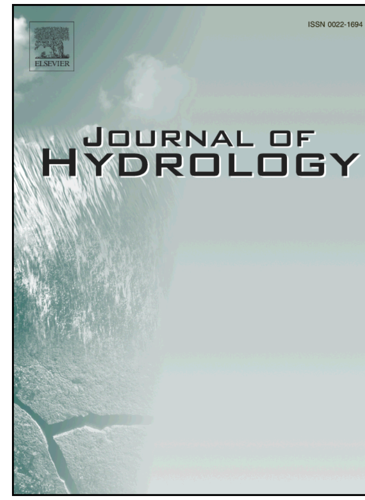
Accepted Manuscript



Research papers

Evaluating the reliability of stormwater treatment systems under various future climate conditions

Kefeng Zhang, Desmond Manuelpillai, Bhupendra Raut, Ana Deletic, Peter M. Bach

PII: S0022-1694(18)30827-8
DOI: https://doi.org/10.1016/j.jhydrol.2018.10.056
Reference: HYDROL 23220

To appear in: *Journal of Hydrology*

Received Date: 5 June 2018
Revised Date: 24 September 2018
Accepted Date: 24 October 2018

Evaluating the reliability of stormwater treatment systems under various future climate conditions

Kefeng Zhang[1,2*], Desmond Manuelpillai[1], Bhupendra Raut[3,4], Ana Deletic[1,2], Peter M. Bach[2,5,6]

[1)] Water Research Centre, School of Civil and Environmental Engineering, UNSW Sydney, NSW 2052 Australia
[2)] Monash Infrastructure Research Institute, Department of Civil Engineering, Monash University Clayton 3800 VIC, Australia
[3)] School of Earth, Atmosphere and Environment, Monash University, Melbourne, Australia
[4)] Indian Institute of Tropical Meteorology, Pune, India
[5)] Swiss Federal Institute of Aquatic Science & Technology (Eawag), Überlandstrasse 133, Dübendorf, Switzerland
[6)] Institute of Environmental Engineering, ETH Zürich, 8093, Zürich, Switzerland

**Corresponding author email: kefeng.zhang860225@gmail.com*

Abstract

Water Sensitive Urban Design (WSUD) stormwater systems, also known as Low Impact Development (LID) systems or Nature Based Solutions (NBS), are currently implemented based on the underlying assumption of statistical stationarity of rainfall, which threatens to become outdated under climatic uncertainty. This paper applies a new downscaling method to examine the implications of climate change on future rainfall and evaluate the reliability of WSUD stormwater infrastructure in pollution reduction, flow frequency mitigation and reliability as an alternative water supply. A variety of future atmospheric scenarios are considered as part of this comprehensive assessment by analysing an ensemble of eight different downscaled General Circulation Models (GCMs). High resolution catchment-scale rainfall projections for Melbourne, Australia were generated using a scheme called High-resolution Downscaling of Rainfall Using STEPS (HiDRUS) at a fine 1 km and 6-minute scale for more precise analysis with uncertainty estimates. Statistical analyses show that, in general, the climate models predict a drier future with fewer rainfall events and longer dry periods when comparing the simulated near future (2040-2049) periods against the base-line period (1995-2004). The difference simulated between historical and future rainfall projections show minimum difference of WSUD performance in pollution removal and flow frequency reduction, with slightly lower harvesting reliability (<3%) observed under future climate; high variabilities, however, were observed across GCM simulations, indicating big uncertainties of system reliability under various conditions, *e.g.* design wetland sizes may vary from 2.5% to 4.0% of the impervious catchment area according to different future projects across GCMs. Larger WSUD systems are recommended to ensure reliable performance of pollution removal, as well as harvesting reliability under simulated future conditions. The significance of considering an ensemble of different GCMs as opposed to many scenarios generated by a single 'best' climate model was also demonstrated for the robust estimation of uncertainty in future WSUD reliability. This work highlights important considerations for the future design, management and quantitative evaluation of WSUD reliability.

Keywords: General Circulation Model (GCM); climate downscaling; stormwater management; pollutant removal; stormwater harvesting; Water Sensitive Urban Design (WSUD)

1. Introduction

Urbanisation and climate change are growing concerns (Semadeni-Davies *et al.*, 2008; IPCC, 2013) and pose challenges to water professionals. Numerical studies have been conducted to investigate the impact of future climate and rapid urbanisation on urban catchment hydrology requiring fine resolution data (e.g. Andréasson *et al.*, 2004; Denault *et al.*, 2006; Olsson *et al.*, 2009; Zahmatkesh *et al.*, 2015). Future climate scenarios with fine resolutions can be generated in many ways, via artificially decreasing/increasing historical rainfall time series (*e.g.* Bach *et al.*, 2013; Urich *et al.*, 2013), or through adjusting Intensity-Frequency-Duration (IFD) curves (e.g.McIntyre *et al.*, 2007); different downscaling techniques, including dynamic downscaling (e.g. Thoeun, 2015), and statistical downscaling (e.g. Hewitson and Crane, 1996; Rummukainen, 1997) have also been used to generate future projections from GCMs. These methods are often used to investigate the climate change on urban hydrological processes (Denault *et al.*, 2006; Olsson *et al.*, 2009) and conventional urban drainage systems (Prudhomme *et al.*, 2002; Semadeni-Davies *et al.*, 2008) in a regional / city scales. It is frequently found that higher rainfall intensities and peak flows are expected in urban areas in the future thereby impacting the performance of conventional urban drainage networks with larger flows, longer event durations and more frequent pluvial floods (Ashley *et al.*, 2005; Rosenberg *et al.*, 2010; Kang *et al.*, 2016; Wang *et al.*, 2017). The studies regarding the impact of climate change on local catchment scales are often limited due to the lack of finer scale of space-time rainfall predictions.

Recently, Raut *et al.* (2018) developed a multiplicative random cascade model – HiDRUS, which is able to generate very fine spatial and temporal resolution rainfall projections (in one kilometre and six minutes); HiDRUS is the first kind of method that allows for GCM outputs to be downscaled to a local catchment scale, taking into account the topographical features of the local landscape.

Sustainable solutions, *e.g.* stormwater biofilters, constructed wetlands and ponds, developed under the concept of Water Sensitive Urban Design (WSUD, also known as LID or NBS) (Fletcher *et al.*, 2015), were designed to mimic natural hydrological processes removed through urbanisation (Semadeni-Davies *et al.*, 2008), and they have also been found to effectively counter the adverse effects of climate change and urbanisation (Bach *et al.*, 2013; Wang *et al.*, 2017). However, these systems are traditionally designed and operated under the basis of stationary (*e.g.* using historical data) to achieve required treatment targets (VSC, 1999). This, despite the fact that climate change leads to more extreme rainfall (across most of Australia) and longer dry periods (in southeast and southwest Australia) (Steffen *et al.*, 2017), can impact the structural integrity and function of many of these natural systems; for example extreme events with either long dry periods or large volumes were found to adversely impact the pollution removal performance of stormwater biofilters (Zhang *et al.*, 2014). Hence, with detailed historical hydrological inputs used to design WSUD and their perceived long operational lifespan (e.g. usually goes beyond 20 years), it is questionable whether currently designed WSUD systems can still deliver adequate treatment several years later and, thus, it is necessary to investigate the impact that climate change can inflict.

Relevant research on climate change impact on the performance of WSUD systems is rather limited due to the requirement of high space-time resolution of rainfall data and it is currently unclear whether existing WSUD systems can provide adequate treatment under future climate. Burge *et al.* (2012) conducted a case study in Melbourne, Australia to measure the likely impact of climate change on WSUDs using multiple scenarios of adjusted historical rainfall time series to represent the extremes of a number of projected ranges of climate change scenarios. It was found that potential climate change futures will have minimal impact on the efficiency and effectiveness of WSUD infrastructure, *e.g.* most stormwater treatment devices cope very well with the climate change predictions (worst case scenario was a pollutant load reduction performance of only up to 6% for bioretentions systems) while, in the most cases, the reduction in system reliability as an alternative water supply was within 5% of the base case. Sharma *et al.* (2016) however found that climate change will lead to increased outflow concentrations of total suspended solids and copper from a stormwater retention pond by comparing current climate and a synthetic future climate scenario with increased intensity of rainfall events and longer dry periods. Both studies used simple approaches to account for future climate (*i.e.* rainfall scaling/adjustments) and the contradicting findings trigger the further need for investigating the level of uncertainty of the treatment performance of WSUD systems under future climate and, hence, understanding the reliability of these systems designed according to the current paradigm, against variable future scenarios.

As decentralised WSUD systems are usually designed to treat medium to small events (*i.e.* up to 2 years return periods) and can be located at the smaller street and allotment scales, higher spatial (~ 1 km) and temporal resolution (< hourly) of future rainfall time series are needed (Ochoa-Rodriguez *et al.*, 2015). However, the lack of fine catchment-scale future rainfall time-series impedes the advancement of our understanding of WSUD system reliability under future climate. A further impediment is that there are even fewer studies that consider multiple scenarios in accounting for future climate uncertainties (Jones, 2000). GCMs are diverse in their implementation of the physics of atmosphere and have different couplings of atmosphere-ocean-ice components (IPCC, 2013). Therefore, it is reasonable to hypothesise that investigating multiple scenarios across GCMs would be far superior to exploring various scenarios within a single specific GCM (e.g. Rosenberg *et al.*, 2010); this can also lead to as a more conservative all-encompassing understanding about the degree of uncertainty of future system reliability for WSUD planning and robust decision making.

The aim of this study is to investigate the impact of future climate for three key functions of WSUD: (1) pollutant treatment performance and (2) flow frequency reduction (both are investigated for two common decentralised WSUD technologies – biofilter systems and constructed wetlands), as well as (3) harvesting reliability (based on the design of rainwater tanks). To undertake this study, we used HiDRUS (Raut *et al.*, 2018) to downscale fine spatial and temporal resolution rainfall projections (in one kilometre and six minutes) from an ensemble of eight different GCMs. Statistical investigation of historical and future rainfall projections for different rainfall characteristics was performed and compared to rain gauge data. We identified the extent of the future rainfall variability and the extent to which multiple rainfall time series from a single GCM vs. an ensemble of time series from multiple GCMs should be considered. Using an 'optimum number' of simulations extracted from different GCMs, we then propagated these rainfall projections to WSUD models to investigating their impact on system reliability and the associated design variability.

To the authors' knowledge, this is the first study that evaluates the future reliability of WSUD systems using finer scale space-time predictions of future rainfall downscaled from GCMs against a multitude of system behaviours. This study specifically focuses on Melbourne, Australia as the city has a rich history with WSUD and thousands of assets already implemented (Kuller *et al.*, 2018). Nevertheless, the methods are state-of-the-art and transferable to any other region in the world for which future rainfall data can be obtained. It should be acknowledged that our goal is not to validate whether GCMs and the HiDRUS can reproduce exact historical rainfall patterns, but to investigate variabilities of WSUD reliability impacted by future climate.

2. Methods

2.1 Rainfall data sets

This study specifically focused on one site: Melbourne Regional Office (MRO; Latitude: -37.81 °S Longitude: 144.97 °E), which has observed 6-min time series rainfall data for 1995-2004 (from the Bureau of Meteorology – BoM). The observed data is regarded as the baseline for the evaluation of climate change effects. Ten-year periods have been found to achieve a satisfactory compromise between modelling a short period that represents current climate conditions and a longer period that better illustrates climatic variability (Ashley et al., 2005). In the BoM rainfall data, there were two periods of missing data: 1-30 Nov 1998 and 1-28 Feb 2002; the data in the missing period were recorded as zeros but replaced by the authors using data from a nearby rain gauge station (Bundoora; Latitude: -37.72 °S Longitude: 145.05 °E; about 18 km north-east from MRO).

In this study, eight different GCMs that span the entire variability of the Australian region were selected (Table 1); a similar set of GCMs was also used in other studies within this region (Raut *et al.*, 2016). Selecting only eight GCMs was also was a compromise due to (1) large disk space requirement for storing the downscaled rainfall data-sets that have very fine spatial and temporal resolutions (e.g. > 100 TB), and (2) long computation time for WSUD modelling using the long-term downscaled rainfall series. This study used the RCP8.5 scenario from CMIP5 which represents the CO_2 emission scenario that is highly energy-intensive as a result of high population growth and a lower rate of technology development (van Vuuren *et al.*, 2011).

Table 1 *GCMs used in this study: temporal resolution of all GCMs is daily.*

GCM	Modeling Group, Country	Horizontal Resolution (Latitude × Longitude*)
ACCESS1-0	CSIRO, Australia	$1.25^o \times 1.875^o$
ACCESS1-3	CSIRO, Australia	$1.25^o \times 1.875^o$
BCC-CSM1	Beijing Climate Centre, China	$1.9^o \times 1.9^o$
CMCC-CMS	Centro Euro-Mediterraneo, Italy	$2^o \times 2^o$
CNRM-CM5	Meteo-France, France	$1.9^o \times 1.9^o$
GFDL-CM3	Geophysical Fluid, Dynamics Lab, USA	$2.0^o \times 2.5^o$
MIRO-C5	Centre for Climate System Research, Japan	$1.4^o \times 1.4^o$
MRI-CGCM3	Meteorological Research Institute, Japan	$2.8^o \times 2.8^o$

The HiDRUS model developed by Raut *et al.* (2018), was used to downscale the eight GCMs. The model uses multiplicative random cascades (which is often used in the urban hydrological context, *e.g.* Licznar *et al.* (2015) and Müller and Haberlandt (2018)) from the Short-Term Ensemble Prediction Systems (STEPS, Seed *et al.*, 1999) to generate high-resolution rainfall fields. The rainfall structures follow Lagrangian temporal evolution in an AR2 framework. For details on the model, please refer to *Seed et al. (1999), Raut et al. (2012)* and Raut *et al.* (2018). Briefly, the model is capable of downscaling rainfall from time and space resolutions of several hours and a hundreds of kilometres to scales of minutes and kilometres (which is required for catchment-scale modelling). The model can successfully reproduce the frequency distribution of 6-minute rainfall intensities, storm durations, inter-arrival times and autocorrelation function against radar data at 12 locations in the Greater Melbourne area (Raut *et al.*, 2018). The spatial variation in rainfall accumulation is realised by using multiplicative bias factors computed from the observed data (Raut *et al.*, 2018). The model was run with historical parameters estimated from the Melbourne radar data during the observation period 2008-2015. It should be noted that the period/data used to estimate historical parameters (2008-2015;

radar data) is different from the historical period of this current study (1995-2004; rain gauge data); further studies by Raut *et al.* (Submitted manuscript) on testing HiDRUS has validated that the parameter transferability was acceptable across different periods and datasets.

Using HiDRUS, 100 continuous simulations of rainfall at 1 km spatial and 6 minutes temporal resolutions were generated from each GCM over the periods of 1995-2004 (historical projections) and 2040-2049 ("near-future" projections – for simplification "future" is used throughout the paper instead of "near-future") at MRO. The 100 different simulations provide estimation of the 'within-GCM' variabilities of different scenarios; further analysis was done to evaluate the number of simulations needed to represent the variability in Section 2.2.2.

2.2 Rainfall data analysis

2.2.1 Rainfall characteristics

Model simulations (*i.e.* historical and future rainfall projections) and BoM rain gauge data were compared using a set of common rainfall characteristics at 6-min time steps. The purpose was to better understand the behaviour of each GCM so that subsequent changes in WSUD reliability could potentially be explained. Characteristics investigated included: *Annual Rainfall* (mm), *Number of Rainfall Events per year* (a count of the number of events that are recorded; a threshold of a minimum six hours were taken to separate two consecutive rainfall events), *Event Duration* (hours), *Average Rainfall Intensity* (mm/hr; average rainfall intensity across all the rain events recorded - propagated from 6-minute data), *Maximum Rainfall Intensity* (mm/hr; the average of the highest rainfall intensity of all the rain events recorded), *Average Dry Period* (hrs; average length of dry periods between two events), and *Annual Rainy Days* (a day with recorded total rainfall ≥1mm).

Box plots were created for each of the selected rainfall characteristics to compare downscaled and rain gauge data over the same periods. These not only provided an insight into the variability of scenarios being modelled, but how accurately GCMs can reproduce historical rainfall. An analysis of trends would also signal important findings about the GCM model-generated rainfall time series and future climate predictions in general.

2.2.2 Capturing variability in model rainfall characteristics

A large ensemble of simulations from models is highly desirable to robustly account for uncertainties in hydrological design (Raut *et al.*, 2018), *e.g.* in this study 100 simulations were generated from each GCM. However, it is often not convenient or practical to use the entire rainfall simulation data-set generated for WSUD modelling. It was therefore required to develop a pragmatic approach to the modelling by using fewer simulations while still ensuring that the extracted sample is representative of the original data. In this study, subsets of the 100 scenarios from each GCM (*i.e.* subsets of 5, 10, 25 and 50 scenarios) were randomly selected 10,000 times to generate large number of possible scenario combinations. These subsets were statistically compared to the original 100 GCM simulations. Each subset was compared against the original 100 scenarios: (i) to ascertain whether the reduced sample was statistically representative of the original data (*i.e.* has similar median) and (ii) was clearly derived from the same continuous distribution at the 5% significance level. This was done using Wilcoxon rank-sum and two-sample Kolmogorov-Smirnov tests (Gibbons and Chakraborti, 2011).

2.3 Propagation of future rainfall projections to stormwater models

2.3.1 Stormwater treatment performance

The Model for Urban Stormwater Improvement Conceptualisation (MUSIC) (eWater, 2014) and a third party software package, known as the DaCapo Design Curve Generator (Bach and Dotto, 2016), were employed to evaluate the change in WSUD treatment performance (presented as design curves) under future climate using the downscaled rainfall projections as input with the number of simulations selected for use. MUSIC is a conceptual model that uses 6-minute continuous rainfall data to simulate rainfall-runoff, pollution generation and treatment processes across a user-defined catchment and WSUD treatment train, widely used in the Australian urban water industry to evaluate the treatment performance of WSUD assets. As multiple scenarios and a large number of system designs were required, DaCapo was used. The software systematically generates MUSIC simulations files though varying the parameter values and batch runs them to produce data with which to plot WSUD system performance curves (*i.e.* pollutant removal efficiency vs. system size).

In this study, two WSUD technologies – stormwater biofilters and wetlands were selected for investigation, as they are currently the most widely used WSUD systems for stormwater treatment (Hatt *et al.*, 2006), and their treatment performance, as well as flow mitigation capacities have been well reported (Gogate *et al.*, 2017). System sizes were varied from 0.01 – 5% of the catchment impervious area (taken here as 1 ha, 100% impervious) for biofilters and 0.01 – 10% for wetland systems. Two different biofilter designs with extended detention depths of 0.1m and 0.4m and wetland designs having water depths of 0.1m and 0.5m were tested. Other input parameters are summarised in Table 2. Four key indicators were simulated through MUSIC: flow reductions, load reductions of Total Suspended Solids (TSS), Total Phosphorus (TP) and Total Nitrogen (TN). Results were expressed in terms of these four indicators and plotted as performance curves against varying system sizes using both BoM rain gauge data and rainfall projections for all selected GCMs. It should be noted that biofilter and wetland systems were investigated individually in this study, *i.e.* one single WSUD technology in a catchment, to understand the specific impact of climate change on individual system.

Table 2 *Key design inputs used in MUSIC to model pollutant removal efficiency for biofilters and wetlands*

System Type	Design Parameter(s)
Biofilter System	
System Surface Area [% catchment impervious area]	0.01 – 5 (1)
Extended Detention Depth [m]	0.1 and 0.4
Submerged Zone Depth [m]	0.4
Filter Depth [m]	0.5
Media Saturated Hydraulic Conductivity [mm/hr]	180
System Exfiltration Rate [mm/hr]	0.0 (2)
Constructed Wetland System	
System Surface Area [% catchment impervious area]	0.01 – 10 (1)
Extended Detention Depth [m]	0.1 and 0.5
Permanent ponding depth [m]	0.35
System Exfiltration Rate [mm/hr]	0.0 (2)
Detention Time [hrs]	72
Treatment Targets to be meet (3):	
Total suspend solids (% load reduction)	80
Total Nitrogen (% load reduction)	45
Total Phosphorus (% load reduction)	45

(1) 30 different surface areas were modelled, increasing exponentially from 0.01 to 5 (for biofilter) or 10 (for wetland); (2) Both WSUD systems were assumed to be lined, i.e. water in the systems was not allowed to exfiltrate into the groundwater; (3) according to by Best Practice Environment Management (BPEM) guideline in Victoria, Australia (VSC, 1999)

2.3.2 Impact on flow frequency

WSUD is known to reduce peak stormwater flows. Hence, in this study, the potential of biofilters and wetlands to reduce flow frequency in urban areas was estimated, mainly in terms of managing the frequent flood events (*e.g.* return periods up to 2 years); the 6-min time series flow data exported from MUSIC were examined to identify all the peak daily discharges (excluding zero flows that resulting from no rain) for partial frequency analysis (for detailed methods please refer to Pilgrim (2007)), for the above investigated catchment. In addition to the urbanised catchment (no WSUD implementation), two WSUD catchments were further explored: (1) biofilter sized as 1% of catchment area with extended detention depth (EDD) of 0.1 m and (2) wetland sized as 3% of catchment area with EDD of 0.5 m; both were typical biofilter/wetland design size in Melbourne. The generated historical and future rainfall projections, as well as observed BoM rain gauge data were used to generate the flow frequency curves to show the likelihood of daily peak flow changes.

2.3.3 Storage-Behaviour analysis

To assess changes in the reliability of alternative water supply through stormwater harvesting using rainwater tanks under future climate, a hypothetical storage-behaviour analysis was undertaken with continuous simulation of inflow, outflow and changes in storage volume of a simple rainwater tank according to mass balance principles (Fewkes and Butler, 2000; Liaw and Tsai, 2004; Mitchell *et al.*, 2008). The storage-behaviour model was simulated at a daily time step using aggregating 6-minute GCM downscaled rainfall. Imteaz *et al.* (2011) and Mitchell et al. (2008) both demonstrated that using a daily time step in the water balance model can sufficiently determine a realistic tank volume.

For this analysis, the runoff coefficient of the roof catchment was set to 1.0 where the generated runoff from a designated household roof area of 200 m^2 was assumed to divert completely to the connected storage tank. A constant daily demand (D) of 0.43 m^3/day was set according to the reported daily water demand for in Melbourne (Coombes and Kuczera, 2003); tank sizes between 1 to 10 m^3 were simulated to express results as a storage-reliability curve. Water demand, tank storage, spillage (using a 'yield after spill' order Mitchell *et al.*, 2008) and supply were then calculated following the sequence of equations listed below (Eq. 1 to 4). On a particular day, if the water storage is greater than the tank capacity C, the excess water will spill over and the tank storage level at the end of the day is reset equal to C. The amount supplied is then equal to the demand, D, or limited to the volume remaining in the tank, S_i, depending on how much water is available.

$$V_i = R \times A \tag{1}$$
$$S_i = V_i + S_{i-1} \tag{2}$$
$$V_{spill} = S_i - C \quad \text{for } S_i > C \tag{3}$$
$$Su_i = \begin{cases} D_i \rightarrow S_i = S_i - D \text{ for } S_i > D \\ S_i \rightarrow S_i = 0 \text{ for } S_i < D \end{cases} \tag{4}$$

where, V_i is the harvested rainwater inflow on the i[th] day (m^3), R is the daily rainfall (mm), A is the roof area (m^2), S_i is the water stored in the rainwater tank (m^3), S_{i-1} is the amount of water at the beginning of time step/day before (m^3), V_{spill} is the spilled amount of water from the tank (m^3), C is the capacity of the rainwater tank (m^3), Su_i is the amount of water supplied each day, and D_i is the daily rainwater demand (m^3).

Volumetric reliability is used to quantify the performance of the rainwater harvesting system and was calculated as:

$$R_e = \frac{\sum S_u}{\sum D} \times 100 \tag{5}$$

where, R_e is the volumetric reliability of the tank, D is the cumulated demand over the simulated 10-year period and Su is the total amount of water supplied in response to this demand. Reliability curves are developed using the output from Eq. 4 with reliability (R_e) plotted against storage volume, *i.e.* tank capacity (C) to visually discern the effect of future climate conditions on stormwater harvesting reliability.

3. Results & Discussion

3.1 Statistical analysis of rainfall data

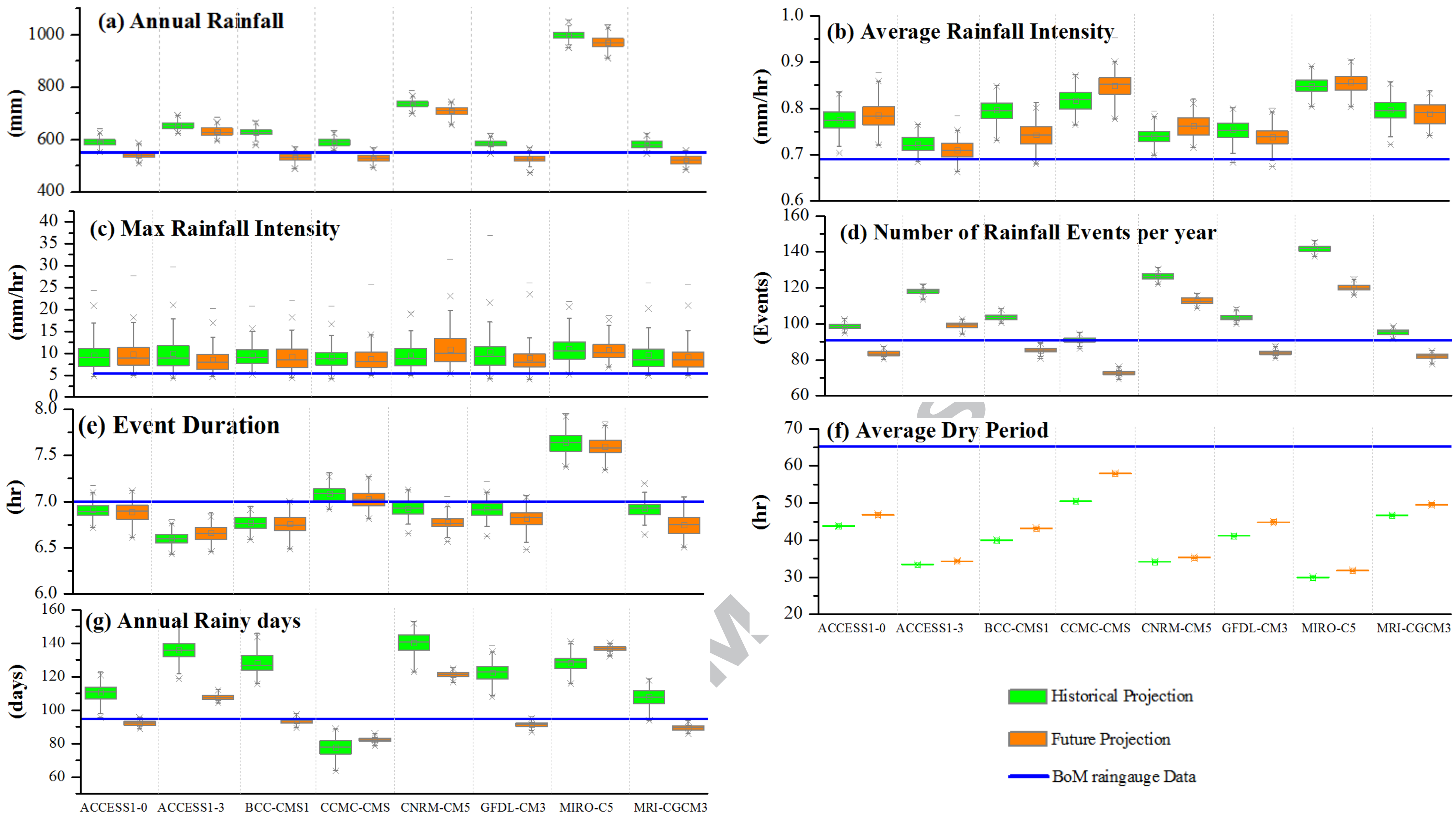
Figure 1 compares different rainfall characteristics (*e.g.* annual rainfall, intensities, durations, rainy days, etc.) across the different data sets and time periods. By comparing historical projections and future projections, a drier future was simulated by all GCMs, evidently by less annual rainfall (average of 579.4 mm of all GCMs), smaller number of rainfall events per year (93), as well as longer dry periods (43 hours) compared to that of historical projections: 667.8 mm annual rainfall, 110 events, and 40 hours dry periods respectively. The results also indicate less rain days each year for most of the GCMs, except for CMCC-CMS and MIRO-C5. In addition, the observed trends for both the average and maximum rainfall intensities vary across different GCMs, with some simulated higher extreme maximum intensities (*e.g.* ACCESS1-1, BCC-CMS1, CMCC-CMS and CNRM-CM5); this have good agreement with previous studies that suggest higher rainfall intensities in the future (e.g. Rosenberg *et al.*, 2010). Other GCMs, however, generated lower extreme intensities in the future. Across all eight GCMs, MIRO-C5 is found to be the 'wettest' model (annual rainfall between 950 and 1000 mm) and GFDL-CM3 the 'driest' (annual rainfall between 500-650 mm). Within each GCMs, the variation between 100 simulations is, however, small except for maximum intensity (which has a similar range between GCMs). This indicates that variations between GCMs should be prioritised rather than variations within a single GCM to understand future climate variability.

It was found that the rainfall characteristics of historical rainfall projections from the studied GCMs did not have perfect match with that of observed BoM rain gauge data (Figure 1). Discrepancies between rain gauge data and the projections from GCMs is evident, especially for *Average Dry Period* (underestimated: 30-50 hours by GCM compared to 65 hours by BoM data; Figure 1); both average and maximum rainfall intensities estimated from GCM projections were generally higher than that from BoM data. Differences are expected as these GCMs are usually not calibrated directly to the observed rainfalls but rather the atmospheric conditions to produce the global-scale atmospheric processes (IPCC, 2013). Hence while GCMs may have good fit with historical atmospheric conditions, the rainfall produced may be very different from historical measurements. The downscaling process attempts to reflect the spatial variability within a single spatial domain. HiDRUS has been able to reproduce spatial patterns of rainfall across Melbourne, Australia's domain as has been demonstrated (Raut *et al.*, 2018). However, further distinguishing a single point (*i.e.* Melbourne Regional Office) within this spatial domain for reasons of comparison with historical data will exacerbate uncertainties. The authors believe that further calibration of the data against observations may actually risk altering the model's intended behaviour entirely. As such, instead the uncertainties of each GCM were embraced by exploring their 'within model' variability as well as their use as an ensemble future rainfall data set. All selected GCMs are therefore included for further analysis; as it provides a more robust assessment of the variability of WSUD reliability under future climate. Given the above arguments further over-calibration of the downscaled time series are refrained, but the following analysis of the results are performed while keeping in mind the actual differences between downscaled historical GCM simulations and BoM observations.

3.2 Variability analysis of model simulations

The frequency distribution curve in Figure 2 shows that a high level of similarity is retained when reducing the sample size from the original 100 simulations downscaled from each GCM with even 5 simulations yielding significant proportions of *p*-values greater than 0.05. This is true for all models as only slight variations are observed. It is therefore suggested that, for 10-year future scenario rainfall time series, 5 to 10 simulations from a single GCM's downscaled outputs would be sufficient to capture intrinsic model variability and maximise computational efficiency, also allowing modellers to consider a greater variety of GCMs or climate scenarios in their assessments. These findings hold true for Melbourne, but would have to be assessed for other locations. Findings may also differ for shorter time series.

326



327
328 *Figure 1: Box plots of rainfall characteristics at Melbourne Regional Office using 6-minute ensemble of GCM simulations and rain gauge data over the*
329 *baseline period of 1995-2004 (historical projections) and future period of 2040-2049 (future projections). 100 simulations from each GCM were used for this*
330 *investigation.*

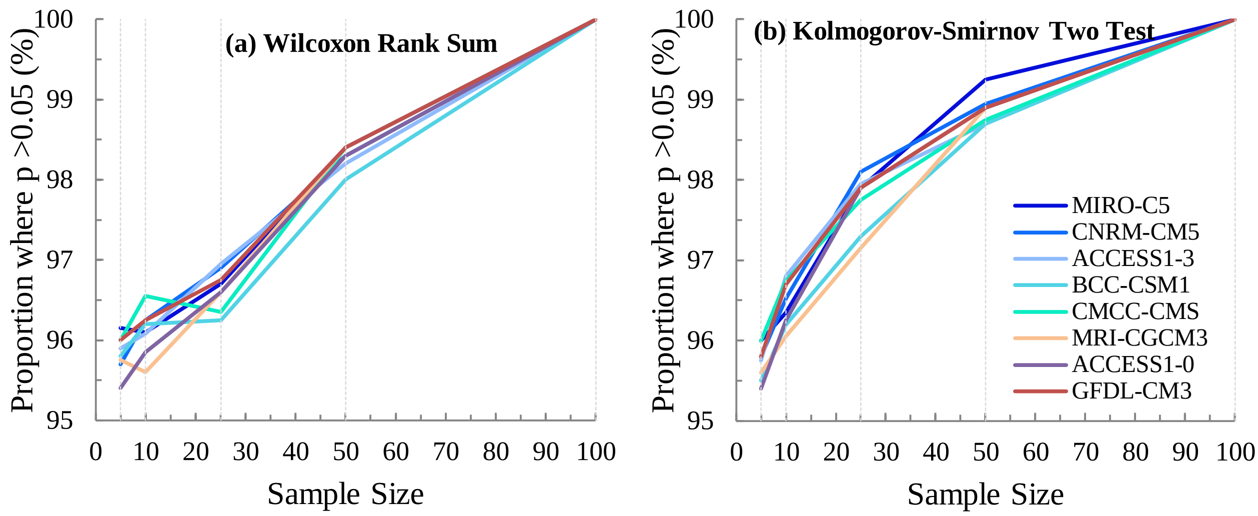


***Figure 2**: Frequency distribution curve for proportion of sample sizes (5, 10, 25, 50 and 100) that show significant similarity with the original 100 ensemble simulations for each GCM. GCMs are indicated with legends following a colour gradient from wettest (top dark blue: MIRO-C5) to driest (bottom orange: GFDL-CM3).*

3.3 Stormwater treatment performance

Figure 3 presents the design curves generated using historical and future rainfall projections, with the comparison to one generated based on BoM observation data (only volume and TN curves are presented here as examples due to the high similarities between the pollutants). Interestingly, despite the statistical differences of rainfall characteristics identified in Figure 1 between historical and future rainfall projections, the design curves in Figure 3 show high similarities between the historical and future ones. However, slightly higher runoff volume reduction (~3%) is simulated from future rainfall projections compared to historical conditions (Figure 3). This finding is further confirmed by the direct comparison of WSUD treatment performance estimated from historical and future rainfall projections (Figure 4) — all points are closely centred around the 1:1 "*no-impact*" line (*i.e.* future equals to history) for pollution reduction, while majority of the points (>95%) for runoff volume reduction are above the "*no-impact*" line. The increased performance of runoff volume reduction in the simulated future is likely due to decreased inflow volumes in the drier future, which was found previously to lead to higher volume reductions (Hatt *et al.*, 2009); pollutant removal in WSUD systems, however, is influenced by multiple factors *e.g.* it was reported that longer dry periods can adversely impact system performance in removing sediments and nutrients in stormwater (Hatt *et al.*, 2009; Zhang *et al.*, 2015). Nevertheless, other design elements/operational factors of biofilters (*e.g.* plants, submerged zones and maintenance), which are not accounted in MUSIC may become crucial when considering future climate. For example, a notable increase in dry period or decrease in the number of rainy days (see Figure 1 f and g) will require a change in design of how water is retained in the system or more frequent maintenance to ensure that plants within the system are healthy.

Figure 3 also indicates variability among different scenarios within a single GCM and different GCMs. Bigger variations of reductions are observed between GCM predictions than within a GCM. GCM projections produce broader bands for wetland systems compared to biofilters, showing that wetlands appear more susceptible to future climate. This could also be due to of larger system size and scale, *e.g.* wetlands are usually an order of magnitude larger than biofilters.

Table 3 summarises the desired system size to meet treatment targets required by BPEM guideline in Victoria, Australia (Table 2) estimated using the BoM curve as well as the GCM curves in Figure 3. According to the estimates, in general, large uncertainties exist considering various historical and future scenarios with the latter ones providing higher variability. It can be seen that if wetlands are to be designed to achieve BPEM targets, they should be sized to 3.5% of catchment impervious area according to the BoM curve (which currently underpins system design and compliance checking in practice); depending on different future scenarios, system sizes can vary from 2.5% to 4.0% of the catchment area, indicating that the current design may be sufficient but under some GCM scenarios (*e.g.* MIRO-C5), larger systems may be required. In terms of biofilters, the current designed system size (0.8% of the catchment impervious area) is at the lower boundary of the variability bands estimated from all GCMs (0.8-1.5% of catchment area; Table 3) to deliver the same treatment performance. While it is often not necessary, biofilters could be sized to 1.5% of the catchment area for ensuring 100% reliability under future conditions according to the current analysis. MICRO-C5, which is the ‘wettest’ GCM, always guides the critical system design for both systems.

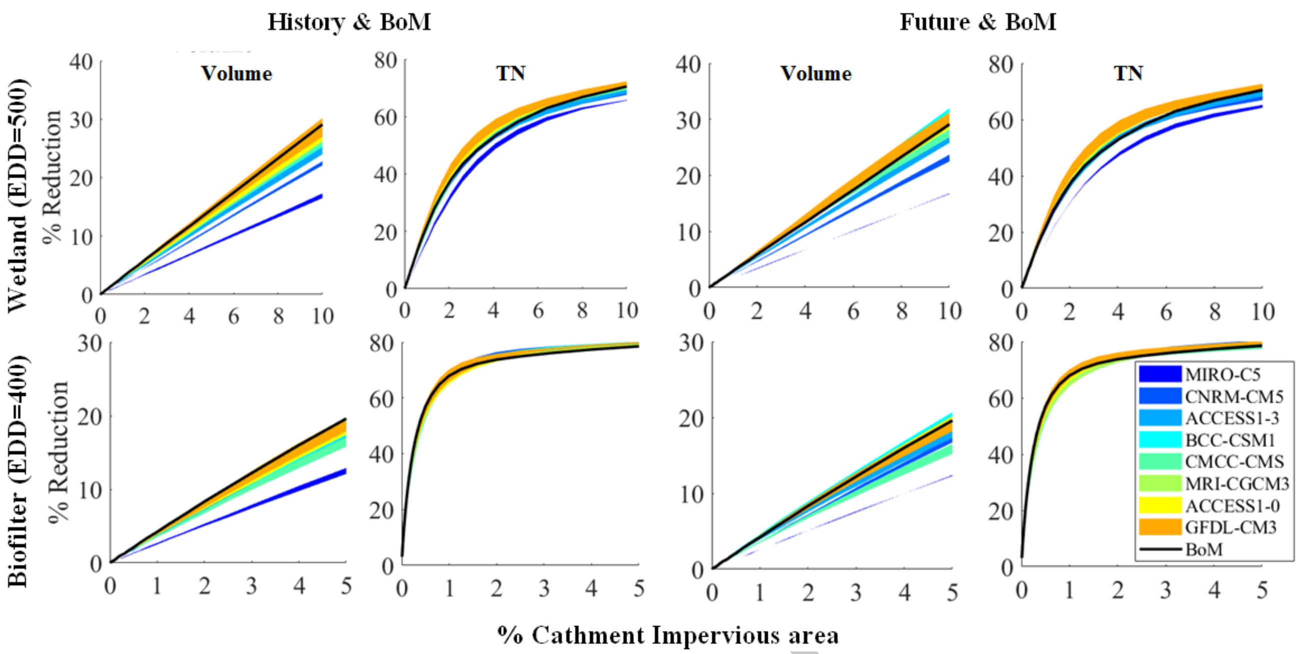


Figure 3: *Wetland (EDD=500mm) and biofilter (EDD=400mm) design curves. The shaded bands indicate the variation between the 5 simulations used in each GCM. GCMs are indicated with legends following a colour gradient from wettest (dark blue: MIRO-C5) to driest (orange: GFDL-CM3).*

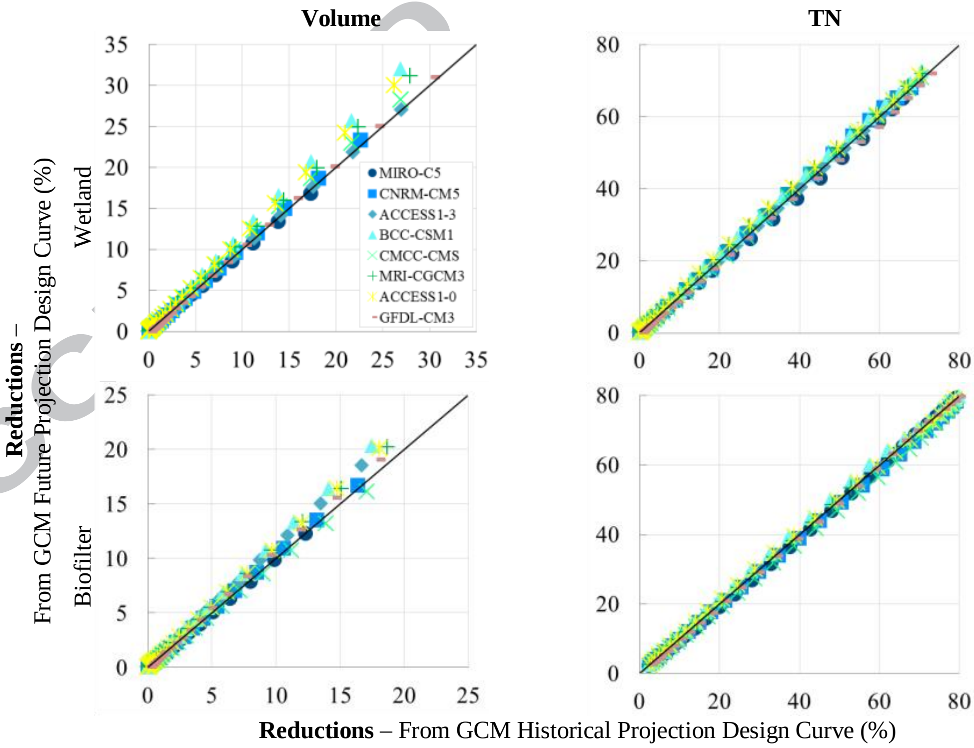


Figure 4*: Comparison of treatment performance estimated using historical design curves and future curves for wetland (EDD= 500mm) and biofilter (EDD = 400 mm).*

Table 3 *Estimated sizes of the systems required to meet BPEM targets (i.e. 80% TSS, 45% TP and 45% TN load reduction) using historical and future projections, as well as BoM rainfall*

	Wetland[1)] (*Historical*)	Wetland (*Future*)	Biofilter (*Historical*)	Biofilter (*Future*)
MIRO-C5	3.8-4.0%[3)]	3.7-4.0%	1.4%	1.4-1.5%
CNRM-CM5	3.3-3.4%	3.1-3.5%	1.1-1.2%	1.1-1.2%
ACCESS1-3	2.9-3.5%	3.2-3.6%	1.0-1.1%	0.9-1.1%
BCC-CSM1	2.8-3.2%	2.8-3.2%	0.9-1.0%	0.9-1.0%
CMCC-CMS	3.1-3.2%	3.2-3.4%	0.9-1.0%	0.9-1.0%
MRI-CGCM3	3.1-3.2%	2.9-3.1%	0.9-1.0%	0.9-1.0%
ACCESS1-0	2.8-3.1%	2.7-3.2%	1.0-1.1%	0.9-1.1%
GFDL-CM3	2.5-3.1%	2.5-3.1%	0.8-1.0%	0.8-0.9%
BoM	**3.5%**	**3.5%**	**0.8%**	**0.8%**

Note: [1)] the critical pollutant for wetland design is TSS; [2)] the critical pollutant for wetland design is TP; [3)] sizes are expressed as percentage of catchment area;

3.4 Impact on flow frequency

The comparisons of peak flows estimated by MUSIC using historical and future rainfall projections, as well as observed BoM rainfall are presented in Figure 5. As indicated, in general climate change simulated through HiDRUS from different GCMs shows high similarity of flows between historical and future scenarios and with the direct comparisons in Figure 5c&f show limited impact for small flows (*i.e.* <0.1 m^3/s) with low variability, closely aligning to the 1:1 "*no-impact*" line. This explains why minimal differences of pollution removals were estimated between historical and future projections found in Figure 3 as WSUD systems are usually designed for small events with flows up to 3-month return period (=0.25 [Years], peak flow is equivalent to approximately 0.1 m^3/s) for pollution treatment. While for higher flows (>3-month return period), points scatter further away from the "*no-impact*" line and the differences vary according to different GCMs, indicating higher variabilities between GCMs. For example, the GCMs with both higher average and maximum rainfall intensities in the future projections (*e.g.* CNRM-CM5, Figure 1) often have higher estimated flows than the ones with lower estimated average and maximum rainfall intensities (*e.g.* BCC-CMS1).

The flow frequency behaviour of a fully urbanised catchment with no WSUD systems implementation (estimated from the BoM rainfall by MUSIC) was also plotted in Figure 5. Biofilters or wetlands can provide the greatest benefit up to a return period of around 3-month, after which there are minimal differences with that of an urbanised catchment in terms of peak flows, potentially indicating the ineffectiveness of these systems under such significant events; biofilters and wetlands are able to reduce the peak flow of 3-month return period rainfall event from 0.085m^3/s to 0.075m^3/s and 0.070m^3/s, respectively. Again, considering future climate, large uncertainties are observed between GCMs, and even with WSUD implementation, the estimated flows often exceed the values of the urbanised catchment at present day. The estimated 3-month flows vary from 0.06 - 0.12 m^3/s for the WSUD catchments and the wettest model (MIRO-C5) represented the most critical design scenario (*i.e.* = 0.12 m^3/s); this is expected as MIRO-C5 has the highest intensities and longest event durations (Figure 1). This also explains why the BoM curves are closer to the lower bounds of the curves as BoM rainfall data has relatively lower rainfall intensities. Comparative evaluation also illustrates another risk of selecting a singular "best" model as if selecting wettest GCM MIRO-C5 might lead to overly conservative predictions of larger, more frequent flows.

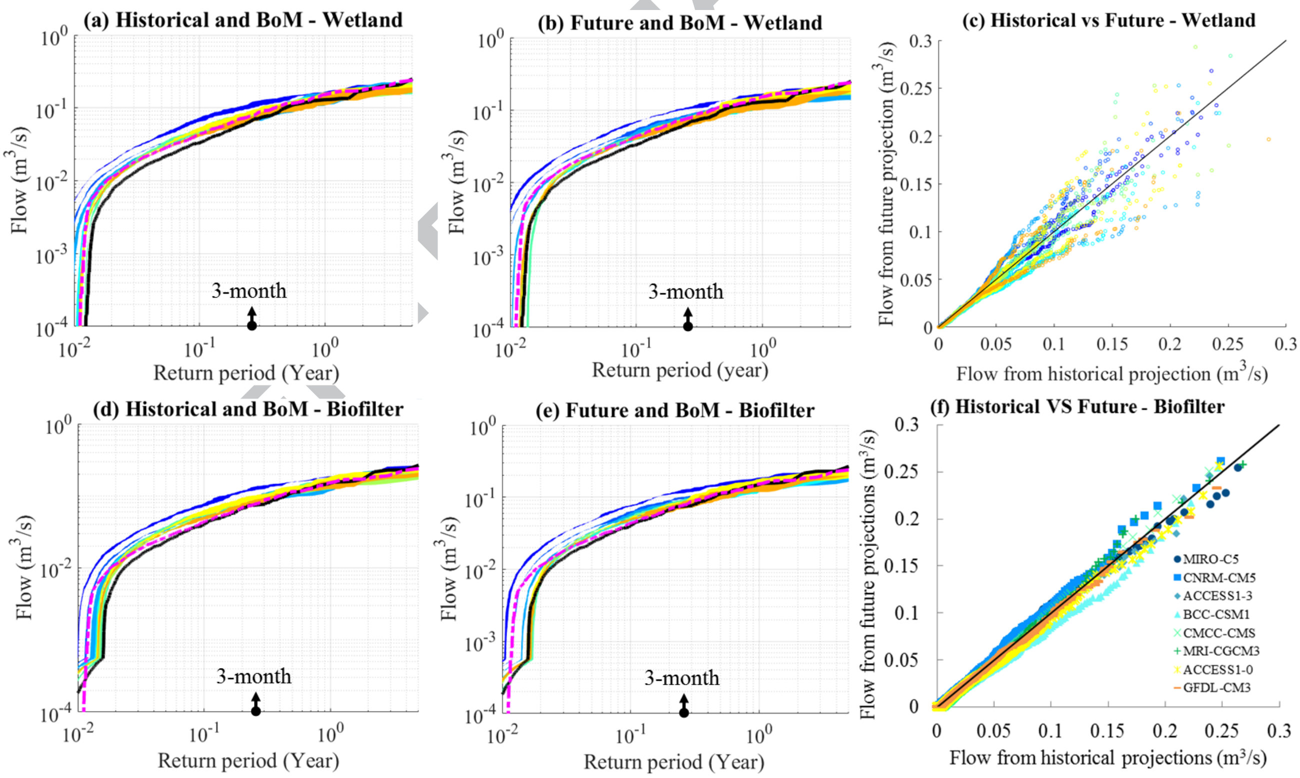


Figure 5: *Comparisons of daily peak flows generated from historical and future projections, as well as observed BoM rainfall; Year in log used in X-axis; 'BoM (Urbanised – No WSUD)' refers to the with no biofilter or wetland systems).*

3.5 Stormwater harvesting reliability

Figure 6 shows the storage-behaviour analysis results (plotted as reliability against tank size) estimated using both historical and future projections, and BoM rainfall. Different from treatment performance and flow reduction that have high similarity between historical and future conditions, simulated climate change always lead to slightly lower harvesting reliabilities (<3%) for all GCMs (see Figure 6), very likely due to drier conditions in the future. A clear difference in harvesting reliability between GCMs is observed, with higher reliability estimated for wetter GCMs. For example, for a 3 m^3 rainwater tank, the estimated reliability in the future can vary from 45% (CMCC-CMS) to 72% (MIRO-C5), indicating significant future uncertainties for harvesting reliability through rainwater tanks. The variation between simulations within each GCM is still minor (e.g. <3% for 10 m^3 tank) and large variation evidenced across GCMs. The resulting critical scenario for this case (CMCC-CMS) is different to that for treatment performance and flow frequency impact where the MIRO-C5 is the critical GCM (Figure 3; Figure 6). This may be attributed to the lower rainfall, greater dry period and less rainy days recorded for CMCC-CMS (Figure 1).

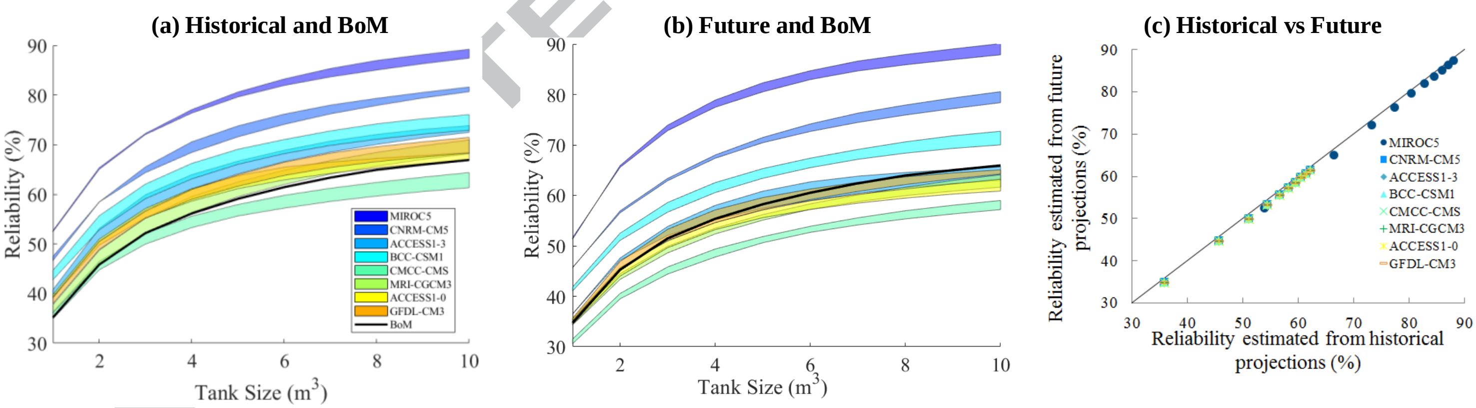


Figure 6*: Storage behaviour curves illustrating the relationship between rainwater tank size and reliability under the different climate scenarios.*

4. Practical implications and limitations of the work

Adapting to climate change will require rejecting basic assumptions about stationary conditions that have historically underpinned flood, water, and conservation management (Milly *et al.*, 2008). Some argue that simply coping with present climate variability is enough of a challenge (Washington et al., 2006), however, with the understanding that infrastructure is generally less costly and disruptive if necessary measures to mitigate climate change are taken well in advance of anticipated changes.

Given the complexity of global warming and climate uncertainty, which is well exemplified by distinctive statistical characteristics of each GCM in Figure 1, it is necessary to look beyond the notion of a singular “best model” and towards the use of an ensemble in order to provide assurance and robust assessment of system reliability under a wide range of potential conditions. The fact that there was no unambiguously superior model observed (*i.e.* downscaled historical GCM projections did not match perfectly with observations) and, most importantly, that the critical design scenario varied according to the performance indicators (*i.e.* treatment vs harvesting) supports this hypotheses that it is better to use less simulations and more models than less models and more simulations. It was found that using a large number of realisations from each downscaled GCM is not necessary if designers wish to most efficiently and comprehensively assess the reliability of their systems over the long term. In this study only 5 realisations were used, and the actual number for different cases shall be determined according to specific rainfall data, the study area and the objective criteria.

The results in this paper suggest that the WSUD performance in pollutant removal and flow reductions under simulated future climates has minimal difference to that under simulated historical conditions. Stormwater harvesting systems are expected to have slightly lower reliabilities (<3%) under future climate. Significant uncertainties exist according to future rainfall projections across various GCMs; looking into these uncertainties can provide insights to the performance variabilities, and assisting the adaptation of WSUD systems into unknown future. The analysis reveals that while WSUD systems may stay resilient in providing treatment performance under the simulated future, they could be also sized bigger to account for future climate uncertainties. Notably, this analysis only considered impact of future rainfall in treatment performance, while in fact the stormwater pollution concentration may increase due to urbanisation (Wang *et al.*, 2017), hence further studies are recommended to understand the impact of urbanisation on WSUD reliability. Moreover, peak flows (especially for return periods of < 3 months) based on BoM rainfall are often underestimated compared to the simulated future climate variability; this was probably due to the higher estimated rainfall intensities from both historical and future projections compared to BoM observations. Larger rainwater tank sizes may be needed to provide same level of harvesting reliability considering future climate change scenarios projected from any GCM, *e.g.* to achieve 50% of reliability, 3 m^3 tank size is typically enough at present according to BoM rainfall data, it however needs to be increased to at least 4.5 m^3 to keep the same reliability under the future scenarios based on the CMCC-CMS (Figure 6).

It was noted that there were differences between downscaled historical data and observations, which may potentially influence the results (*e.g.* the underestimation of peak flows from BoM rainfall). Moreover, this study was limited to the results of just eight climate model projections at a single location, with only one CO_2 emission scenario – RCP8.5 scenario form CMIP5 that represents no action in climate change mitigation. Further, the uncertainties in the models used have not been discussed. Therefore the findings from the current study should be taken as preliminary, future research should be directed towards replicating and expanding this study and reproducing these results for a larger database with more locations and models to further justify these claims. Also, studies can also be expanded to involve more WSUD technologies, *e.g.* ponds and swales; nevertheless, all these analysis indicate the necessity of considering future uncertainties when designing stormwater management systems; and the method in this study can be used to quantitatively design a conservative WSUD system to cope with future climate variability.

5. Conclusion

In this paper, a multiplicative cascade model (HiDRUS) was employed to generate high resolution rainfall projections at 1 km and 6 minute interval for an ensemble of General Circulation Models (GCMs) at Melbourne Regional Office to offer, for the first time, comprehensive insights into the adaptability of Water Sensitive Urban Design (WSUD) under a variety of future climate scenarios on a much fine scale. Four critical contributions and conclusions from this study were found:

1. Downscaled rainfall projections from eight different GCM's have been statistically

characterised; the results indicate the likelihood of a drier future with less rainfall events and longer dry periods;

2. Despite their validity at the atmospheric level, rainfall projections do not necessary match the observed rain gauge data. Nevertheless, these discrepancies and uncertainties associated with their predictions should be embraced and propagated to WSUD reliability rather than discarded;
3. Using an ensemble of GCMs is preferable over a "best model, many scenarios" approach to account for the variety of potential scenarios; it is found that for this location and setup only 5 rainfall simulation scenarios (10 years in length) are needed to capture the variability of the downscaled rainfall simulations from each tested GCM;
4. Minor differences of WSUD performance in removing pollutants, flow frequency reduction, and slightly lower harvesting reliability were observed when comparing simulated historical conditions with future scenarios; however high variabilities do exist and larger WSUD systems are suggested to cope with the high variabilities in the simulated future scenarios, in order to ensure that the treatment targets as well as harvesting reliability are still met.

Acknowledgments

This project was funded by the Cooperative Research Centre (CRC) for Water Sensitive Cities as part of The Water Sensitive Cities Toolkit Project (Project D1.5). The authors would like to thank Cintia Dotto, Ross Allen, Christian Jakob and Michael Reeder for their contributions to this project.

References

Andréasson, J., Bergström, S., Carlsson, B., Graham, L.P. and Lindström, G., 2004. Hydrological change–climate change impact simulations for Sweden. *AMBIO: A Journal of the Human Environment* 33(4), 228-234.

Ashley, R.M., Balmforth, D.J., Saul, A.J. and Blanskby, J., 2005. Flooding in the future–predicting climate change, risks and responses in urban areas. *Water science and technology* 52(5), 265-273.

Bach, P.M., Deletic, A., Urich, C., Sitzenfrei, R., Kleidorfer, M., Rauch, W. and McCarthy, D.T., 2013. Modelling Interactions Between Lot-Scale Decentralised Water Infrastructure and Urban Form - a Case Study on Infiltration Systems. *Water Resources Management* 27(14), 4845-4863.

Bach, P.M. and Dotto, C.B.S., 2016. Da Capo Project Overview [Online] Available: www.petermbach.com/dacapo.

Burge, K., Browne, D., Breen, P. and Wingad, J., 2012. Water Sensitive Urban Design in a changing climate: estimating the performance of WSUD treatment measures under various climate change scenarios, 119, *Engineers Australia*.
Coombes, P.J. and Kuczera, G., 2003. Analysis of the performance of rainwater tanks in Australian capital locations. *Proceedings of the 28th International Hydrology and Water Resources Symposium*.
Denault, C., Millar, R.G. and Lence, B.J., 2006. Assessment of possible impacts of climate change in an urban catchment. *JAWRA Journal of the American Water Resources Association* 42(3), 685-697.
eWater, 2014. Model for urban stormwater improvement conceptualisation (MUSIC Version 6).
Fewkes, A. and Butler, D., 2000. Simulating the performance of rainwater collection systems using behavioral models. *Build. Serv. Eng. Res. Technology* 21(2), 99-106.
Fletcher, T.D., Shuster, W., Hunt, W.F., Ashley, R., Butler, D., Arthur, S., Trowsdale, S., Barraud, S., Semadeni-Davies, A., Bertrand-Krajewski, J.-L., Mikkelsen, P.S., Rivard, G., Uhl, M., Dagenais, D. and Viklander, M., 2015. SUDS, LID, BMPs, WSUD and more – The evolution and application of terminology surrounding urban drainage. *Urban Water Journal* 12(7), 525-542.
Gibbons, J.D. and Chakraborti, S., 2011. Nonparametric statistical inference. *Springer*, USA.
Gogate, N.G., Kalbar, P.P. and Raval, P.M., 2017. Assessment of stormwater management options in urban contexts using Multiple Attribute Decision-Making. *Journal of Cleaner Production* 142, 2046-2059.
Hatt, B.E., Deletic, A. and Fletcher, T.D., 2006. Integrated treatment and recycling of stormwater: a review of Australian practice. *Journal of Environmental Management* 79(1), 102-113.
Hatt, B.E., Fletcher, T.D. and Deletic, A., 2009. Hydrologic and pollutant removal performance of stormwater biofiltration systems at the field scale. *Journal of Hydrology* 365(3-4), 310-321.
Hewitson, B. and Crane, R., 1996. Climate downscaling: techniques and application. *Climate Research* 7(2), 85-95.
Imteaz, M.A., Rauf, A. and Aziz, M.A., 2011. eTank: A Decision Support Tool for optimizing rainwater tank size Perth, Australia.
IPCC, 2013. Climate Change 2013: The Physical Science Basis. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
Jones, R.N., 2000. Managing uncertainty in climate change projections–issues for impact assessment. *Climatic Change* 45(3-4), 403-419.
Kang, N., Kim, S., Kim, Y., Noh, H., Hong, S. and Kim, H., 2016. Urban Drainage System Improvement for Climate Change Adaptation. *Water* 8(7), 268.
Kuller, M., Bach, P.M., Ramirez-Lovering, D. and Deletic, A., 2018. What drives the location choice for water sensitive infrastructure in Melbourne, Australia? *Landscape and Urban Planning* 175, 92-101.
Liaw, C.H. and Tsai, Y.L., 2004. Optimum Storage Volume of Rooftop Rain Water Harvesting Systems for Domestic Use. *Journal of The American Water Resources Association*.
Licznar, P., De Michele, C. and Adamowski, W., 2015. Precipitation variability within an urban monitoring network via microcanonical cascade generators. *Hydrol. Earth Syst. Sci.* 19(1), 485-506.
McIntyre, N., Al-Qurashi, A. and Wheater, H., 2007. Regression analysis of rainfall–runoff data from an arid catchment in Oman. *Hydrological Sciences Journal* 52(6), 1103-1118.
Milly, P.C.D., Betancourt, J., Falkenmark, M., Hirsch, R.M., Kundzewicz, Z.W., Lettenmaier, D.P. and Stouffer, R.J., 2008. Stationarity Is Dead: Whither Water Management? *Science* 319(5863), 573-574.
Mitchell, V.G., McCarthy, D.T., Deletic, A. and Fletcher, T.D., 2008. Urban stormwater harvesting – sensitivity of a storage behaviour model. *Environmental Modelling & Software* 23(6), 782-793.
Müller, H. and Haberlandt, U., 2018. Temporal rainfall disaggregation using a multiplicative cascade model for spatial application in urban hydrology. *Journal of Hydrology* 556, 847-864.
Ochoa-Rodriguez, S., Wang, L.-P., Gires, A., Pina, R.D., Reinoso-Rondinel, R., Bruni, G., Ichiba, A., Gaitan, S., Cristiano, E., van Assel, J., Kroll, S., Murlà-Tuyls, D., Tisserand, B., Schertzer, D., Tchiguirinskaia, I., Onof, C., Willems, P. and ten Veldhuis, M.-C., 2015. Impact of spatial and temporal resolution of rainfall inputs on urban hydrodynamic modelling outputs: A multi-catchment investigation. *Journal of Hydrology* 531, 389-407.
Olsson, J., Berggren, K., Olofsson, M. and Viklander, M., 2009. Applying climate model precipitation scenarios for urban hydrological assessment: A case study in Kalmar City, Sweden. *Atmospheric research* 92(3), 364-375.
Pilgrim, D., 2007. Australian Rainfall and Runoff: A Guide to Flood Estimation. *Institution of Engineers, Australia* Canberra, Australia.
Prudhomme, C., Reynard, N. and Crooks, S., 2002. Downscaling of global climate models for flood frequency analysis: where are we now? *Hydrological Processes* 16(6), 1137-1150.
Raut, B., de la Fuente, L., Seed, A., Jakob, C. and Reeder, M.J., 2012. Application of a space-time stochastic model for downscaling future rainfall projections, , 579-586, ACT Australia.

Raut, B.A., Reeder, M.J. and Jakob, C., 2016. Trends in CMIP5 Rainfall Patterns over Southwestern Australia. *Journal of Climate* 30(5), 1779-1788.

Raut, B.A., Reeder, M.J., Jakob, J. and Seed, A.W., Submitted manuscript. Stochastic Space-Time Downscaling of Rainfall Using Event-Based Multiplicative Cascade Simulations. *Journal of Geophysical Research: Atmospheres*.

Raut, B.A., Seed, A.W., Reeder, M.J. and Jakob, C., 2018. A Multiplicative Cascade Model for High-Resolution Space-Time Downscaling of Rainfall. *Journal of Geophysical Research: Atmospheres* 123(4), 2050-2067.

Rosenberg, E.A., Keys, P.W., Booth, D.B., Hartley, D., Burkey, J., Steinemann, A.C. and Lettenmaier, D.P., 2010. Precipitation extremes and the impacts of climate change on stormwater infrastructure in Washington State. *Climatic Change* 102(1), 319-349.

Rummukainen, M., 1997. Methods for statistical downscaling of GCM simulations. *Swedish Meteorological and Hydrological Institute*, Norrköping, Sweden.

Seed, A.W., Srikanthan, R. and Menabde, M., 1999. A space and time model for design storm rainfall. *Journal of Geophysical Research: Atmospheres* 104(D24), 31623-31630.

Semadeni-Davies, A., Hernebring, C., Svensson, G. and Gustafsson, L.-G., 2008. The impacts of climate change and urbanisation on drainage in Helsingborg, Sweden: Suburban stormwater. *Journal of Hydrology* 350(1–2), 114-125.

Sharma, A.K., Vezzaro, L., Birch, H., Arnbjerg-Nielsen, K. and Mikkelsen, P.S., 2016. Effect of climate change on stormwater runoff characteristics and treatment efficiencies of stormwater retention ponds: a case study from Denmark using TSS and Cu as indicator pollutants. *SpringerPlus* 5(1), 1984.

Steffen, W., Hughes, L., Alexander, D. and Rice, M., 2017. Cranking up the intensity: Climate change and extreme weather events *Climate Council of Australia Ltd*.

Thoeun, H.C., 2015. Observed and projected changes in temperature and rainfall in Cambodia. *Weather and Climate Extremes* 7, 61-71.

Urich, C., Bach, P.M., Sitzenfrei, R., Kleidorfer, M., McCarthy, D.T., Deletic, A. and Rauch, W., 2013. Modelling cities and water infrastructure dynamics. *Engineering Sustainability* 166(5), 301-308.

van Vuuren, D.P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G.C., Kram, T., Krey, V., Lamarque, J.-F., Masui, T., Meinshausen, M., Nakicenovic, N., Smith, S.J. and Rose, S.K., 2011. The representative concentration pathways: an overview. *Climatic Change* 109(1), 5.

VSC, 1999, Urban Stormwater: Best Practice Environmental Management Guidelines (BPEMG). CSIRO Publishing.

Wang, M., Zhang, D.Q., Su, J., Trzcinski, A.P., Dong, J.W. and Tan, S.K., 2017. Future Scenarios Modeling of Urban Stormwater Management Response to Impacts of Climate Change and Urbanization. *CLEAN – Soil, Air, Water* 45(10), 1700111-n/a.

Zahmatkesh, Z., Karamouz, M., Goharian, E. and Burian Steven, J., 2015. Analysis of the Effects of Climate Change on Urban Storm Water Runoff Using Statistically Downscaled Precipitation Data and a Change Factor Approach. *Journal of Hydrologic Engineering* 20(7), 05014022.

Zhang, K., Randelovic, A., Aguiar, L.M., Page, D., McCarthy, D.T. and Deletic, A., 2015. Methodologies for Pre-Validation of Biofilters and Wetlands for Stormwater Treatment. *PLoS ONE* 10(5), e0125979.

Zhang, K., Randelovic, A., Page, D., McCarthy, D.T. and Deletic, A., 2014. The validation of stormwater biofilters for micropollutant removal using in situ challenge tests. *Ecological Engineering* 67, 1-10.

Highlights

- High resolution catchment scale rainfall predictions were generated from 8 GCMs
- Multiple GCMs is preferable over a best model to account for the future variability
- Simulated climate change has limited impact on pollutant treatment performance
- Larger WSUD systems are recommended to account for the future variability
- Bigger rainwater tank is suggested for same harvesting reliability in the future