Uncertainty Analysis and Calibration of SWMM Using a Formal, Bayesian Methodology

Misgana K. Muleta

Department of Civil and Environmental Engineering, California Polytechnic State University, San Luis Obispo, California, 93407; email: mmuleta@calpoly.edu.

Abstract

Importance of uncertainty analysis (UA) to estimate the degree of reliability associated with model predictions is being understood. Consequently, literature that describes various Bayesian methods for the assessment of parameter and model predictive uncertainty has been steadily rising. Applications dealing with urban stormwater management are, however, very limited. This study demonstrates successful application of a formal Bayesian methodology for UA of the U.S. EPA Stormwater Management Model (SWMM), a widely used urban stormwater management model, and illustrates the methodology using a highly urbanized watershed in southern California. DREAM$_{(ZS)}$, a recently developed effective and efficient sampling algorithm, and a generalized, formal likelihood function that addresses the assumptions commonly made regarding error structure including independence, normality and homoscedasticity are used for the UA. Results will include comparison of the simulated error structure with the assumptions made by the likelihood function, histogram of the parameters posteriors, bounds of the 95 percent confidence interval, and the maximum likelihood (ML) predictions. A conventional calibration attempted to compare the ML results derived from the UA with the optimal solutions identified by the single objective calibration will also be presented. Besides illustrating state-of-the-art in UA, the study will highlight application of the methodology to developing a watershed management model to mitigate stormwater quantity and quality problems associated with urbanization.

Introduction

The objective of this study is to examine effectiveness of a formal Bayesian approach for uncertainty analysis and calibration of a Storm Water Management Model (SWMM5) (Rossman, 2005). The a recently developed efficient MCMC sampling scheme known as DREAM$_{(ZS)}$ (Schoups and Vrugt, 2010) and a generalized flexible likelihood function (GL) has been used to identify parameter posteriors and to estimate runoff prediction uncertainty. The methodology is illustrated using Ballona Creek watershed, a heavily urbanized watershed located in Los Angeles, California. Effectiveness of the UA method in removing heteroscedasticity and temporal correlation, and in identifying representative pdf for the residuals has been scrutinized. To examine robustness of the UA method to also identify the optimal solutions typically sought by classic calibration approaches, the UA solution (i.e., the maximum likelihood (ML) parameter set and the associated runoff predictions) have been compared with the solutions determined by an automated calibration algorithm known as Dynamically Dimensioned Search (DDS) (Tolson and Shoemaker, 2007).

Most UA studies in water resources that applied MCMC technique within the Bayesian framework used lumped-conceptual models for rainfall-runoff analysis of rural watersheds (Kuczera et al., 2006; Vrugt et al., 2009; Schoups and Vrugt, 2010). Few studies have been reported in more challenging areas such as spatially distributed modeling (Feyen et al. 2008; Muleta et al., 2011b). Applications to urban watersheds are very limited. Ball (2009) underlined

the need for UA based approaches for evaluation of urban drainage models in his discussion of the conventional calibration effort reported by Barco et al. (2008) on Ballona Creek watershed. Freni et al. (2008, 2009a) applied GLUE to an urban drainage model, and tested sensitivity of the solutions to likelihood measures (Freni et al., 2009a) and acceptability thresholds (Freni et al., 2008). In a separate study, Freni et al. (2009b) compared performance of Bayesian Monte Carlo method to that of GLUE. Mannina and Viviani (2010) applied GLUE for UA of stormwater quality using a conceptual, urban drainage model they developed in-house. All these applications of UA to urban drainage models used GLUE, an informal approach whose statistical validity has been questioned (Mantovan and Todini, 2006; Stedinger et al., 2008; Vrugt et al., 2008).

Uncertainty Analysis and Calibration Methods

Besides the likelihood function, a sampling scheme that efficiently identified posterior pdfs is crucial for effective application of Bayesian based UAs. Markov-Chain Monte Carlo (MCMC) schemes are often used for this application, and improving efficiency of MCMC schemes has been one of focuses of UA research in the past few years. In this regard, $DREAM_{(ZS)}$, an MCMC algorithm that capitalized on the strength of the DiffeRential Evolution Adaptive Metropolis (DREAM) (Vrugt et al., 2008), has been developed by Schoups and Vrugt (2010). Effectiveness and efficiency of $DREAM_{(ZS)}$ for posterior sampling has been reported in several studies. Schoups and Vrugt (2010) applied the GL function and $DREAM_{(ZS)}$ for rainfall-runoff analysis of two watersheds using a lumped conceptual model. Application of GL function and $DREAM_{(ZS)}$ to a more complex and spatially distributed watershed simulation model has been reported by Muleta et al. (2011b). This study examines $DREAM_{(ZS)}$ and GL for UA of SWMM5 using the Ballona Creek watershed which is one of the most urbanized watersheds in the world with about 83 percent imperviousness (Bay et al., 2003). The reader is referred to Schoups and Vrugt (2010) for further description of $DREAM_{(ZS)}$.

   Single objective automated calibration was performed, primarily, to compare solutions of the conventional model calibration technique to those identified by GL and $DREAM_{(ZS)}$. The Dynamically Dimensioned Search (DDS) (Tolson and Shoemaker 2007) was used to identify optimal values of SWMM5 runoff parameters. DDS has been developed to improve efficiency of calibrating computationally demanding models. DDS is a simple, single-objective, heuristic search method that starts by globally searching the feasible region and incrementally localizes the search space as the number of simulation approaches the maximum allowable number of simulations (the only stopping criteria used by the algorithm). Progress from global to local search is achieved by probabilistically reducing the number of model parameters modified from their best value obtained thus far. New potential solutions are created by perturbing the current parameter values of the randomly selected model parameters only. The best solution identified thus far is maintained, and are updated only when a solution with superior value of the objective function is found.

   DDS requires minimal algorithmic parameter tweaking as the only parameters to set are the maximum number of model evaluations and the scalar neighborhood size perturbation parameter (r) that defines the random perturbation size standard deviation as a fraction of the decision variable range. The recommended value of 0.2 (Tolson and Shoemaker 2007) has been used for r in this study. Efficiency and effectiveness of DDS has been reported by Tolson and Shoemaker (2007) and Muleta (2010) who compared its performance to that of widely used optimization methods including the Shuffled Complex Evolution-University of Arizona (*SCE-*

UA) (Duan et al., 1992) and the Genetic Algorithms (Holland,1975). For this study, DDS has been integrated with SWMM5 to calibrate runoff for the study watershed.

Stormwater Management Model (SWMM)

SWMM was first developed in 1971 and it continues to be widely used throughout the world for planning, analysis and design of stormwater runoff, combined sewers, sanitary sewers, and other drainage systems (Rossman, 2010). SWMM5, the latest version of SWMM5, simulates hydrology, hydraulics and water quality of urbanized and non-urbanized watersheds. The hydrologic processes modeled include precipitation (rainfall or snow fall), evaporation, surface runoff, infiltration, groundwater flow, and snow packs and snow melt. Both single event and long-term (continuous) simulations can be performed accounting for spatial and temporal variability in the climate, soil, land use and topography in the watershed. Surface runoff is estimated using the non-linear reservoir method where surface runoff occurs only when the depth of the overland flow exceeds the maximum surface storage provided by ponding, surface wetting, and interception, in which case the runoff is calculated using Manning's equation. Horton, Green-Ampt and the Curve Number methods are available to model infiltration.

Runoff quality including buildup and washoff of pollutants can be simulated using various approaches from both developed and non-developed land uses. The runoff quantity and quality simulated from a subwatershed and the wastewater loads (if any) assigned to the receiving nodes are added, and then transported using either steady, kinematic wave or dynamic wave routing through a conveyance system of pipes, channels, storage/treatment devices, pumps, and hydraulic regulators such as weirs, orifices, and other outlet types. Hydraulic conditions of any level of complexity including those experiencing backwater effect, flow reversal, and pressurized flow can be accommodated. In addition, the capability to model the commonly used Low Impact Development (LID) types including porous pavements, bioretention cells, infiltration trenches, vegetative swells and rain barrels has been recently added to SWMM5. For this study, source code of SWMM5 has been integrated with the UA and single objective calibration methods previously described.

The Ballona Creek Watershed and Data

The Ballona Creek watershed is used to illustrate the methods described in this study. Drainage area of the watershed is about 337 km^2, and its land use consists of 64% residential, 8% commercial, 4% industrial, and 17% open space (LACDPW, 2011). The upper 217 km^2 of the Ballona Creek watershed (i.e., the portion that drains to the streamflow gaging station used in this study) has been modeled here. Extensive networks of storm drains collect stormwater from the watershed and convey it to the Ballona Creek, a nine-mile long flood protection channel that discharges to the Santa Monica Bay (LACDPW, 2011). The watershed has been identified as the major source of non-point source pollution to the Santa Monica Bay (Stenstrom and Strecker,1993). The data needed to build SWMM5 have been collected from various sources. Digital elevation model, land use map and imperviousness map were obtained from the United States Geological Survey (USGS) seamless data warehouse (http://seamless.usgs.gov/), and SSURGO soil map has been obtained from the Natural Resources Conservation Service (NRCS) soil data mart (http://soildatamart.nrcs.usda.gov/).

Because of the difficulty to accurately delineate urban subwatersheds from digital elevation models alone (Gironas et al., 2010), the subwatershed delineation obtained from the

Los Angeles County Department of Public Works (LACDPW) were used for this study. The LACDPW delineation divided the watershed in to 134 subwatersheds. The number of subwatersheds was further reduced to 92 by merging smaller subwatersheds (area less than 0.41 km^2 or 100 acre) to the closest adjoining subwatershed. Subcatchment information such as area, slope, and flow length were extracted from the digital elevation model. The soil, land use and imperviousness maps were superimposed onto the subwatersheds to extract runoff modeling parameters including percent imperviousness, infiltration parameters and Manning's roughness coefficient. Rainfall data at three gages and streamflow data for a monitoring station that drains about 65 percent of the Ballona Creek watershed were obtained from the LACDPW for fifteen years (i.e., 1996-2010) at fifteen minute interval. Proximity and altitude criteria were used to define the raingage that represents each subcatchment.

The watershed consists of thousands of storm drains (underground pipes and open channels) designed for flood protection purposes. With the assumption that overland flow from each of the 92 subwatersheds directly flows to a storm drain inlet located at the outlet of the subwatershed, only 72 larger storm drains were considered in this model. Because each subwatershed may contain numerous streets, swales and minor storm drains that can play significant roles in routing of runoff and contaminants within the subcatchment, modeling the major drainage systems alone, as done in this study, can modify hydraulics of the drainage system, for example by slowing down the travel time. The approach used herein is, however, commonly used to simplify modeling complexity and to reduce the cost associated to data collection (Gironas et al., 2010). The study by Burian et al. (2000) was used for storm drain information including shape, size, slope and length. The storm drain data obtained from the LACDPW lacks these crucial data.

Methodology

Both streamflow and rainfall data are available at fifteen-minute interval. The curve number method was selected for infiltration modeling as the *CN* values (primary parameter for the curve number method) can be determined more readily, compared to Horton or Green-Ampt parameters, from the land cover and soil maps available for the watershed. Because the GL-DREAM$_{(ZS)}$ algorithm used for the UA requires running SWMM5 repetitively(up to tens of thousands of runs) to converge, computational time is a significant concern. As such, kinematic wave routing was selected to reduce computational burden of the dynamic wave routing option. Likewise, continuous (long-term) simulation was not considered because of computational demand. Short-term simulation (i.e., duration of 7 days for calibration and 6 days for verification) were used for both single objective calibration and uncertainty analysis. Most studies that reported on calibration of urban drainage models used single-event simulations with typical duration of a day or less (Barco et al., 2008; Fang and Ball, 2007). The simulation durations considered in this study for both calibration and verification cases are therefore, significant improvements compared with single event simulations. While single-event and the short-duration simulation pursued here may suffice applications such as flood control, continuous (e.g., multiple year duration) simulation models are more appropriate for applications that are sensitive to long-term watershed characteristics (e.g., contaminant buildup and washoff processes).

Total of eleven SWMM5 runoff parameters were considered for calibration and uncertainty analysis. The parameters were assumed to follow uniform distribution as done in Muleta and Nicklow (2005), and lower and upper bounds were assigned based on literature

(Rossman, 2010; Barco et al., 2008). Values of the parameters vary from subwatershed to subwatershed depending on soil, land use, imperviousness map, topography and/or other characteristics of the subwatershed. Baseline values for the parameters have been extracted for each subwatershed from the soil, land use, imperviousness and topography maps using GIS. During both calibration and uncertainty analysis, these baseline values were altered by multiplying the parameter by the values proposed by the calibration and the UA algorithm. This way, the baseline values would be scaled up or down while preserving the spatial variability determined from watershed characteristics.

The streamflow measured at the Swatelle station was used for calibration and uncertainty analysis. Mean absolute error (*MAE*) was used as objective function for the single objective calibration. *MAE* was selected as objective function based on the finding of Muleta (2011a) that compared relative effectiveness of the efficiency criteria commonly used in hydrologic modeling to describe goodness of model performances. According to the study, efficiency criteria such as MAE that describe the absolute deviation between observations and model simulations were found robust. Goodness of the calibration result was further assessed using additional efficiency criteria including the Nash-Sutcliffe efficiency (*NSE*) criterion (Nash and Sutcliffe, 1970), Percent Bias *(PBIAS)* and total volume of runoff.

Posterior distribution of the eleven runoff parameters was estimated with $DREAM_{(ZS)}$ and the GL function. Six additional error model parameters were considered for the GL. A total of 100,000 SWMM5 simulations were used to sample posterior distribution of the parameters, including the six error model parameters. Convergence of $DREAM_{(ZS)}$ to a stable posterior pdf was monitored using the *R* statistic of Gelman and Rubin (1992). Convergence is declared when $R_j \leq 1.2$ for all $j = 1,\ldots,d$, where *d* is the number of model parameters being analyzed (i.e., 17 in this study). The last 5,000 GL-$DREAM_{(ZS)}$ runs that meet the convergence criteria were extracted and parameter posteriors were determined and reported for each individual parameter.

Once the posterior distribution of the model parameters is known, runoff predictive uncertainty can be estimated by propagating the different samples of the posterior distribution through the SWMM5 model, and reporting the respective prediction uncertainty ranges (e.g., 95% confidence interval). This prediction interval, however, represents parameter uncertainty only; it doesn't consider other sources of error, including model structural, forcing data, and calibration data uncertainty. Total predictive uncertainty was calculated using the methodology described in Schoups and Vrugt (2010) based on the error model parameters determined by GL-$DREAM_{(ZS)}$. The ML parameter values determined by GL-$DREAM_{(ZS)}$ are benchmarked against the calibration results obtained using DDS, and are also compared in terms of their ability to fit different parts of the hydrograph. Further details of the methodology and the UA and calibration results will be provided during the oral presentation.

References

Ball, J. E. (2009). Discussion of "Automatic Calibration of the U.S. EPA SWMM Model for a Large Urban Catchment" by J. Barco, K. M. Wong, and M. K. Stenstrom. J. Hydraulic Engineering,135 (12): 1108-1110.

Barco, J., Wong, K.M., and Stenstrom, M.K. (2008). Automatic Calibration of the U.S. EPA SWMM Model for a Large Urban Catchment. J. Hydraulic Engineering,134 (4): 466-474.

Burian, S.J., McPherson, T.N., Brown, M.J., and Turin, H.J. (2000). "Development of a stormwater model for the Ballona Creek watershed." Los Alamos Unclassified Report (LA-UR-00-1849), Los Alamos National Laboratory, Presented at 1st Ballona Wetlands Symposium, May 17-18, 2000, Los Angeles, California.

Fang, T., and Ball, J. E. (2007). "Evaluation of spatially variable parameters in a complex system: An application of a genetic algorithm." J. Hydroinform., 9_3_, 163–173.

Freni, G., Mannina, G., Viviani, G., 2008. Uncertainty in urban stormwater quality modelling: the effect of acceptability threshold in the GLUE methodology. Water Research 42 (8–9), 2061–2072.

Freni, G., Mannina, G., Viviani, G., 2009. Uncertainty in urban stormwater quality modelling: The influence of likelihood measure formulation in the GLUE methodology. Science of the Total Environment 408 (2009), 138–145.

Freni, G., Mannina, G., Viviani, G., 2009b. Urban runoff modelling uncertainty: Comparison among Bayesian and pseudo-Bayesian methods. Environmental Modelling & Software 24 (2009),1100–1111.

Gelman, A. and D. B., Rubin (1992), Inference from Iterative Simulation Using Multiple Sequences. Statistical Science, 7, 457-472.

Kuczera, G., D. Kavetski, S. Franks, and M. Thyer (2006), Towards a Bayesian total error analysis of conceptual rainfall runoff models: Characterizing model error using storm dependent parameters, J. Hydrol., 331, 161–177, doi:10.1016/j.jhydrol.2006.05.010.

LACDPW, 2011. Ballona Creek Watershed, http://ladpw.org/wmd/watershed/bc/, retrieved on July, 26, 2011.

Mantovan, P., and E. Todini (2006), Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology, J. Hydrol., 330, 368– 381.

Muleta, M.K., 2011a. Improving Model Performance Using Season Based Evaluation. J. Hydrologic Engrg., ASCE, in press.

Muleta, M.K., 2011b. Sensitivity of Model Performance to the Efficiency Criteria Used as Objective Function during Automatic Calibrations. J. Hydrologic Engrg., ASCE, in review.

Muleta, M.K., Nicklow, J.W., 2005. Sensitivity and uncertainty analysis coupled with automatic calibration for a distributed watershed model. J. Hydrol., 306, 1-19.

Rossman, L.A. (2010). "Stormwater Management Model User's Manual, Version 5", U.S. EPA, Cincinnati, OH.

Schoups, G., and J. A. Vrugt (2010), A formal likelihood function for parameter and predictive inference of hydrologic models with correlated, heteroscedastic, and non‐Gaussian errors, Water Resour. Res., 46, W10531, doi:10.1029/2009WR008933.

Stedinger, J. R., R. M. Vogel, S. U. Lee, and R. Batchelder (2008), Appraisal of the generalized likelihood uncertainty estimation (GLUE) method, Water Resour. Res., 44, W00B06, doi:10.1029/2008WR006822.

Stenstrom, M. K., and Strecker, E. _1993_. Assessment of storm drain sources of contaminants to Santa Monica Bay: Annual pollutants loadings to Santa Monica Bay from stormwater runoff. Rep. No. UCLA-ENG-93-62, Vol. I, 1–248, Univ. of California, Los Angeles.

Tolson, B. A., Shoemaker, C. A.(2007). "Dynamically dimensioned search algorithm for computationally efficient watershed model calibration." *Water Resour. Res*., 43, W01413, doi:10.1029/2005WR004723.

Vrugt, J. A., C. G. H. Diks, H. V. Gupta, W. Bouten, and J. M. Verstraten (2005), Improved treatment of uncertainty in hydrologic modeling: Combining the strengths of global optimization and data assimilation, Water Resour. Res., 41, W01017, doi:10.1029/2004WR003059.

Vrugt, J. A., C. J. F. ter Braak, C. G. H. Diks, D. Higdon, B. A. Robinson, and J. M. Hyman (2009), Accelerating Markov chain Monte Carlo simulation by differential evo-lution with self-adaptive randomized subspace sampling, Int. J. Nonlinear Sci. Numer. Simul., 10(3), 273–290.

Vrugt, J. A., C. J. F. ter Braak, H. V. Gupta, and B. A. Robinson (2008), Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches to hydrologic modeling, Stoch. Env. Res. Risk A., 23, 1059–1060, doi:10.1007/s00477-008-0274-y.