Generating projected rainfall time series at sub-hourly time scales using statistical and stochastic downscaling methodologies

S. Molavi 1, H. D. Tran 1,2, N. Muttil 1

1School of Engineering and Science, Victoria University, PO Box 14428, Melbourne, VIC. 8001, Australia; 2Institute of Sustainability and Innovation, Victoria University, Melbourne, Australia
Email: shahram.molavi@live.vu.edu.au; dung.tran@vu.edu.au; nitin.muttil@vu.edu.au

Abstract
Many climate studies in the recent past have revealed an obvious variation in climate compared to the past. Recent extreme events such as flash flooding, bushfires and drought provide ample evidence of these variations. In addition to the natural cycle of the climate, the anthropogenic effects of human development are no longer negligible. Emission of greenhouse gases and other aerosols into the atmosphere have led to warmer temperatures and consequently to more extreme events.

Urban stormwater systems will particularly be influenced by climate change. Conventionally, to design and develop the stormwater collection systems, it has been assumed that events that occurred in the past would happen in the future. The change in climatic patterns has led to a new approach of considering the future variation in climate for the assessment of stormwater systems.

In this study statistical and stochastic approaches to downscaling climate variables from Global Climate Models (GCMs) are discussed and a convenient approach to downscale daily rainfall data into sub-hourly timescales is presented. The Statistical Downscaling Model (SDSM) has been selected to downscale GCMs spatially at the site location. SDSM provides results in a daily time base. A disaggregation approach has been modified and simplified to generate sub hourly time scale rainfalls from daily rainfalls. Availability of these data is essential for the assessment of stormwater system functionality against future variability. This research is an ongoing attempt to develop a new statistic stochastic approach to increase the accuracy of the model in re-sampling of the observed data especially at the sub-hourly time scale.

Key Works: Climate Change, Global Climate Models, Statistical Downscaling Model SDSM, Stochastic Disaggregation.

Introduction
Urban drainage systems have been designed and developed to handle large amounts of rainfall that is converted to stormwater runoff after a catchment reaches its infiltration capacity. The anthropogenic effects of imperviousness of catchments, mostly due to urbanization, have led to stormwater being concentrated and collected more rapidly. Conventionally, the recorded historic data of rainfall and flows from the catchment have been treated as a reliable basis for the design of stormwater collection system infrastructure. The design rainfalls applied were obtained through frequency analysis of the stationary
rainfall data. Nowadays, the measured hydro-climatic variables provide us with many anomalies when compared to the past recorded data. This evidence of the anomalous climatic events can be considered that the climate is changing.

Different methods have been used to implement the effects of climate change in impact assessment studies. Estimates of future climate change scenarios represent a change between present-day climate scenarios and some future plausible climate scenarios. The creation of future climate change scenarios has been undertaken using Global Climate Models (GCMs), analog studies, sensitivity analyses, and trend analyses (Denault et al., 2006). Except for GCMs, these methods rely highly on recorded data and thus are unable to provide an image of driver of change factors in the future. GCMs simulate climate using 3 types of information, namely, socio-economic data, climate observations and other environmental variables. Six emission scenarios have been derived from the Special Report on Emission Scenarios (SRES) representing different future technological and economic developments that might influence greenhouse gases (Nakicenovic et al., 2000). SRES is implemented in GCMs for describing socio-economic development.

The information on climate change that is available in GCMs is coarse in spatial and temporal resolution and thus modellers are unable to resolve important sub-grid scale features such as rainfall and temperature (Wilby et al., 2002). The resolution of GCMs is probably hundreds of kilometres. Several techniques have been used to spatially downscale GCM results to a smaller regional scale. Regional Climate Modelling or RCM, also known as Dynamic Downscaling uses outputs from global models to produce initial conditions and time-dependent meteorological boundary conditions. These models are mostly complex and computationally expensive (Maerens et al., 2003). Fine spatial resolution also could be obtained from statistical methods, in processes known as Statistical Downscaling. Regional or local climate information is derived by first determining statistical models that relate large-scale climate variables (or “predictors”) to regional and local variables (or “predictands”) (Wilby et al., 2004). To develop a statistical downscaling methodology, the stationary data and output are re-gridded to GCM scale in order to simulate local climate variables (Hay et al., 1991). Weather classification methods and regression models are major statistical downscaling methodologies (Wilby et al., 2004). An advantage of statistical downscaling methods is their low computational demand; nevertheless they are highly sensitive to empirical relations between predictors and predictands (Wilby et al., 1998, Winkler et al., 1999).

Another issue with the direct use of GCM outputs is the temporal resolution. Particularly in stormwater drainage systems, it is essential to specify rainfall events in suitable time intervals. To estimate the time of concentration of runoff from an urban catchment, time intervals in minutes are needed to define fast cumulating water hydrographs. Simple statistical methodologies are mostly unable to produce strong correlations between large scale climate variables at a large time scale (e.g. daily and monthly) and weather variables at sub-hourly time scales. In this paper, an approach for downscaling using conventional methodologies is presented. In this approach, which will be explained further in two stages, the data from GCMs have been downscaled and disaggregated.
Methodology

The methodology used in this study consists of two major stages: (1) a spatial downscaling methodology to downscale low-resolution climate variables into local scale; and (2) disaggregation of the daily rainfall into half-hour time scale for current and future periods under different climate change scenarios. Figure 1 shows the flowchart of the methodology. In this paper, the model has been crosschecked using 2009 rainfall data from Moorabbin airport rain gauge station.

The Statistical Downscaling Model (SDSM) (Wilby et al. 2002) has been used in the first stage for spatial downscaling. This aims to develop a regression based model between global scale climate variables provided by GCMs with daily rainfalls at a local scale.

Temporal downscaling is later applied in the second stage using disaggregation to generate half hourly rainfall time series from daily rainfall series.

Stage 1 SDSM spatial downscaling

SDSM is a hybrid of a stochastic weather generator and a regression-based downscaling method. Large-scale circulation patterns and atmospheric variables are used to linearly regress to the local-scale weather generator parameters (e.g. precipitation occurrence and intensity) (Wilby et al., 1998). Four discrete processes are involved: (1) screening of predictors; (2) model calibration; (3) synthesis of observed data; (4) generation of climate change scenarios. Details of the computations in SDSM can be found in Wilby et al. (2002) and Karamouz et al. (2010).

The present and future rainfall time series would be synthesised once Stage 1 is undertaken. The application of the two stages of the methodology is presented in the following sections.

Springvale, within the City of Greater Dandenong is selected as the case study site. The GCM climate variables and at-site daily rainfall data available at Moorabbin Airport rain gauge have been collected. The selected GCMs are HadCM3 A2 and HadCM3 B2 (Pope et al, 2000) for the 1960-1990 period as the baseline and future periods of 2011-2040, 2041-2070 and 2071-2099. A2 and B2 represent the potential future development in greenhouse gas emissions (Nakicenovic et al, 2000) in a high and mild rate respectively.
The details of the computations and the results of the Stage 1 process can be found in Molavi et al. (2011).

Figure 2 presents the results of Stage 1, which show the percentage difference between the synthetic future rainfall and the present or base line rainfall. From the figure, it can be seen that undertaking the A2 scenario would project a dramatic decline (down to -14%) in average rainfall. B2 shows some positive anomalies, although the trend is decreasing, but there are still some increasing anomalies, particularly in May and August.

As mentioned above, the SDSM methodology is only capable of downscaling the climate variables spatially from large grids to a local scale (i.e., the exact location of the rain gauge). The precipitation is projected at a daily timescale. To downscale the timescale from daily to sub-hourly, a stochastic methodology has been modified and applied, as presented in the next section.

**Stage 2 - Disaggregation using stochastic downscaling**

As shown in Figure 3, SDSM will provide the rainfall time series at a daily timescale. This rainfall time series is, however, too coarse to be used for urban flood assessment. In this stage daily rainfall is disaggregated to sub-hourly time scale.
There are various disaggregation models available in the literature, for instance, the Bartlett-Lewis Rectangular Pulses (BLRP) model (Rodriguez-Izarbe et al., 1987) and the Disaggregated Rectangular Intensity Pulse (DRIP) model (Heneker et al. 2001). Both these models more or less carry out a stochastic approach to disaggregate rainfall values from a larger to a smaller timescale. The first model (BLRP) assumes that each rainfall event and associated inter event pulses follow a Poisson process. The duration of each event and pulse are assumed to be exponentially distributed. The intensity of each pulse is randomly generated from a fitted probability distribution. The BLRP model should be calibrated to the site of a study using sub-hourly samples. In the other model, DRIP, the main characteristics of the rainfall are assumed to happen stochastically. The duration of the storms, the durations between storms and the average intensity of the storms are generated through a random process using Generalised Exponential Distributions. In the next step, the average intensity is disaggregated into the proper scale using a random walk process.

The time step to be applied in this approach is half an hour. Using equal time steps for the random pulses simplifies the approach. Generating the wet periods of the events through a random process automatically produces the number of rainfall pulses. Figure 4 illustrates the stochastic assumptions of the disaggregation. As shown in the figure, the Dry Duration between two independent events, the Wet Duration or duration of an event, and the number and values of the pulses in each event are all assumed to follow a random process.

According Heneker et al. (2001), any dry duration of more than 2 hours distinguishes two independent wet events. As the time steps of the pulses are 30 minutes, it is assumed that there would not be any pulse with a zero value in each event. To generate a random process for each variable occurrence, it is essential to fit each variable along with a marginal Probability Distribution Function (PDF). For this
reason, four probability distributions, the Normal, GEV, Weibull and Gamma distributions, have been fitted to a sample of 30 minutes rainfalls from Moorabbin Airport Rain Gauge for the year 2010.

Figure 5 shows the marginal probability distributions fitted to the 30 minute rainfall pulses. Only the last three months of 2010 have been illustrated. Each month was independently analysed to allow for the effects of seasonality. The Weibull and Gamma distributions provide the best matches to the data. A similar approach was carried out for Wet and Dry Durations.

Figure 5 Marginal fits of the candidate probability distributions for three monthly periods

Figure 6 presents the probability distribution fits for Wet and Dry durations. As seen in this figure, Wet Durations are matched best by the Gamma distribution, but Dry Durations are matched best by the Exponential distribution.
The parameters of all the probability distribution functions have been computed and derived using MATLAB’s built in Probability Distribution Fit Tool. Figure 7 presents the flowchart for the rest of the procedure to disaggregate daily rainfall into a half an hour timescale rainfalls. Since the future rainfall intensities and distributions are not available, it is essential to calibrate the stochastic model to a wide range of the recorded data. This incorporates a shortcoming of the procedure due to the stationary assumption, which assumes that some of the statistical characteristics of the rainfall will not change in the future. To test the validity of the approach, it is applied to daily rainfall values for 2009 at Moorabbin Airport. The results have then been compared with the half hour recorded data from the rain gauge.

Results and Discussion

As mentioned earlier, the validity of the model has been tested using 2009 data from Moorabbin Airport. Figure 8 presents a comparison of the recorded and modelled rain pulses for the month ten of 2009.
As shown in Figure 8, the model is able to capture the pattern of the recorded data. The next step of the disaggregation procedure presented in the flow chart in Figure 7 is to upscale or downscale the sum of the randomly generated values of the rainfall pulses to the exact amount of the daily rainfall. This step provides a direct relationship between the daily rainfalls and the sub-hourly rainfall pulses from the stochastic model. The other adjustment that is being made is through the model calibration by applying the case study data. It is observed that the model could not generate the outstanding peaks of some extreme events. To provide a better picture of the model’s capability of reproducing the synthetic pulses, Figure 9 presents a blow up of the wet events.

A systematic evaluation of the model is needed to determine the compatibility of the approach. To identify the most suitable coefficients of the probability distributions, a number of different yearly datasets are required. It is clear that the model is very sensitive to the PDF’s parameters. The peaks still could not be captured by the model.
Conclusion

A statistical as well as stochastic downscaling approach is proposed in this study, to be used to produce rainfall data for the evaluation of urban drainage systems. The proposed approach is used to generate daily projected rainfalls through the Statistical Downscaling Model. In the next stage, the spatially downscaled daily rainfall is disaggregated into half hour timescale rainfalls using the stochastic methodology proposed in this study. The proposed stochastic model is calibrated using rainfall data from Moorabbin Airport for 2010. The model has then been validated using rainfall data from 2009. Further studies are underway to evaluate the reliability of the proposed downscaling approach using more data.

References