Development of an Automatic Calibration Framework for Hydrologic Modelling Using Approximate Bayesian Computation

A. Chowdhury, P. Egodawatta, J. M. McGree, A. Goonetilleke

Abstract—Hydrologic models are increasingly used as tools to predict stormwater quantity and quality from urban catchments. However, due to a range of practical issues, most models produce gross errors in simulating complex hydraulic and hydrologic systems. Difficulty in finding a robust approach for model calibration is one of the main issues. Though automatic calibration techniques are available, they are rarely used in common commercial hydraulic and hydrologic modelling software e.g. MIKE URBAN. This is partly due to the need for a large number of parameters and large datasets in the calibration process. To overcome this practical issue, a framework for automatic calibration of a hydrologic model was developed in R platform and presented in this paper. The model was developed based on the time-area conceptualization. Four calibration parameters, including initial loss, reduction factor, time of concentration and time-lag were considered as the primary set of parameters. Using these parameters, automatic calibration was performed using Approximate Bayesian Computation (ABC). ABC is a simulation-based technique for performing Bayesian inference when the likelihood is intractable or computationally expensive to compute. To test the performance and usefulness, the technique was used to simulate three small catchments in Gold Coast. For comparison, simulation outcomes from the same three catchments using commercial modelling software, MIKE URBAN were used. The graphical comparison shows strong agreement of MIKE URBAN result within the upper and lower 95% credible intervals of posterior predictions as obtained via ABC. Statistical validation for posterior predictions of runoff result using coefficient of determination (CD), root mean square error (RMSE) and maximum error (ME) was found reasonable for three study catchments. The main benefit of using ABC over MIKE URBAN is that ABC provides a posterior distribution for runoff flow prediction, and therefore associated uncertainty in predictions can be obtained. In contrast, MIKE URBAN just provides a point estimate. Based on the results of the analysis, it appears as though ABC the developed framework performs well for automatic calibration.

Keywords—Automatic calibration framework, approximate Bayesian computation, hydrologic and hydraulic modelling, MIKE URBAN software, R platform.

A. Chowdhury and P. Egodawatta are with the School of Earth, Environment and Biological Sciences, Queensland University of Technology, Australia (e-mail: a.chowdhury@qut.edu.au, p.egodawatta@qut.edu.au).

J. M. McGree is with the School of Mathematical Sciences, Queensland University of Technology (e-mail: james.mcgree@qut.edu.au).

A. Goonetilleke is with the School of Civil Engineering and Built Environment, Queensland University of Technology (e-mail: a.gooonetilleke@qut.edu.au).

I. INTRODUCTION

HYDROLOGIC models are commonly used to simulate rainfall-runoff processes. Hydrologic model is generally a combination of runoff generation model and runoff routing model. Runoff generation model converts rainfall to surface runoff and runoff routing model translates the surface runoff into the catchment outlet. Most urban hydrologic models typically accompany a hydraulic model component in order to simulate stormwater runoff in pipe and channel systems. Hydraulic model mainly replicates the conveyance of flow using the concepts of open channel flow hydraulic theories [1].

Hydrologic models are developed based on different conceptualizations such as time-area method and storage reservoir routing method [1]. Irrespective of the conceptualization used, the model is built based on mathematical equations that replicate common hydrologic processes in a catchment, utilizing a range of parameters to represent different catchments, rainfall and drainage characteristics. Though there is a specific physical meaning, there are situations where these parameters are difficult to derive using geographic, climatic or catchment characteristics [2]. Hence, typical values are often used during simulation of models.

Due to critical involvement of model outputs in decision making, there is a necessity to develop an appropriate approach to obtain estimates of model parameters. In this regard, model calibration plays a vital part. A range of researchers have argued that the approach to be used for model calibration is a critical factor influencing the accuracy of model outcomes [3]. In this context, developing an automatic calibration procedure for urban hydrologic models can be regarded as important. Having an automatic calibration procedure can eliminate the 'trial and error' processes that are adopted in manual calibration so that there is no influence exerted by subjective selections. Furthermore, automatic calibration procedure can produce more reliable model outcomes. However, most commercial modelling software does not provide automatic calibration facilities [4]. These may be due to the involvement of large number of parameters and large datasets in the calibration process. It is also difficult to integrate user selected automatic calibration procedures with these software packages due to their 'black box' nature of conceptualization and model structure. This highlights the importance of developing a specialized automatic calibration procedure that can be used as a tool to obtain reliable model

parameters.

Most of the existing parameter estimation methods are formed typically based on a likelihood function. However, in scenarios where the model is a complex integration of equations with a range of parameters such as a hydrologic model, the likelihood function may be challenging and/or computationally expensive to compute [5]. To overcome this, this paper introduces an innovative calibration framework using Approximate Bayesian Computation (ABC) for use in urban hydrologic models. The main advantage of using this framework is that ABC enables one to perform Bayesian inference in calibration. Hence, the developed ABC framework can be applied for automatic calibration of desired parameters in hydrologic modelling.

II. MATERIALS AND METHODS

A. Study Sites

The study sites were selected primarily based on the availability of geographical and meteorological data relating to the catchment, drainage networks and rainfall, and monitored data relating to stormwater runoff. The selected study sites are situated in Gold Coast region, south east Queensland. They are three small catchments, namely, Alextown, Gumbeel and Birdlife Park, situated within the Highland Park residential area are shown in Fig. 1. The baseline data relating to catchment and drainage networks were obtained from Gold Coast City Council (GCCC) data bases. Rainfall and runoff data were obtained from the catchment monitoring program established in these three catchment from 2002 to 2004.
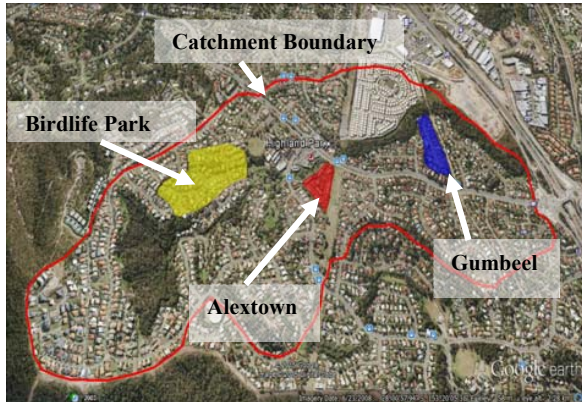

Fig. 1 Locations of Study Catchments within Highland Park Residential Area [21]

The required catchment data for hydrologic simulation included land use, slope, catchment areas, the proportion of impervious areas and land cover. This information was collected from the review of previous study conducted by [7]. The required drainage system network data for hydraulic simulation included gully pit locations, size, pipe diameter, length, ground elevation and invert elevation.

B. Specialized Modelling Using R Package

Hydrologic model was developed using R software based on time-area conceptualization. Modelling with R is more flexible than using other programming software. This is because R provides a powerful platform and language for statistical computing and graphics. Currently R is widely used for research in a diversity of sectors due to its flexibility in data manipulation, usability and additional functionality facilitated by user created software packages. For example, the "EasyABC" package within R enables the application of ABC methods for data analysis [8]. In this regard, the use of R in combination with the R-Studio interface provides greater flexibility [9].

The developed model uses time-area method as the basic conceptualisation. The time-area diagram represents the relationships between contributing catchment area and the runoff rate. To develop a time-area diagram, the catchments time of concentration is divided into a number of equal time steps. For each time step, contributing catchment area is estimated and demarcated as zones delimited by isochrones line [10]. The delimited subareas by the isochrones can be measured and plotted as time-area diagram are shown in Fig. 2.
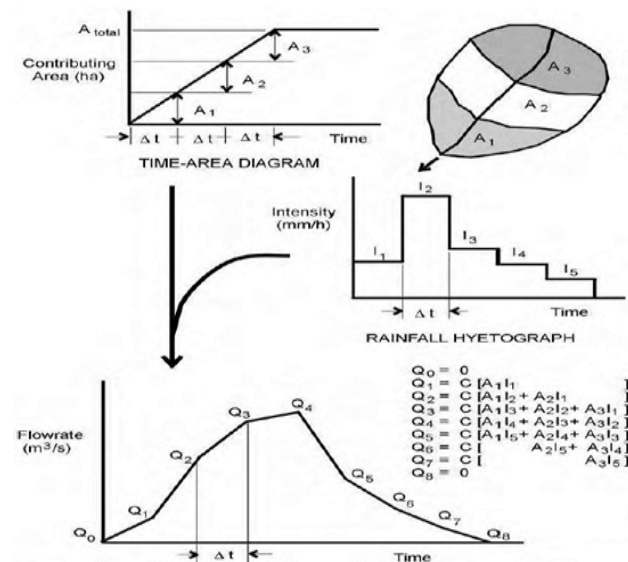

Fig. 2 Time-area routing model [10]

Time-area relationship is typically considered as linear. However, it may be considered as concave or convex shape depending on the catchment geometry and other influential factors to represent non-linearity [10]. For consistency of the model, the time step of the effective storm hyetograph is considered as same as the time step of the time-area diagram as shown in Fig. 2. The routing procedure adopted in time-area method calculates each flow time separately depending on the physical characteristics of each delineated zone [10]. At the end of each time interval, the partial flow is computed by the product of effective rainfall and the contributing subarea of the catchment. The desired runoff hydrograph is then produced by the summation of the partial flows as illustrated in Fig. 2.

The advantage of time-area routing models is twofold. Firstly, the user has the freedom to specify the time-area diagram prior to any modelling being undertaken. Secondly, only a few parameters require physical interpretation [4]. Due to this, the time-area routing model provides greater flexibility and requires less computational effort compared to other physically and conceptually based models. MIKE URBAN (MOUSE), ILSAX and DRAINS are commonly used hydrologic models in Australia that have been developed based on the time-area routing procedure [4], [10].

Hydraulic models are commonly used for pipe and channel flow routing. The hydraulic model developed in R software was based on the time-lag method. Time-lag method computes the maximum travel time separately through a pipe or channel from inlet points to outlet for each partial flow [11]. Then the individual partial flows were lagged using corresponding lag time and aggregated to produce the final hydrograph at the pipe outlet.

To enable the use of the EasyABC package with developed models in R, datasets were prepared in specific formats. The developed model in R was associated with four calibration parameters. These are initial loss (IL), time of concentration (TC), reduction factor (RF) and time-lag (TL). The calibration framework was developed by combining the developed model with EasyABC package functions in R. Hence, the developed framework was used for automatic calibration through Bayesian inference. The calibration was an iterative process and R facilitates the parameters estimation by enabling a large number of iterations within the given parameters range [12]. The posterior distribution of parameters can be displayed graphically by drawing a histogram or density curve. The accuracy of simulation result can also be evaluated using the upper and lower 95% credible intervals in R.

C. Mike Urban Modelling

To validate the model developed in R, a separate set of models were also developed using commercial modelling software, MIKE URBAN, developed by the Danish Hydraulic Institute (DHI). MIKE URBAN is an Arc GIS-supported hydrologic and hydraulic process simulation software. The software provides MOUSE engine which enables hydrologic simulation based on selected time-area routing procedures. Accordingly, three separate models for the selected three catchments were developed using MIKE URBAN software based on the collected catchment and drainage network data. An extensive review of published literature and desktop study was undertaken to facilitate the MIKE URBAN modelling task. The software is widely used and accepted among the research community [4].

D. Bayesian Inference via Approximate Bayesian Computation

Bayesian inference is a special technique of statistical inference developed based on Bayes' theory. In typical practice, Bayesian inference determines the posterior distribution of the parameters by the combination of two probability distributions: the prior distribution and the likelihood function. Typical expression for posterior distribution is presented in (1) [13]:

$$P(\theta \mid y) = \frac{p(y \mid \theta)\pi(\theta)}{p(y)} \qquad (1)$$

where, $\pi(\theta)$ is the prior distribution of parameters, $P(\theta|y)$ is the posterior distribution of parameters, $p(y|\theta)$ is the likelihood, and $p(y)$ is the marginal likelihood or model evidence.

In all model-based applications that use Bayesian statistical inference, the likelihood function is of particular concern. For simple models, likelihood function can typically be determined straight forwardly. However, most hydrologic models are complex and contain a range of parameters directly expressing features of natural processes. As such, the likelihood function becomes more difficult to define and compute [5], [14]. This leads to complexities in deriving or sampling of parameters from posterior distributions. To overcome this challenge, Approximate Bayesian Computation (ABC) is used. ABC is a simulation-based approach to avoid evaluation of the likelihood function by comparing observed data with simulated data [6]. Due to this, ABC has gained popularity over the last few decades in different sectors of biological sciences such as population genetics, epidemiology, ecology, and systems biology [5], [6], [15].

The likelihood free computation constitutes a process for accepting parameter values from a chain of parameter proposals once they satisfy the criteria for comparing observed and simulated data [16]. The process is dependent on the probability of matching observed and simulated data [14]. If the probability of matching data is very small, a tolerance limit between simulated and observed data may be allowed to increase at an acceptable rate. The tolerance limit is the distance that determines the level of discrepancy between each pair of observed and simulated data in each simulation. If the data are continuous and in highly dimensional, then one needs to reduce the dimensionality of problem as poor acceptance rates will be observed or high tolerance levels will be needed, both of which are not desirable. The dimensionality problem arises if a sample consists of many univariate or multivariate measurements in analysis. Hence, as such situations summary statistics can be used to reduce the dimensionality of the problem [17]. Summary statistics reduce the dimensionality of data by generating datasets with small distance to observed and simulated data. Generally, low dimensional summary statistics are used to capture required information in comparison between observed and simulated data to increase more acceptance rate under low tolerance levels. Hence posterior predictions of parameters and runoff results are obtained through ABC from accepted parameters values and corresponding simulation results. For this study, standard rejection algorithms were developed as noted below, based on the tolerance limit and summary statistics [17].

Given observed data y, the following algorithm is repeated until a total number of simulations have been performed
1. Draw $\theta \sim \pi(\theta)$.
2. Simulate $x_i \sim p(x|\theta_i)$.

3. Accept θ_i if $\rho(S(x_i), S(y)) < \epsilon$.
4. Otherwise, reject and repeat the process.

E. Statistical Validation

In hydrologic modelling, model's performance is judged to be acceptable if it satisfies specific statistical evaluation criteria as shown in (2)-(4) [18]. These statistical evaluations include coefficient of determination (CD), root mean square error (RMSE) and maximum error (ME). These evaluations are necessary to judge the best fit of calibrated parameters with the developed model by comparing measured and simulated results. The mathematical expressions of these measurements are given below:

$$CD = 1 - \left[\frac{\sum_{i=1}^{n}(S_i - O_i)^2}{\sum_{i=1}^{n}(O_i - \bar{O})^2} \right] \qquad (2)$$

$$RMSE = \sum_{i=1}^{n} \left[\frac{(S_i - O_i)^2}{n} \right]^{\frac{1}{2}} \qquad (3)$$

$$ME = Max \left| S_i - O_i \right|_{i=1}^{n} \qquad (4)$$

where, n = Total number of observations, S = Simulated posterior mean results, O = Observed results, Ō = Mean of the observed results. CD value is a measure of the proportion of the total variance of observed data explained by the predicted data. ME value measures the maximum errors for any specific time steps from the total observation. The over-estimation or under-estimation in comparison to observed values are indicated by the RMSE value. CD, RMSE and ME values close to 1, 0 and 0, respectively, means negligible error between the observed and simulated results [19].

The evaluation of model performance by graphical presentations can also be useful for viewing distribution patterns. The graphical presentations can include trends analysis of measured and simulated data [18]. For example, credible intervals are an estimate in statistics for determining parameter uncertainty in broader application of models. The upper and lower 95% credible intervals are typically used to judge the accuracy of model calibration and statistical validation of the model. These methods were adopted for comparison of outputs from the model developed in R with MIKE URBAN simulation outcomes.

III. RESULT AND DISCUSSIONS

Required rainfall and runoff data for model calibrations were selected by a careful inspection of the available measured rainfall events and corresponding runoff data. Other input data such as catchment subdivision areas and impervious fractions were also prepared in the required form. Details of the selected input data for three study catchments are shown in Table I. The same input data were used for MIKE URBAN simulations so that the model developed in R can be validated. The developed model has four calibration parameters. For

each parameter (IL, TC, RF and TL), suitable log-normal distributions were chosen as prior distributions based on the historical rainfall patterns, and to draw the positive parameter values. During calibration, four parameter values were drawn from prior distributions for each independent sampling. By this way, 200,000 samples were drawn independently from each distribution.

TABLE I
SELECTED RAINFALL, RUNOFF EVENTS, CATCHMENT SUBDIVISIONS, IMPERVIOUS FRACTIONS FOR CALIBRATION AND CALIBRATED PARAMETER VALUES

Symbol	Alextown		Gumbel		Birdlife Park	
Selected rainfall events	10		10		10	
Selected runoff events	10		10		10	
Catchment subdivisions	27		2		54	
Impervious fractions (%)	57.2		45.8		40.7	
Calibrated parameters value	Range	Mean	Range	Mean	Range	Mean
IL	0.82-1.12	1.03	0.81-1.12	0.97	0.09-0.10	0.095
TC	19.9-30.6	23.8	27.9-43.0	36.4	15.6-31.2	23.0
RF	0.65-1.0	0.87	0.10-0.23	0.15	0.62-1.45	1.0
TL	0.67-1.61	1.05	11.8-17.8	14.3	0.08-0.12	0.10

IL = initial loss, TC = time of concentration, RF = reduction factor, TL = time-lag.

Calibration was performed on single event basis using high performance computing system (HPC) undertaking 200,000 iterations using the standard rejection ABC algorithm. The standard rejection ABC algorithm shown in Section D above was performed in four steps. Generally, ABC rejection method does not require to specify tolerance limit because it is implied once to choose the proportion of samples to keep. Accordingly, best 40,000 samples and corresponding simulation results were retained. Each single event calibration produces posterior distributions of four calibrated parameters. The minimum and maximum range of calibrated parameter values were given in Table I. Finally, the posterior mean values for the selected four parameters were determined by combining the individual posterior distributions for the selected rainfall events for the three catchment models. The distribution of the calibrated parameters for Alextown catchment is presented as histograms and the mean values are indicated by a vertical line in Fig. 3.

As shown in Fig. 3, all parameters have positive skewed distributions. For validation purpose the posterior mean values of calibrated parameters were used as input for MIKE URBAN model simulations for the three study catchments.

Model validation is the step followed to confirm that the model developed for specific rainfall events is capable of accurately predicting runoff flow for different rainfall events [20]. In this study, the developed models using R were validated by comparing its' simulated mean flow obtained from ABC posterior prediction with MIKE URBAN simulation outcomes. Figs. 4-6 show example outputs from calibrated models for one of the selected events in Alextown, Gumbel and Birdlife Park catchments. As seen in Figs. 4-6, the calibration results of R show strong agreement with observed runoff data and mostly within the upper and lower

95% credible intervals. It is also noticeable that MIKE URBAN model outcomes are closely comparable to R model outcomes, and also falls within the upper and lower 95% credible intervals.
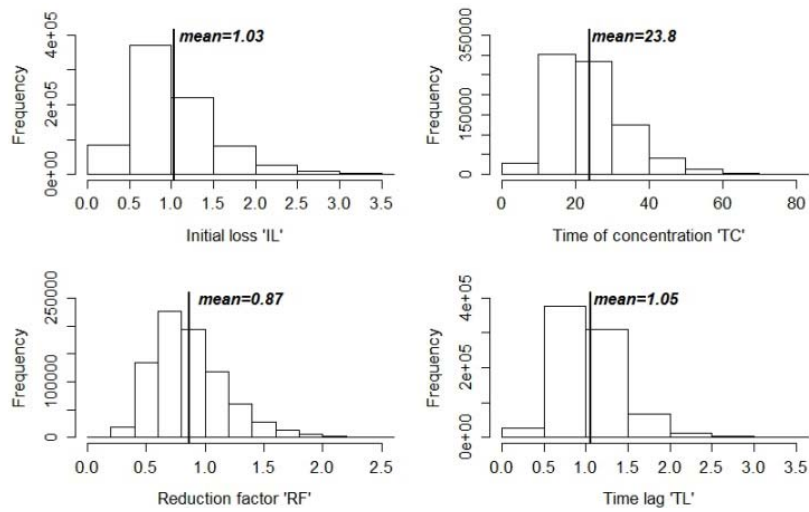

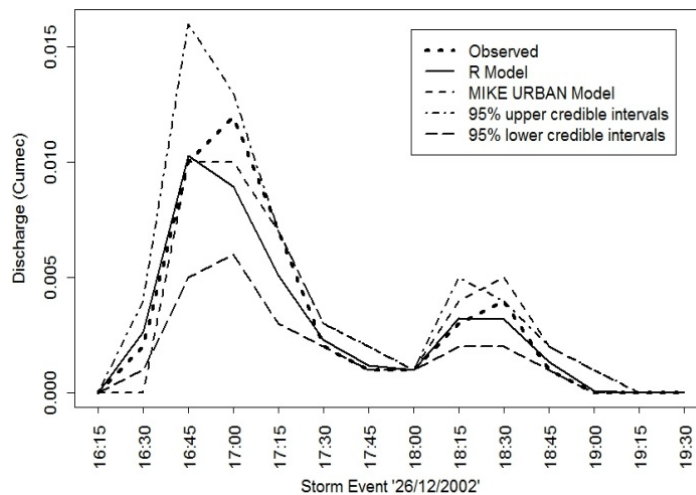Fig. 3 Histogram of calibrated parameters for Alextown

Fig. 4 Calibration plot of single rainfall event for Alextown catchment
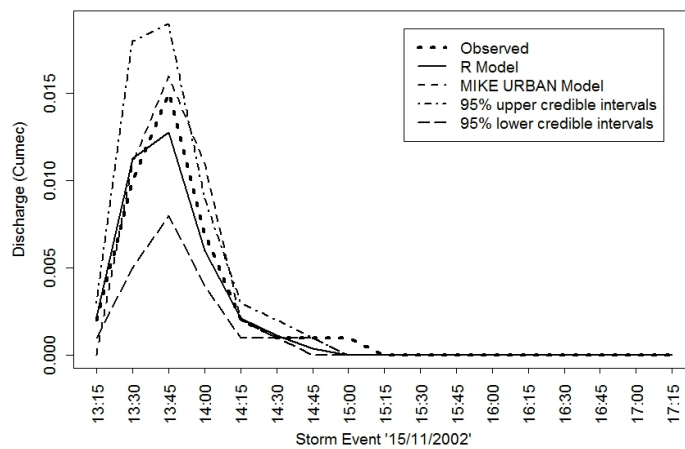

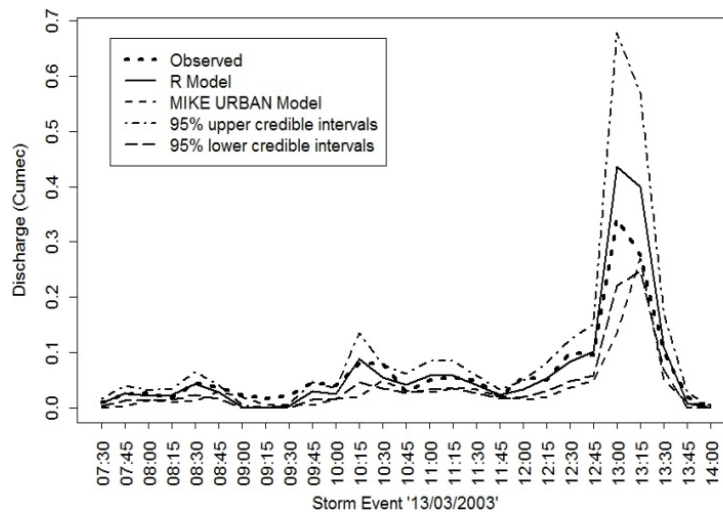Fig. 5 Calibration plot of single rainfall event for Gumbel catchment

Fig. 6 Calibration plot of single rainfall event for Birdlife Park catchment

TABLE II
STATISTICAL VALIDATION OF RUNOFF CALIBRATION RESULT

Study Sites	Rainfall Events	CD	RMSE	ME
Alextown	21 - 08 - 2002	0.90	0.015	0.014
	10 - 12 - 2002	0.91	0.015	0.009
	26 - 12 - 2002	0.93	0.002	0.003
	1 - 02 - 2003	0.91	0.004	0.006
	1 - 03 - 2003	0.91	0.007	0.012
	7 - 03 - 2003	0.96	0.003	0.004
	12 - 03 - 2003	0.94	0.005	0.010
	27 - 04 - 2003	0.90	0.018	0.010
	24 - 10 - 2003	0.93	0.020	0.027
	14 - 12 - 2003	0.92	0.013	0.024
Gumbel	28 - 04 - 2002	0.83	0.002	0.002
	5 - 05 - 2002	0.92	0.001	0.000
	2 - 06 - 2002	0.92	0.001	0.001
	16 - 06 - 2002	0.81	0.002	0.002
	21 - 09 - 2002	0.84	0.001	0.002
	13 - 11 - 2002	0.86	0.005	0.004
	15 - 11 - 2002	0.97	0.002	0.002
	10 - 12 - 2002	0.92	0.003	0.005
	13 - 03 - 2003	0.88	0.002	0.001
	24 - 10 - 2003	0.93	0.004	0.004
Birdlife Park	28 - 04 - 2002	0.92	0.041	0.039
	29 - 04 - 2002	0.89	0.020	0.018
	3 - 05 - 2002	0.93	0.034	0.021
	4 - 05 - 2002	0.86	0.021	0.023
	2 - 06 - 2002	0.91	0.009	0.011
	4 - 06 - 2002	0.93	0.007	0.009
	27 - 08 - 2002	0.95	0.004	0.008
	27 - 10 - 2002	0.91	0.017	0.014
	13 - 11 - 2002	0.90	0.088	0.100
	27 - 04 - 2003	0.90	0.033	0.050

CD = coefficient of determination, RMSE = root mean square error, ME = maximum error.

For statistical validation of the developed models, selected statistical measures were performed for all the observed and simulated runoff events and the results are presented in Table II. The CD values for the three catchments are close to 1, and

RMSE and ME values ranges from 0.001 to 0.088 and 0.000 to 0.050, respectively. Statistical measures obtained are graphically presented in the form of Box-Whisker plots in Figs. 7-9. As seen in Fig. 7, median CD value falls above 0.90 for all three catchments. Most of the CD values fall within the 3rd quartile (corresponds to 75thpercentile) zone for Alextown and Birdlife Park catchments, and for Gumbel catchment, falls within the1stquartile (corresponds to 25thpercentile) zone. This means that the fluctuations in CD values were greater in Gumbel than the other catchments and ranges within 0.81 to 1. As seen in Figs. 8-9, medians of RMSE and ME are close to 0.02 for Birdlife Park catchment, and relatively lower for the other catchments. The variation is within 0.004 to 0.088. The minimum value for both RMSE and ME was found to be smaller for Gumbel catchment and fluctuates within 0.001 to 0.005. The statistical evaluation outcomes suggest that the agreement between observed and simulated results obtained for flow discharges is appropriate. This highlights that the developed automatic calibration framework is capable of producing reliable results for all of the rainfall events selected for the three catchments.
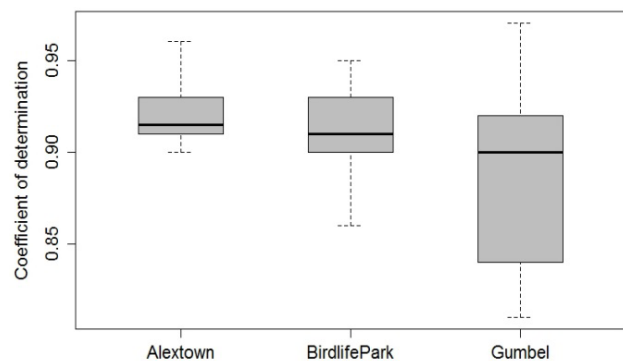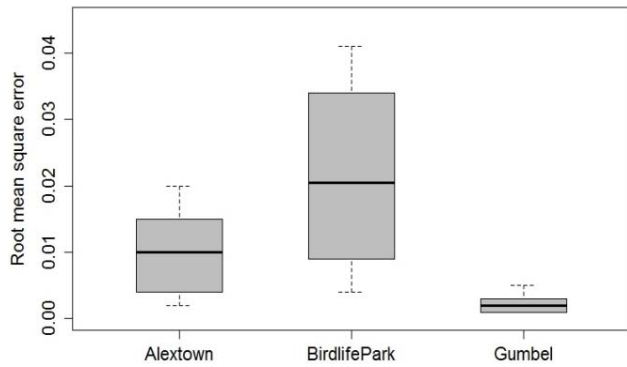

Fig. 7 Plot of CD for study catchments

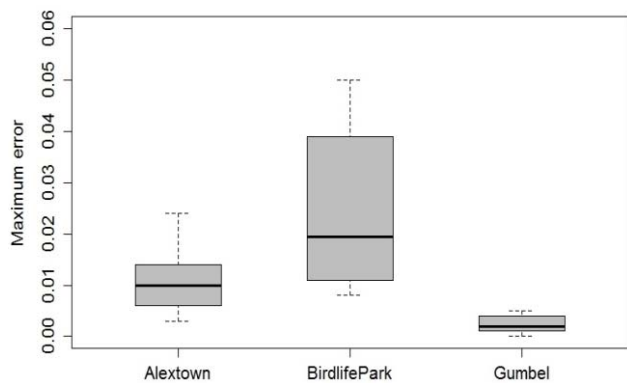Fig. 8 Plot of RMSE for study catchments

Fig. 9 Plot of ME for study catchments

IV. CONCLUSIONS

The study presented the development of an automatic calibration framework in R software using ABC for urban hydrologic modelling. To develop the automatic calibration framework, hydrologic and hydraulic models were developed in R software and then combined with EasyABC package function. The analysis results revealed that the developed calibration framework performs well for all rainfall events selected at the three study catchments within the upper and lower 95% credible intervals. The results are influenced by the choice of priors during calibration. Choice of summary statistics influences the computational efficiency of ABC method and may lead to the source of the ABC error. Hence, as typical practice, mean of observed and simulated flow are used as lower dimensional summary statistics to capture the relevant information in comparing observed and simulated flow. For validation of the model, three separate models were also developed using MIKE URBAN software. The simulation outcomes of the developed model in R were compared with MIKE URBAN model outcomes for all the study catchments. The graphical comparison showed strong agreement with MIKE URBAN simulation results. To judge the accuracy of models, statistical measures of simulated and observed results were evaluated. The statistical measures showed reasonable agreement between observed and simulated flow. For the three study catchments, CD, RMSE and ME values were found close to 1, 0 and 0, respectively. This suggests that the developed automatic calibration framework is a robust approach when compared to other classical statistical approaches often used in hydrologic modelling.

From the outcomes of this study, it can be concluded that the developed ABC calibration framework is suitable to be used with complex hydrologic models where likelihood is intractable and/or expensive to compute, and where is it straight forward to simulate data from the model. ABC is an approximate method and has an additional source of error. Specifically, this additional source of error comes from the tolerance level. When a small tolerance level is used, this additional error is reduced. However, it may be difficult to simulate parameter values that yield simulated data which are similar to the observed data. Consequently, ABC typically suffers from small acceptance rates, which leads to a larger computational burden to obtain a reasonable number of samples from the ABC posterior. ABC also overcomes the shortcomings of tedious trial and error processes in manual calibration. Validation results with MIKE URBAN confirmed that the developed model can be used as an alternative for commercial modelling platforms which seek automatic calibration for model calibration and parameter estimation.

REFERENCES

[1] C. Zoppou, "Review of urban storm water models", *Environmental Modelling & Software*, vol. 16, no. 3, pp. 195-231, 2001.
[2] H. Medsen, G. Wilson, H. C. Ammentorp, "Comparison of different automated strategies for calibration of rainfall-runoff models", *Journal of Hydrology*, Vol. 261, pp. 48-59, 2002.
[3] D. Pilgrim, "Model evaluation, testing and parameter estimation in hydrology", *National Symposium on Hydrology*, Australian Academy of Science, Canberra, 1975.
[4] M. URBAN, "Mike Urban Model User Manual", Danish Hydraulic, 2012.
[5] R. Rasmussen, and G. Hamilton, "An approximate Bayesian computation approach for estimating parameters of complex environmental processes in a cellular automata", *Environmental Modelling & Software*, vol. 29, no. 1, pp. 1-10, 2012.
[6] K. Csilléry, M. G. Blum, O. E. Gaggiotti, and O. François, "Approximate Bayesian computation (ABC) in practice", *Trends in ecology & evolution*, vol. 25, no. 7, pp. 410-418, 2010.
[7] P. Egodawatta, "Translation of small-plot scale pollutant build-up and wash-off measurements to urban catchment scale", Queensland University of Technology, Brisbane, PhD Thesis, 2007.
[8] F. Jabot, T. Faure and N. Dumoulin, "EasyABC: performing efficient approximate Bayesian computation sampling schemes using R", *Methods in Ecology and Evolution*, vol. 4, no. 7, pp. 684-687, 2013.
[9] P. Torfs, and C. Brauer, "A (very) short introduction to R", *Hydrology and Quantitative Water Management Group*, Wageningen University, The Netherlands, 2014.
[10] G. O'Loughlin and B. Stack, "DRAINS user manual", Watercom Pty Ltd, 2014.
[11] S. C. Service, "Time of Concentration", In National Engineering Handbook: Natural Resources Conservation Service, 2010.
[12] K. Csilléry, M. G. Blum, O. E. Gaggiotti, and O. François, "Approximate Bayesian computation (ABC) in practice", *Trends in ecology & evolution*, vol. 25, no. 7, pp. 410-418, 2010.

[13] A. E. Raftery, and J. Gill, "Bayesian statistics: one-day course for the American Sociological Association", *Bayesian statistics*, pp. 1-61, 2002.

[14] M. A. Beaumont, "Approximate Bayesian computation in evolution and ecology" *Annual review of ecology, evolution, and systematics*, vol. 41, pp. 379-406, 2010.

[15] M. A. Beaumont, W. Zhang, and D. J. Balding, "Approximate Bayesian computation in population genetics", *Genetics*, vol. 162, no. 4, pp. 2025-2035, 2002

[16] P. Marjoram, J. Molitor, V. Plagnol, and S. Tavaré, "Markov chain Monte Carlo without likelihoods", Proceedings of the National Academy of Sciences, Vol. 100, No. 26, pp. 15324-15328, 2003.

[17] J. K. Pritchard, M. T. Seielstad, A. Perez-Lezaun, and M. W. Feldman, "Population growth of human Y chromosomes: a study of Y chromosome microsatellites", *Molecular Biology and Evolution*, vol. 16, no. 12, pp. 1791-1798, 1999.

[18] K. Loague and R. E. Green, "Statistical and graphical methods for evaluating solute transport models: overview and application", *Journal of contaminant hydrology*, vol. 7, no. 1, pp. 51-73, 1991.

[19] I. Green and D. Stephenson, "Criteria for comparison of single eventmodels", *Hydrological Sciences Journal*, vol. 31, no. 3, pp. 395-411, 1986.

[20] H. J. Henriksen, L. Troldborg, P. Nyegaard, T. O. Sonnenborg, J. C. Refsgaard and B. Madsen, "Methodology for construction, calibration and validation of a national hydrological model for Denmark", *Journal of Hydrology*, vol. 280, pp. 52-71, 2003.

[21] G. Maps, Download link: https://www.google.com.au/maps/place/28%C2%B000%2754.4%22S+153%C2%B020%2710.1%22E/@-28.0146414,153.3353743,17z/data=!3m1!1e3