annual) scales. In flood studies, daily rainfall is a key input into hydrologic models that estimate flow. The use of historical rainfall data provides only one realisation of the past climate. Stochastic rainfall data provide alternative realisations that are equally likely to occur, and can be used as input into hydrologic models to quantify uncertainty associated with climatic variability. The quantification of modelling uncertainty is central to risk-based design, system operation and environmental management decisions [McMahon et al., 1996].

Stochastic rainfall data are random numbers that are generated to preserve the statistical characteristics (e.g. mean, variance, long-term persistence, auto-correlations) of historical data. Different characteristics are important for different applications (e.g. extreme rainfall for floods, dry spell for droughts). Each stochastic realisation (replicate) is different and has different characteristics compared to the historical data, but the average of each characteristic from all the stochastic replicates should be similar to the historical data. Generation of daily rainfall at a single site is a well-researched area. However, for accessing environmental systems at the regional scale, the spatial dependence of rainfall must be accounted for, in addition to the preservation of statistical properties of rainfall series at each site. Realistic stochastic spatial rainfall that preserves at-site statistical characteristics and accounts for spatial correlations are needed as inputs into calibrated hydrologic and hydrodynamic models for regional flood risk assessment [e.g. Grayson et al., 2004; Tan et al., 2005a].

There are two main approaches for stochastic daily spatial rainfall generation: multi-site rainfall models; and space-time rainfall models. Multi-site rainfall models are extensions to single site stochastic rainfall models that simulate multi-site
rainfall concurrently using serially independent but spatially correlated uniform random numbers [e.g. Wilks, 1998]. Stochastic space-time rainfall models may be based on cluster point process [e.g. Northrop, 1998; Cowpertwait et al., 2002] or scaling-based multiplicative random cascade [e.g. Jothisityangkoon et al., 2000]. Tan et al. [2005b] compared a multi-site two-part model (one of the stochastic models in Stochastic Climate Library, http://www.toolkit.net.au/scl, a product in the Catchment Modelling Toolkit designed to facilitate stochastic climate data generation), and a transition probability matrix-random cascade model (TPM-CAS). The TPM-CAS model was found to simulate 1-day and 3-day rainfall (and flow) AEP characteristics better than the multi-site two-part model, and is therefore potentially a better model for regional flood studies.

This paper describes the application of the TPM-CAS model for regional flood risk assessment. The approach consists of two components: a daily temporal areal rainfall generation model based on a first-order transition probability matrix (TPM) method; and a daily spatial rainfall disaggregation model based on the non-homogeneous multiplicative random cascade (CAS) concept. The approach is tested using 100 years of historical catchment average rainfall (derived from grided rainfall data) across a synoptic scale region covering the Gippsland Lakes catchments. The results are assessed by comparing the key spatial and temporal statistics in the stochastic replicates with those of the historical data.

2. STUDY AREA AND DATA

The Gippsland Lakes in southeast Australia have a water surface area totalling almost 400 km² and contributing catchment area of over 20,000 km². Several of the region’s large towns are in close proximity to water edges of the Lakes and significant flooding of commercial, residential and agricultural properties has occurred, most recently in June 1998 when large parts of Lakes Entrance and Paynesville were inundated, with an estimated AUD$77.5 million damage in the entire East Gippsland Shire. Prior to June 1998, the most recent major flood occurred in April 1990 along the Mitchell River, causing an estimated AUD$18 million damage [EGSC, 2001].

Figure 1 shows two meso-scale square regions (to suit the random cascade modelling approach) of 128 km x 128 km each (i.e. 32 x 32 cells of 4 km x 4 km per cell), which are devised to adequately and tightly cover the six major catchments: the eastern region (Tambo, Nicholson and Mitchell rivers flowing into Lake King), and the western region (Avon, Thomson/Macalister and Latrobe rivers flowing into Lake Wellington). The two lakes are hydraulically separated by both the elongated Lake Victoria and McLennan Strait. This allows the two square regions to be conveniently distinguished in the flood risk modelling framework. We have used historical 0.05° x 0.05° daily grided rainfall data recorded over a 100-year period (1901-2000) [QDNRM, 2000] in this investigation.

![Figure 1. The Gippsland Lakes catchments with two meso-scale square regions.](image)

3. MODEL DESCRIPTION

3.1 Transition Probability Matrix (TPM)

A daily stochastic temporal areal rainfall model is used based on the first-order transition probability matrix (TPM) model of Srikanthan and McMahon [1985]. TPM models are an extension of the Markov chain concept to multi-state models. The first state is a zero rainfall state followed by several non-zero intermediate rainfall states, with the largest rainfall state unbounded. Transition probabilities among all possible pairs of states and the distribution parameters of the largest rainfall state are estimated from the data, and subsequently used in the generation. The performance of TPM models lies in the choice of the number of states, their upper and lower thresholds, and the distributions used for the intermediate states and the largest states. These involve relatively large number of parameters, hence a long historical record is required to estimate the parameters reliably [Wilks and Wilby, 1999]. Several researchers have modelled daily rainfall at a single site successfully using TPM method [e.g. Allen and Hann, 1975; Selvatingam and Miura, 1978; Gregory et al., 1993]. Chapman [1994, 1998] compared the Srikanthan and McMahon [1985] TPM model with the best-selected two part models and found that the TPM model generally performs better (in producing the mean, standard deviation, skewness and number of wet days).
In this study, the TPM model is modified to generate daily areal (regional) average rainfall ($R_d$) in two adjoining meso-scale regions concurrently. $R_d$ can occur in one of the ten rainfall states: state 1 is dry ($R_d \leq 0.1$ mm/day), states 2 to 9 are intermediate states with lower and upper bounds, and state 10 is the largest state with no upper bound ($R_d > 25$ mm/day). A linear distribution is used to model $R_d$ for states 2 to 9. A shifted Gamma distribution is used to model $R_d$ in the unbounded largest state in the eastern region, while $R_0$ in the unbounded largest state in the western region is generated based on the relationship between the eastern and western rainfall so that cross-correlations between extreme rainfalls in the two regions are preserved. A nonparametric technique is used to generate (by resampling) a western $R_0$ if the eastern $R_0$ is within the range of historical record in that calendar month, otherwise a hybrid regression/non-parametric technique is used to generate (by extrapolation) a western $R_0$ (Figure 2). The third largest historical eastern $R_0$ is used to define the upper limit of the range of historical record (vertical line in Figure 2). This is a compromise to maximise the use of historical east-west rainfall relationship for resampling the western $R_0$, and to avoid a few extreme eastern $R_0$ (which can be much larger than most of the other eastern $R_0$) from dominating the resampling.

**Figure 2.** Relationship between eastern and western regional average rainfall for July.

The parameters in the model, which are estimated from the historical data, are: transition probabilities of being in a particular rainfall state given the state on the previous day; two Gamma distribution parameters for the largest rainfall state; and linear regression parameters describing the east-west rainfall relationship in the largest state. The seasonality in occurrence and magnitude of daily rainfall are taken into account by considering each month separately (three-running month is used here). The Boughton [1999] adjustment is used to reproduce the rainfall inter-annual variability.

### 3.2 Random Cascade (CAS) Model

A non-homogeneous multiplicative random cascade (CAS) model is used here to disaggregate the generated daily regional average rainfall ($R_d$) into a daily spatial rain field. Stochastic multiplicative random cascade models utilise scaling invariance features such as extreme variability and strong intermittence seen in the observed rain fields to model space-time rainfall [Lovejoy and Schertzer, 1990; Gupta and Waymire, 1990]. Theoretical arguments and empirical evidence suggest that spatial and temporal organisation of rain fields tend to exhibit self-similarity in their patterns at different scales, and can be modelled within the multifractal framework [Seed, 2003]. This self-similarity property enables parsimonious parameterisations of rain fields over a wide range of scales, hence circumventing the problem of separate parameterisation at each scale in the cluster point process approach [Lovejoy and Schertzer, 1990].

The conceptual basis of multiplicative random cascades originates from the turbulence theory, where a cascade of turbulent eddies is seen as transferring kinetic energy from a large energy scale progressively to smaller dissipation scales [Over and Gupta, 1996]. The analogy to rainfall is that the total mass of rainfall is disaggregated in a scaling hierarchical manner (Figure 3), such that an area of higher intensity is embedded in larger areas of lower intensity, which are part of even larger structures of even lower intensity.

![Figure 3. Schematic of a 2-d construction of discrete multiplicative random cascades.](image)

The CAS model used here is a modified version [Tan, 2004] of the non-homogeneous multiplicative random cascade model of Jothityangkoon et al. [2000], which improves the realism of simulated rain fields, notably for extreme rainfall events. To estimate the model parameters for each square meso-scale region, all historical daily rain fields for each month are grouped according to the historical daily regional average rainfall ($R_0$), $6 \leq R_0 \leq 25$, and $R_0 > 25$ mm/day). Monthly systematic variations in spatial rain fields (e.g. due to orographic effects) for each grouping are extracted as the non-homogeneous spatial deterministic
layers. An appropriate spatial deterministic layer is then applied to each historical daily rain field (depending on \( R_0 \)) to obtain the daily residual (random) spatial rain field, from which the monthly multiplicative random cascade parameters (\( \beta \) and \( \sigma^2 \)) for each day are estimated and plotted against \( R_0 \) (i.e. \( \beta - R_0 \) and \( \sigma^2 - R_0 \), see Figure 4).

**Figure 4.** Relationships between the cascade parameters and western regional average rainfall.

In the simulation mode, the spatial rain field for each day in each square meso-scale region is simulated using the monthly cascade parameters resampled from the historical \( \beta - R_0 \) and \( \sigma^2 - R_0 \) relationships using a non-parametric k-nn resampling approach [Lall and Sharma, 1996; Tan, 2004] based on the \( R_0 \) generated by the TPM model. Appropriate spatial deterministic layer (depending on \( R_0 \)) is then re-applied to the simulated rain field to account for orography (while preserving the simulated regional rainfall amount). The resampling approach ensures that the variability of cascade parameters seen in the empirical relationships is preserved, instead of being smoothed away using the 'best fit' curves, hence leading to more realistic rain field simulation.

**4. RESULTS**

Figures 5 and 6 show the AEP plots and spatial correlation plots (for July) of 1-day and 3-day regional average rainfall in the eastern and western regions for the historical and generated (50 replicates of 100-year) data using the TPM model.

**Figure 5.** AEP plots of historical (solid) vs. generated (hollow) regional average rainfall.

**Figure 6.** Spatial correlations of historical (solid) vs. generated (hollow) regional rainfall for July.

Figure 7 shows typical AEP plots of 1-day and 3-day catchment average rainfall between the Tambo/Nicholson and Latrobe catchments for the historical and simulated (20 replicates of 100-year) data using the TPM-CAS model. Results for other catchments are similar and are not shown here. The spatial correlations (for July) of 1-day and 3-day rainfall (for total rainfall >10 mm and >25 mm, respectively) between the six major Gippsland Lakes catchments for historical and simulated data are given in Table 1.

**Table 1.** Spatial correlations of historical (his) and simulated (sim) catchment average rainfall for July.

<table>
<thead>
<tr>
<th></th>
<th>Tambo/Nicholson</th>
<th>Mitchell Up</th>
<th>Mitchell Low</th>
<th>Avon</th>
<th>Thomson/Macalister</th>
<th>Latrobe</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-day (&gt;10mm)</td>
<td>His 0.23</td>
<td>0.80</td>
<td>0.91</td>
<td>0.63</td>
<td>0.31</td>
<td>-0.14</td>
</tr>
<tr>
<td>3-day (&gt;25 mm)</td>
<td>Sim 0.80</td>
<td>0.91</td>
<td></td>
<td>0.79</td>
<td>0.19</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tambo/Nicholson</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mitchell Up</td>
<td>0.30</td>
<td>0.73</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mitchell Low</td>
<td>0.70</td>
<td>0.89</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avon</td>
<td>0.81</td>
<td>0.89</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thomson/Macalister</td>
<td>0.33</td>
<td>0.80</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Latrobe</td>
<td>-0.04</td>
<td>-0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note: Correlation values indicate the strength of the relationship between the historical and simulated data.*
Figure 8. Typical observed and simulated daily spatial rain fields of a moderate/heavy rainfall day.

Figure 8 shows the simulated daily spatial rainfall pattern of a typical moderate/heavy rainfall day across the whole modelled region, in comparison with the historical daily spatial rainfall pattern for a similar rainfall condition of the same month.

5. DISCUSSION

5.1 Annual Exceedence Probability

For meaningful flood risk assessment, a key feature that must be preserved in a stochastic rainfall generation model is the extreme rainfall characteristics. Figure 5 shows that the TPM model reproduces the 1-day storm characteristics satisfactorily for both regions. The model also simulates the 3-day regional rainfall satisfactorily for the eastern region, but slightly underestimates the 3-day regional rainfall in the western region. This underestimation may be due to the abrupt changes in the historical 3-day western regional rainfall AEP curve to a steeper slope for events rarer than 5% AEP. One possible explanation could be that historically, some of the larger 3-day extreme rainfall events in the western region resulted from consecutive high daily rainfall arising from the largest rainfall state ($R_c > 25$ mm/day). In this case, the TPM model is unable to accommodate serial correlation in the daily rainfall generation if rainfall on consecutive days comes from the same rainfall state.

The 1-day and 3-day annual maxima catchment average rainfall AEP curves for Tambo/Nicholson and Latrobe (Figure 7) show that the TPM-CAS model reproduces catchment storm characteristics satisfactorily. Results for the other four catchments are similar (not shown here), but with the overall slight underestimation in 3-day western regional average rainfall (from the TPM model) accentuated in the catchment average rainfall. The generated storm characteristics in the Avon (not shown here) are also more attenuated, probably because Avon is located between the two regions in which the regional average rainfall are disaggregated into spatial rainfall separately in the CAS model.

5.2 Spatial Correlation

Spatial correlations at short time scales (e.g. daily, 3-day total) are important for regional flood studies, while correlations at longer time scales (e.g. annual) are important for regional water resources assessment. Figure 6 indicates that the TPM model can generate 1-day and 3-day regional rainfall spatial correlations satisfactorily. When the generated regional average rainfall are subsequently disaggregated into spatial rainfall using the CAS model, the simulated 1-day and 3-day catchment average rainfall spatial correlations (Table 1) are mostly higher than the correlations in the historical data (only simulated correlations of Thomson/Macalister with other catchments are lower than the correlations in the historical data). This is expected, as the multiplicative random cascade approach is known to produce simulated rain fields that tend to decorrelate too quickly and hence appear to be less clustered (Tan et al., 2005b, see also discussion in Section 5.3). This also leads to more simulated rainy days with light drizzle than the historical rainy days at the catchment scale. For these reasons, CAS will always overestimate spatial correlation with poorer stochasticity.

5.3 Daily Spatial Rainfall Pattern

The simulated daily spatial rainfall patterns should look like the historical rain fields for similar rainfall conditions in the same season/month. Figure 8 compares a typical simulated and a historical daily spatial rainfall pattern of a moderate/heavy rainfall day in April. The figure shows that spatial characteristics such as spatial patchiness and non-homogeneity in the rain field are captured reasonably well, especially if one were to look at the large-scale pattern rather than the small-scale details. In principle, the simulated rain field images should be interpreted in a stochastic sense, rather than as a prediction of future events.

Since there is no linkage along the model boundary between the two meso-scale regions, the inability of the spatial rainfall disaggregation technique in maintaining a smooth transition across the two regions leads to apparent discontinuity along the common boundary (especially on days with contrasting simulated $R_c$ in the two regions). Apart from this, the simulated rain fields also tend to decorrelate too quickly and hence appear to be blocky. A similar problem has
been encountered and reported by other researchers. For example, using radar-observed rain fields at Darwin, Australia, Seed et al. (1999) applied a multiplicative random cascade model and found that reasonable fit was achieved for 10-min instantaneous and hourly spatial correlation functions, but the simulated daily rain fields tend to decorrelate too quickly.

Serial correlation is captured during the generation of daily rainfall at the regional scale in the first-order TPM model. The serial correlation is propagated into the simulated daily rainfall at the catchment scale since the CAS model only disaggregates the generated daily regional rainfall into daily spatial rain field. However, the lack of daily space-time correlations (no memory between spatial rain fields over consecutive days) in the CAS model is a challenge for further research into space-time coupling in daily rain field simulation. The inability of the CAS model to simulate the clustering (i.e., spatial correlation) of a daily rain field during storms (within the same day) could be the reason why the simulated extreme daily rainfall in the catchment scale is being underestimated.

6. CONCLUSIONS

This paper describes the development and testing of a stochastic daily spatial rainfall simulation approach for regional flood risk assessment. The approach consists of two components: a daily temporal area rainfall generation model (TPM) based on the first-order transition probability matrix; and a daily spatial rainfall disaggregation model (CAS) based on the non-homogeneous multiplicative random cascade.

The TPM-CAS approach is tested using 100 years of historical catchment average rainfall across a synoptic scale region covering the Gippsland Lakes catchments in southeast Australia. The results are assessed by comparing the key spatial and temporal characteristics in the stochastic replicates with those of the historical data among the two meso-scale regions and the six major catchments. The characteristics assessed are: 1-day and 3-day rainfall AEPs and spatial correlations, and typical daily spatial rainfall pattern for a moderate/heavy rainfall day. The results indicate that the TPM-CAS approach can be used for regional flood risk assessment, although it slightly underestimates the AEPs and simulates higher spatial correlations in most of the catchments tested. It also reproduces daily spatial storm patterns, which appear to be less clustered. The main limitations of the TPM-CAS model are the absence of space-time correlation of rain fields on consecutive days, and problems in simulating the clustering (spatial correlation) of daily rain field during extreme storms, both of which require significant research to overcome.

7. ACKNOWLEDGEMENTS

This work was funded by the Australian Research Council Linkage Grant (C00106804), East and West Gippsland Catchment Management Authority, Gippsland Ports, Gippsland Coastal Board, Dept. of Sustainability and Environment, and Wellington Shire. Initial assistance by Dr. Senlin Zhou in providing the computer code of non-homogeneous multiplicative random cascade model (based on Jothityangkoon et al. [2000]) is acknowledged.

8. REFERENCES


