On Parameter Estimation of Urban Storm-Water Runoff Model

Pedro Avellaneda[1]; Thomas P. Ballestero[2]; Robert M. Roseen[3]; and James J. Houle[4]

Abstract: An existing accumulation and wash-off model was applied and calibrated on a standard asphalt parking lot located in the northeastern United States. The field measured data consisted of rainfall, flow, and runoff samples taken from over 26 storm events monitored from 2004 to 2006. The contaminants under consideration include: total suspended solids, total petroleum hydrocarbons-diesel range hydrocarbons (TPH-D), dissolved inorganic nitrogen (DIN) (comprised of nitrate, nitrite, and ammonia), and zinc (Zn). The objective of the study was to provide probability distributions of model parameters for contaminants that have not been documented much (TPH-D, DIN, and Zn). The best fitting parameter values were found on a storm by storm basis. Subsequently, the range and variability of these parameters are provided for modeling purposes and other urban storm-water quality applications. A normal distribution was fitted to the optimized model parameter values to describe their distributions. A simulated annealing algorithm was used as the parameter optimization technique. Several examples are given to illustrate the methodology and the performance of the model. Finally, a Monte Carlo simulation was performed to assess the capability of the model to predict contaminant concentrations at the watershed's outlet.

DOI: 10.1061/(ASCE)EE.1943-7870.0000028

CE Database subject headings: Stormwater management; Water quality; Calibration; Uncertainty principles; Runoff; Urban areas; Parameters.

Introduction

Storm-water contaminant runoff models are commonly used for urban storm-water quality applications (DeCoursey 1985; Tsihrintzis and Hamid 1997; Zoppou 2001). These models are usually a combination of accumulation and wash-off equations. The accumulation of contaminants on impervious surfaces is nonlinear and follows an exponential increase as it approaches to a maximum value (Alley and Smith 1981). The total amount of contaminants is a function of the initial mass on the surface area and the length of the antecedent storm dry period. The common accumulation model is entirely deterministic, follows an exponential time history, and does not take into account the spatial distribution of the contaminants (Alley and Smith 1981).

Sartor et al. (1974) performed several field experiments on street surfaces when investigating a mathematical expression for simulating the wash off of contaminants. The study revealed that an exponential decay model was able to reproduce measured observations. This exponential decay model was a function of the available mass and rainfall intensity. Other studies have proposed the usage of the total volume of runoff as opposed to rainfall rate (Haiping and Yamada 1996). Computer models normally include a wash-off exponent in the erosion model (Rossman 2004). The wash-off exponent improves the performance of the model, in particular when the watershed seems to have a high nonlinear response to rainfall.

There are a variety of techniques available for model calibration purposes. The simplest method is the trial and error. Genetic algorithms, gradient-based functions, and simulated annealing (SA) are just a few of the more advanced calibration procedures that have been proposed to perform the optimization process (Hopgood 2001). All these methods stochastically explore the domain of the objective function by using a different goodness-of-fit criterion.

Alley and Smith (1981) provided an understanding of model sensitivity to the parameters for an urban runoff quality model. Water quality constituents in their study included total nitrogen, total lead, and suspended solids. To assess the parameter sensitivity, mathematical expressions were derived for each model parameter by direct differentiation of the analytical accumulation and wash-off equations. The wash-off coefficient of the erosion model was found to be the parameter with more variability. This study recommended the additional application of these models on runoff quality data for a better understanding of model assumptions and parameter variability.

Haiping and Yamada (1996) employed an adaptive step-size random search algorithm to calibrate an urban runoff model. The wash-off model used in their study was a linear function of rainfall intensity. The contaminants under consideration were total nitrogen, total solids, and total phosphorus. The model was applied continuously (long-term simulation) over a four month

[1]Graduate Research Assistant, Water Resources, Dept. of Civil Engineering, Univ. of New Hampshire, 230 Gregg Hall, 35 Colovos Rd., Durham, NH 03824 (corresponding author). E-mail: pmavellaneda@gmail.com

[2]Associate Professor, Dept. of Civil Engineering, The UNH Stormwater Center, Univ. of New Hampshire, 35 Colovos Rd., Durham, NH 03824. E-mail: tom.ballestero@unh.edu

[3]Assistant Research Professor, Dept. of Civil Engineering, The UNH Stormwater Center, Univ. of New Hampshire, 35 Colovos Rd., Durham, NH 03824. E-mail: robert.roseen@unh.edu

[4]Outreach Coordinator/Program Manager, The UNH Stormwater Center, Univ. of New Hampshire, 35 Colovos Rd., Durham, NH 03824. E-mail: james.houle@unh.edu

Note. This manuscript was submitted on April 28, 2008; approved on November 11, 2008; published online on February 20, 2009. Discussion period open until January 1, 2010; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Environmental Engineering***, Vol. 135, No. 8, August 1, 2009. ©ASCE, ISSN 0733-9372/2009/8-595–608/$25.00.

simulation period. Bounds of the calibration parameters were provided.

Gaume et al. (1998) focused their study on the estimation and interpretation of parameter uncertainties. A wash-off coefficient was used on the erosion model. This parameter was included to simulate the high nonlineal response to rainfall that is observed in practice. Difficulties during the calibration process were reported, mostly due to the large uncertainty in the parameter values. A lack of recorded data was thought to be the main reason for the calibration failure. This study attempted to calibrate parameter values for suspended solids on a storm by storm basis over the course of eight rainfall events.

For the study herein, storm-water runoff flow and water quality were monitored at a parking lot in Durham, New Hampshire. The objectives were (1) to calibrate an existing accumulation and wash-off model for a range of contaminants that have either not been studied or have little documentation, (2) to identify a probability distribution for each model parameter to describe its variability, and (3) to perform Monte Carlo simulations to estimate contaminant concentrations at the catchment's outlet using the calibrated model. This study does not intend to evaluate different storm-water models. Instead, an existing model was selected for calibration and the results are provided. This study focuses on the calibration of an existing storm water runoff quality model for a group of four contaminants [total suspended solids (TSS), total petroleum hydrocarbons-diesel range hydrocarbons (TPH-D), dissolved inorganic nitrogen (DIN), and Zn] using 26 storm events and a storm by storm approach. Several studies have identified these contaminants to be expected in storm water (Zoppou 2001; Tsihrintzis and Hamid 1997; Sartor et al. 1974). Based on a literature review, it appears that an evaluation of appropriate model parameters for TPH-D, DIN, and Zn has not been performed yet. The TSS continues to be the target contaminant for most storm-water modeling studies. For this study, TSS was selected for comparison purposes.

As a consequence of the awareness of the pollution impact due to storm water, managers need tools to evaluate and control storm water according to water quality criteria (Ahyerre et al. 1998; House et al. 1993). Application of storm water runoff quality models is necessary to improve watershed management, selection of storm-water technologies, parameterization for modeling purposes, understanding of model limitations, and assessment of how much confidence one could have in these models.

Field Site

The study area is a 36,000-m^2 commuter parking lot at the Univ. of New Hampshire in Durham, N.H. The parking lot is curbed, constructed of standard dense mix impervious asphalt, and drained by catch basins. Parking lot usage is a combination of passenger vehicles and routine bus traffic. A total of 786 parking spaces are used to near full capacity throughout the school year. During the summer, the parking lot receives much less use than during the regular school year. Additionally, during the summer the bus service is suspended.

Stormwater runoff flows from catch basins into a central 914-mm diameter reinforced concrete pipe. The runoff sampling station is located at the outlet of this pipe. The runoff time of concentration for the lot is 22 min, with slopes ranging from 1.5–2.5%. Contaminant concentrations are similar to typical values reported in stormwater (Pitt et al. 1995; Zoppou 2001; Minton 2002). The parking lot is subject to plowing, salting, and sanding during the winter. The climatology of the area is characterized as coastal, with an average annual rainfall of 1,220 mm uniformly distributed throughout the year. The average snowfall is approximately 1,600 mm.

Field Methods

A total of 26 discrete rainfall events were monitored between August 2004 and September 2006. Water sampling was performed at the outlet of the 914-mm pipe using a 6712SR ISCO automated sampler (Teledyne Isco, Inc., Lincoln, Neb.) provided with a stainless steel strainer, 9.52-mm vinyl collection tubing (Teledyne Isco, Inc., Lincoln, Neb.), and 24 discrete 1 L polypropylene bottles (Teledyne Isco, Inc., Lincoln, Neb.) and maintained at 4°C. Sampling events were based on storms exceeding 2.5 mm preceded by at least 72 h of dry weather (for most storms). Rainfall was monitored by using an ISCO Model 674 tipping bucket rain gage. The minimum depth that the rain gage could record was 0.254 mm. The rage gage was heated during the winter to guarantee its proper operation under severe weather conditions.

Flow was also monitored at the outlet of the 914-mm pipe, so a hydraulic model to estimate discharge was not necessary. The sampler was triggered on the basis of preset flow conditions. A total of 24 samples were taken for each storm event for a 24-h collection period; however, normally only 8–12 of these samples were sent to the laboratory to be analyzed. Samples were analyzed with the intent of linearizing the runoff concentration graph. The sampling program was designed to collect five samples within the first flush (4-min interval) and spread out the remaining samples over the rest of the hydrograph (24-min interval). ISCO 1 L ProPak disposable sampling bags were used to assure the storm-water samples were kept unaltered.

Runoff constituent analysis included TSS, TPH-D, DIN (comprised of nitrate, nitrite, and ammonia), and zinc (Zn). Selection of these contaminants was based on an initial characterization that included a wide range of petroleum hydrocarbons (gasoline range organics, lube oil, oil, grease), total and dissolved metals (cadmium, copper, iron, lead, mercury), and nutrients (DIN, phosphate, total phosphorus). For this parking lot, concentration levels of TSS, Zn, DIN, and TPH-D appeared to be significant. Storm-water samples were analyzed by a laboratory that is state certified for drinking water and wastewater. Guidance documents on collecting environmental data and the site quality assurance project plan were followed to assure the quality of the results (EPA 2006). Analyses were performed using techniques according to standard methods (APHA et al. 2005). Additional information on watershed characteristics and the sample monitoring program is found in Roseen et al. (2006).

A summary of the monitored storm events is shown in Table 1. This data set includes variations of storm rainfall duration, peak flow, runoff volume, antecedent dry period, and season. Temperatures below 0°C were recorded for the following storm events: 2/10/2005, 3/8/2005, and 12/16/2005. Those storm events occurred under severe weather conditions, with snow and ice covering a considerable percentage of the parking lot area. For the other winter events, weather conditions ranged from mild to moderate (mostly snow accumulation).

Model Structure

Typically, the pollutant accumulation and wash-off model can be described by the following equations (Shaheen 1975; Chen and Adams 2006):

Table 1. Summary of Monitored Storm Events

Rainfall event (m/d/y)	Peak intensity (mm/h)	Rainfall duration (min)	Total rainfall depth (mm)	Peak flow (m^3/day)	Runoff volume (m^3)	Anticedent dry period (days)	Number of samples	Season
9/18/2004	15	1,075	50	5,642	1,364	7.0	8	Fall
10/30/2004	21	705	11	8,678	281	13.0	8	Fall
11/24/2004	9	705	18	4,394	530	3.5	8	Fall
1/14/2005	24	645	17	21,101	1,033	1.3	8	Winter
2/10/2005[a]	6	1,520	32	4,437	795	3.6	10	Winter
3/8/2005[a]	3	1,220	20	2,338	406	5.7	18	Winter
3/28/2005	12	1,685	60	7,675	3,082	3.4	16	Winter
4/20/2005	12	480	15	4,274	1,017	5.9	8	Spring
5/21/2005	6	1,150	23	3,788	919	3.0	8	Spring
6/22/2005	15	95	8	9,120	266	4.0	8	Summer
8/13/2005	24	765	13	18,408	514	10.0	8	Summer
9/15/2005	18	30	5	11,947	444	10.0	9	Fall
9/26/2005	27	400	14	6,585	227	5.0	9	Fall
10/8/2005	6	120	5	9,066	2,137	8.0	8	Fall
11/6/2005	12	100	7	3,376	207	10.8	8	Fall
11/30/2005	9	810	18	3,017	277	5.0	9	Fall
12/16/2005[a]	18	630	35	3,086	470	5.5	8	Winter
1/11/2006	15	320	15	3,499	250	5.8	8	Winter
2/17/2006	12	110	3	2,398	83	2.5	7	Winter
3/13/2006	12	170	7	1,621	101	2.5	8	Winter
5/2/2006	12	1,920	60	4,642	1,251	7.0	9	Spring
5/9/2006	3	565	14	1,622	323	5.6	8	Spring
6/1/2006	125	485	51	12,162	201	10.7	8	Summer
6/21/2006	27	80	5	6,157	129	4.7	7	Summer
7/22/2006	40	50	5	8,333	79	7.5	8	Summer
9/6/2006	30	585	16	6,087	436	4.5	8	Fall

[a]Snow and ice accumulated on the parking lot's surface.

$$\frac{dM_a}{dt} = k_d - k_b M_a \tag{1}$$

$$M_w = M_a(1 - e^{-k_w R}) \tag{2}$$

where M_a=amount of pollutant on the surface; M_w=amount of pollutant removed from the surface during a storm; k_d=constant rate of pollutant deposition; k_b=pollutant removal rate due to wind and traffic; k_w=wash-off coefficient; t=antecedent dry period; and R=total runoff volume. Eq. (2) is called a "first order" model because the exponent of the total runoff volume R is 1. Other studies (Kanso et al. 2003; Chen and Adams 2006) analyzed the performance of a similar model that incorporates a wash-off coefficient as a new parameter. This approach is recommended when the watershed seems to have a high nonlinear response to rainfall. Recent computer models include a wash-off coefficient to improve the quality of the results (Rossman 2004).

Pollutant Accumulation Model

A pollutant buildup model is required to estimate the mass of contaminants on impervious surfaces between storm events. The accumulation of contaminants follows an exponential increase as it approaches to a maximum value (M_m), regardless of the length of the dry period (Alley and Smith 1981; Haiping and Yamada 1996). Integration of Eq. (1) results in the following equation:

$$M_a = M_m(1 - e^{-k_b t}) + M_0 e^{-k_b t} \tag{3}$$

where M_a=mass of pollutant on the parking lot surface (g/m^2); k_b=pollutant removal rate (day^{-1}); M_0=residual amount of pollutant after the previous runoff event (g/m^2); M_m=maximum amount of pollutant buildup (g/m^2); and t=antecedent dry period (days). These units will be used throughout the paper. For this study, M_0 was assumed to be zero. A discussion of this limitation is well described in Haiping and Yamada (1996). It was assumed that each runoff event had enough energy to remove the mass of contaminants accumulated on top of the impervious surface.

Pollutant Wash-Off Model

The pollutant wash-off model describes the removal of contaminants from the impervious surface during a runoff event. Most typical wash-off models are a function of either the discharge or runoff volume (Haiping and Yamada 1998; Alley and Smith 1981; Sartor et al. 1974; Millar 1999; Rossman 2004; Kanso et al. 2003). In this study, the washed-off mass was assumed to be proportional to the available mass and to the discharge. The pollutant wash-off model can be written as follows:

$$\frac{dM_w}{dt} = -k_w Q(t)^w M_a \tag{4}$$

where M_w=washed off mass (g/m^2) at time t; $Q(t)$=discharge (m^3/day); k_w=wash-off coefficient (day^{w-1}/m^{3w}); M_a=mass of

Table 2. Bounds of Calibration Parameters

Contaminant	$M_m(\mathrm{g/m^2})$		$k_b(\mathrm{day^{-1}})$		$k_w(\mathrm{day}^{w-1}/\mathrm{m}^{3w})$		$w(-)$	
	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
TSS	0	10	0	0.2	0	0.2	0	2
TPH-D	0	0.2	0	0.01	0	0.01	0	2
DIN	0	0.2	0	0.01	0	0.01	0	2
Zn	0	0.2	0	0.01	0	0.01	0	2

pollutant on the parking lot surface $(\mathrm{g/m^2})$; and w=wash-off exponent $(-)$. The wash-off exponent w allows the model to have a nonlinear dependency on the discharge, which could be convenient when the pollutograph is nonlinear. The pollutant concentration is calculated by dividing the eroded mass within a time interval Δt by the runoff volume of that same interval. Other studies suggest the combination of the accumulation and wash-off models to obtain a single mathematical expression (Chen and Adams 2006; Haiping and Yamada 1996) that can be used as a tool for continuous simulation. Eqs. (3) and (4) have to be multiplied by the watershed area if the total accumulated mass and total washed off mass need to be calculated.

Objective Function

The aim of this study was to estimate the parameters that best fit the accumulation and wash-off model for four separate contaminants. The best fitting parameters were found by minimizing the sum of squares of residuals

$$O(\Theta) = \min \sum_{i=1}^{m} [C_{\mathrm{obs}}^{i} - C_{\mathrm{est}}^{i}(\Theta)]^2 \tag{5}$$

where O=objective function; Θ=best fitting parameter values; C_{obs}=measured concentrations; $\mathrm{C}_{\mathrm{est}}$=estimated concentrations when using Θ; and m=number of samples analyzed during the storm event.

Optimization Technique

The accumulation and wash-off model was calibrated using field data and an optimization technique implemented to determine the most adequate fitting parameters for the model. Alley and Smith (1981) used Rosenbrock's method to find the best fit values for storm-water applications. A discussion on how parameter interaction affects the optimization technique was also provided. Gaume et al. (1998) investigated the uncertainty of the calibrated parameter values on an urban runoff model similar to Eqs. (3) and (4). In their Paper, the Powell method was used as the optimization technique and a description of the shape of the objective function was provided for several cases. One major difficulty was dealing with narrow valleys in the objective function, which would add more complexity to the search process. The lack of knowledge of the real pollutograph was a contributing factor to the failure of some optimization trials.

Kirkpatrick et al. (1983) proposed SA, a stochastic global optimization technique that finds the global minimum or maximum of a given mathematical function. The SA technique is meant to be used on highly nonlinear multivariable problems. It was initially tested on physical applications but nowadays is extensively used in other scientific fields such as hydrology (Pardo-Iguzquiza 1998; Sumner et al. 1997). Based on a review of literature, it appears that this technique has not yet been used on storm-water runoff. The searching algorithm can be visualized as a bouncing sphere that can travel over the peaks and valleys of a given surface (the objective function). Throughout the iterative process the "energy" at which the sphere bounces decreases as it gets closer to the optimal value. In SA, a new trial solution (new set of parameter values) is accepted when there is a reduction in the current objective function value (Kirkpatrick et al. 1983). First, the probability of accepting a new trial solution is calculated as follows:

$$p(\Delta O) = e^{-\Delta O/T} \tag{6}$$

where ΔO=change in the objective function; T=parameter called temperature (which represents the energy of the bouncing sphere); and $p(\Delta O)$=probability of accepting the proposed set of parameter values. At higher "temperatures" the algorithm extensively explores the parameter space, so the global minimum is likely to be found. When the temperature is high the probability of accepting the proposed set of parameters is high as well. As the temperature drops, the probability of accepting new candidates reduces and the search focuses on previous local optimal parameter values.

The value of T is initially set high and is periodically reduced according to a "cooling schedule" (Hopgood 2001) with a maximum number of steps m. A reduction coefficient $0<a<1$ is used to slowly decrease the value of T. A commonly used simple cooling schedule is

$$T_{t+1} = aT_t \quad t = 0,1,\ldots,m \tag{7}$$

The SA algorithm overcomes the problem of being trapped in a local minimum by accepting a trial solution that, although is not the best, may lead to the true optimal values. The principal advantage of the SA algorithm, when compared to other searching techniques, is that the temperature T keeps the algorithm from getting trapped by permitting uphill (or downhill) moves. The algorithm provides an intriguing instance of "artificial intelligence" in which the computer has arrived almost uninstructed at a solution that might have been thought to require the intervention of human intelligence (Kirkpatrick et al. 1983). Large m values allow the control variable to decrease slowly and then perform the search on a broad area. In this study, the following annealing schedule was applied for most instances: T=20,000, a=0.9, and m=100. The annealing schedule determines the degree of uphill (or downhill) movement permitted during the search, so it is critical to the algorithm's performance. These parameters are problem specific and depend on the scaling of the change in the objective function ΔO. For this study, a trial and error approach was adopted to identify the parameters of the SA algorithm. Different annealing schedules were tested until the same optimal accumulation and wash-off parameters were achieved. However, the proposed annealing scheduled was changed for a few storms in

Table 3. Summary of Samples with Concentrations Reported as BDL

Format	Total number of samples	TPH-D	Zn	DIN	TSS
Discrete	234	69	34	68	38
Percentage	100%	29%	15%	29%	16%

which the algorithm did not seem to have intensively explored the parameter's space (or the temperature did not seem to be high enough). An acceptance probability of 0.7 was kept during the initial stage of the search.

The total number of iterations was fixed (5×10^4) and the optimization process was stopped when the new set of parameter values did not change over 100 iterations. If these parameter values were still changing by the end of the total number of iterations, then the model was rerun again with a higher number of iterations (1×10^5).

Bounds were necessary for the practical implementation of the SA method. The upper and lower bounds for all the parameters are given in Table 2. These values were determined after performing some trial runs with the data and recommendations from the literature (Haiping and Yamada 1996; Chen and Adams 2006; Gaume et al. 1998). The SA results were compared to the values shown in Table 2 to verify that the results were not close to the upper bounds.

Concentration Values below Detection Limit

A concentration value below the analytical detection limit (BDL) is reported when pollutant concentrations are below the analytical reporting limit of the laboratory. The application of the mathematical model requires finite concentration values. The researcher is left with few options when this situation arises: discard valuable information provided by the BDL samples; select an estimate of the concentration (1/2 the detection limit is common); use a mixed probability model (fixed probability density at the detection limit); or generate estimates of the BDL concentration by using a probabilistic approach. An accepted method for generating estimates is to fit a probability distribution based on the data that is above the reporting limit (Helsel and Hirsch 2002). First, a distribution shape is assumed and then the distribution parameters are computed by using conventional methods such as maximum likelihood estimator (MLE) or probability plot procedures. Finally, discrete BDL concentrations (values between 0 and the detection limit) can be randomly generated using the fitted distribution. Table 3 summarizes the number of samples reported as BDL for each contaminant. In this study, the detection limits were typically 0.50 mg/L for TPH-D, 0.03 mg/L for Zn, 0.5 mg/L for DIN, and 10 mg/L for TSS. However, the reported detection limit could vary depending on the actual sample volume and concentrations.

A gamma distribution was employed for this study since it adequately described the collected data (other distributions such as Gumbel and lognormal were tested). The parameters, γ (shape) and θ (scale), of the distribution were determined by using the MLE method and are shown in Table 4. The probability density function for each contaminant is shown in Fig. 1 as well as the respective cumulative probability function (CDFs). The

Table 4. Parameters of the Gamma Distribution Fitted to the Data

Contaminant	Total samples above DL (number)	$\gamma(-)$	θ(mg/l)	$D_{\text{critical}}(-)$	$D_{n,\alpha}(-)$
TPH-D	165	2.82	0.39	0.11	0.11
Zn	200	1.83	0.04	0.10	0.10
DIN	166	2.08	0.31	0.11	0.10
TSS	196	0.79	82.80	0.10	0.14

Note: Level of significance $\alpha=0.05$.

Kolmogorov-Smirnov test (KS test) was used to measure the goodness-of-fit of the fitted probability distribution. A good fit was obtained for TPH-D, Zn, and DIN [Figs. 1(b, d, and f)] since $D_{n,\alpha} \leq D_{\text{critical}}$. The $D_{n,\alpha}$ statistic was computed for a level of significance $\alpha=0.05$ (95% confidence interval) and n concentration values above the detection limit. The D statistic represents the maximum vertical deviation between the empirical and theoretical data probability distributions. Although for TSS the fitted distribution did not pass the KS test at 0.05 level of significance, the fitted probability distribution was used for practical purposes.

This statistical method is meant to preserve the entire distribution of the data both below and above the detection limit. The more data points above the reporting limit, the better the quality of the model from a statistical point of view. For TPH-D and DIN the shape parameter (γ) is likely to be much smaller due to the fact that nearly 30% of the concentrations were reported below detection limit. Note that some of the detection limit issues would have been avoided if a more precise laboratory procedure had been performed.

Calibration Procedure

A broad range of storms (Table 1) were independently used to obtain the optimal values of model parameters for each contaminant. The calibration was performed under the assumption of null residual mass after each rain event ($M_0=0$). This assumption is not true when the storm event is flow limited, in other words, when the storm is not intense enough to wash off the mass of contaminants on the surface. The null residual mass assumption was verified by comparing the total available mass and the mass that was actually washed off (both values are given by the model). For this study, some storm events were identified as flow-limited events. To identify those events, it was necessary to express the total washed off mass (M_w) as a percentage of the total available mass (M_a). A common definition of a flow-limited event dictates that less than 80% of the total pollutant load can be transported by the first 20% of the flow volume (Han et al. 2006; Sansalone and Cristina 2004). In other words, this definition implies that a "first flush" has occurred if more than 80% of the mass has been transported by the first 20% of the flow volume. The full definition of a flow-limited event seemed to be too conservative for the purposes of this research. Therefore, for this study, a flow-limited event was defined as the event that had transported less than 80% of the total available mass at the end of the storm's duration. Table 5 shows the total estimated available mass and the total estimated washed-off mass for all the storms and contaminants. The ratio r_m between the available mass and washed-off mass is reported. For TSS, four events were identified as flow-limited; four for TPH-D; six for Zn; and four for DIN.

Initial parameter values were randomly generated using the bounds shown in Table 2. The root-mean-square error (RMSE)

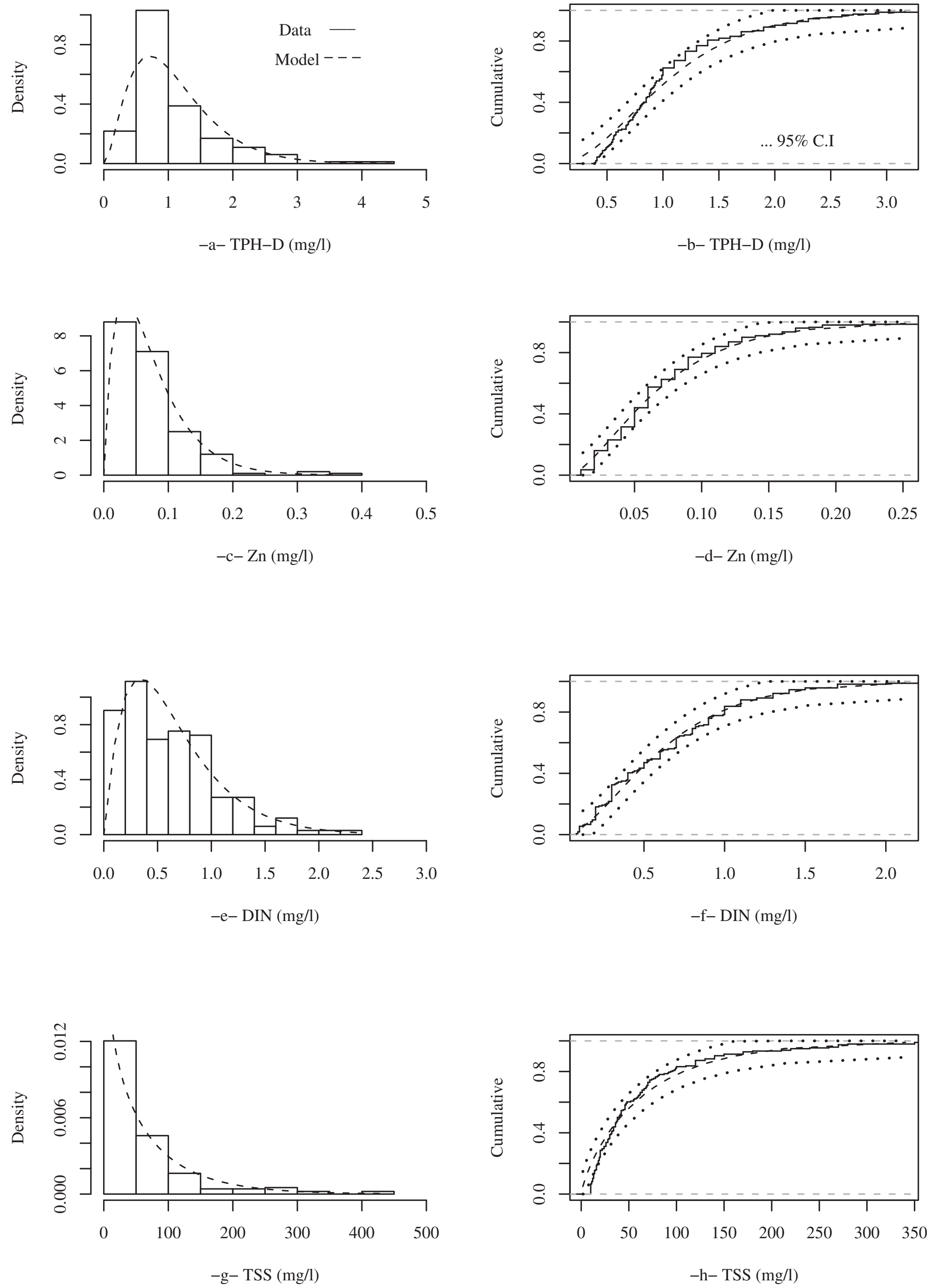

Fig. 1. Empirical and theoretical data probability distributions

was used as the quality criterion. When changes in the objective function remained significant the model was rerun and the total number of iterations increased. On average, convergence was achieved after 2×10^4 iterations. For calibration, concentration values were used as they were reported from the lab and no further statistical analysis was applied to discard anomalous data. The calibration period included storms from 09/18/2004 to 6/01/2006. The last three storm events described in Table 1 were left for validation.

Results

Water Quality Estimation

Figs. 2–4 illustrate model results for some of the storms considered in this study. Each figure shows the hydrograph (flow), hyetograph (rainfall), observed concentrations, estimated concentrations, and the respective relative cumulative mass (CDF). The CDF for each contaminant is actually a mass-based CDF and was

Table 5. Total Estimated Available and Washed-Off Mass for the Storms Used during the Calibration Stage

Rainfall event (m/d/y)	TSS			TPH-D			Zn			DIN		
	Available mass (g)	Washed-off mass (g)	r_m	Available mass (g)	Washed-off mass (g)	r_m	Available mass (g)	Washed-off mass (g)	r_m	Available mass (g)	Washed-off mass (g)	r_m
9/18/2004	27,109	20,823	0.77	316	316	1.00	27	27	0.99	132	132	1.00
10/30/2004	4,115	3,983	0.97	263	187	0.71	31	19	0.62	83	76	0.92
11/24/2004	685	680	0.99	26	26	1.00	1	1	1.00	6	6	1.00
1/14/2005	20,775	20,775	1.00	NA	NA	NA	22	22	0.97	7	7	1.00
2/10/2005	24,140	21,898	0.91	241	240	1.00	22	22	0.97	248	247	1.00
3/8/2005	17,132	15,919	0.93	350	316	0.90	37	33	0.89	162	155	0.96
3/28/2005	25,151	25,031	1.00	227	227	1.00	78	50	0.65	225	225	1.00
4/20/2005	27,615	27,607	1.00	390	360	0.92	63	62	0.99	403	399	0.99
5/21/2005	6,530	5,017	0.77	91	91	1.00	7	7	0.99	46	46	1.00
6/22/2005	9,724	9,724	1.00	130	99	0.76	7	4	0.67	68	44	0.65
8/13/2005	28,611	11,102	0.39	209	181	0.87	19	15	0.81	228	144	0.63
9/15/2005	1,736	1,734	1.00	124	124	1.00	5	5	1.00	48	48	0.99
9/26/2005	1,375	1,375	1.00	151	136	0.91	5	5	0.96	31	31	1.00
10/8/2005	2,529	2,524	1.00	548	526	0.96	16	16	1.00	42	42	1.00
11/6/2005	1,943	1,647	0.85	238	238	1.00	5	5	0.92	63	63	0.99
11/30/2005	4,487	4,461	0.99	98	87	0.89	5	5	1.00	16	16	1.00
12/16/2005	9,783	9,760	1.00	30	30	1.00	31	24	0.77	140	130	0.93
1/11/2006	8,271	8,177	0.99	369	312	0.85	6	6	1.00	40	40	1.00
2/17/2006	14,537	9,165	0.63	170	122	0.72	46	9	0.20	56	42	0.75
3/13/2006	5,817	4,801	0.83	151	99	0.66	4	4	0.88	16	16	1.00
5/2/2006	25,023	24,300	0.97	161	161	1.00	4	4	1.00	19	19	1.00
5/9/2006	13,494	13,134	0.97	146	143	0.98	7	7	1.00	48	48	1.00
6/1/2006	18,207	17,528	0.96	372	372	1.00	16	12	0.73	35	35	1.00

Note: A ratio $r_m < 0.80$ indicates a flow-limited storm event. NA=not available.

obtained by multiplying the estimated concentrations by the synoptically measured flows. For example, Fig. 2(a) shows flow rate changes (below) and rainfall pattern (top). The CDFs for each contaminant are shown in Fig. 2(b). Figs. 2(c–f) show observed and estimated concentrations for DIN, TPH-D, TSS, and Zn (RMSEs are reported). Note that on some figures, the rainfall time series is not completely plotted to allow for a better display of the observed concentrations.

The 03/08/2005 event (Fig. 2) had an almost constant rainfall intensity, and as a result, the hydrograph increased and decreased gradually. The model with calibrated parameters does a very good job for reproducing most of the observed data. The observed concentrations and pollutograph for TPH-D stand out from the other contaminants due to the poorer fit and the scatter of the data. The last seven samples for DIN were reported BDL (<0.5 mg/L). This explains the scatter of the observed values on the lower part of the graph since those values were randomly generated. Nevertheless, the smoothness of the estimated pollutograph does not seem to be affected. For this storm, BDL values were not recorded for TSS, Zn, and TPH-D. The CDFs look very similar for all contaminants. The parameter w was close to 1 because the observed concentrations decreased smoothly. Optimized parameter values for the 03/08/2005 storm event are presented in Table 6.

A type of first flush event is demonstrated in Fig. 3 for the 04/20/2005 storm. The highest intensity occurred early in the storm and a constant rainfall intensity occurred during the rest of the storm duration. The first flush effect is more pronounced for the total suspended solid CDF. Values for the BDL concentration were reported for DIN, TSS, and Zn for the last four, three, and two samples, respectively. Only one BDL value was reported for TPH-D at 356 min. The estimated pollutographs match the higher concentrations at the beginning of the storm. For TPH-D, observed concentrations increase at the very end of the event, however, the model is not capable of estimating this phenomenon. The parameter w ranged from 1.0 and 1.2 for the different contaminants due to the smooth decrease in observed concentrations.

The 01/11/2006 storm was a late-peaking event (Fig. 4). At least two peak flows were observed, one approximately at 160 min and other at 310 min, in direct response to rainfall intensity peaks. The BDL values were reported for Zn and DIN for samples taken at 251 and 341 min. Even though the peak flow occurred

Table 6. Optimized Parameter Values for Selected Storm Events

Rainfall event m/d/y	Parameter	Contaminant			
		DIN	TPH-D	TSS	Zn
3/8/2005	M_m	0.1513	0.1858	0.7119	0.0845
	k_b	0.0052	0.0093	0.1896	0.0021
	k_w	0.0046	0.0005	0.0079	0.0054
	w	1.0684	1.3504	0.9711	1.0013
4/20/2005	M_m	0.1944	0.1993	1.7269	0.1238
	k_b	0.0099	0.0094	0.0980	0.0024
	k_w	0.0026	0.0009	0.0038	0.0050
	W	1.1322	1.1925	1.1539	1.0341
1/11/2006	M_m	0.0207	0.1866	0.3383	0.0681
	k_b	0.0095	0.0096	0.1918	0.0004
	k_w	0.0092	0.0000	0.0057	0.0006
	w	1.2357	1.8571	1.1560	1.5140

Note: Units: $M_m(\mathrm{g/m^2})$; $k_b(\mathrm{day^{-1}})$; $k_w(\mathrm{day^{w-1}/m^{3w}})$; and $w(-)$.

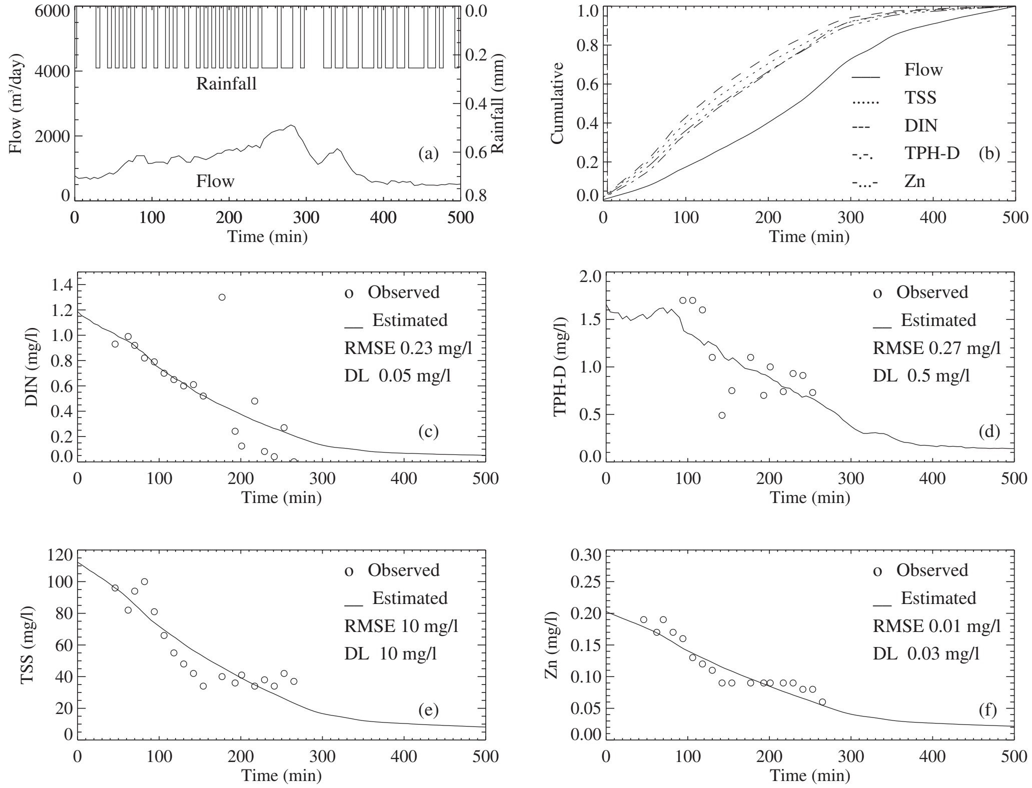

Fig. 2. Results for the 03/08/2005 storm. Optimized parameter values are presented in Table 6.

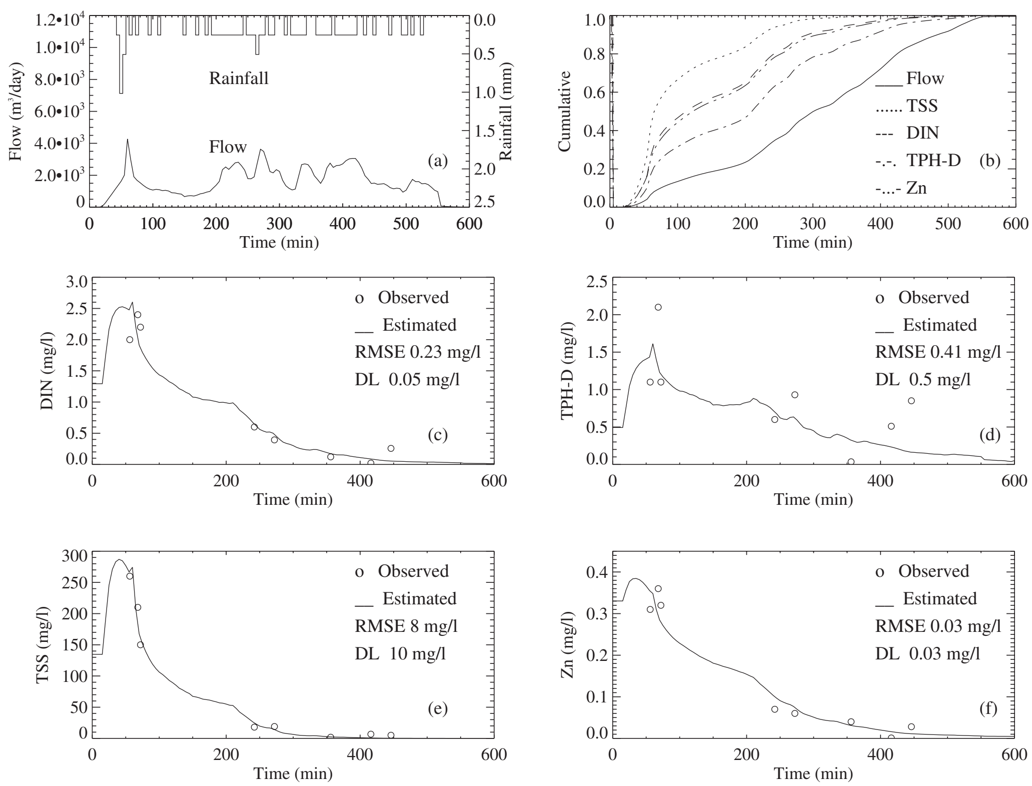

Fig. 3. Results for the 04/20/2005 storm. Optimized parameter values are presented in Table 6.

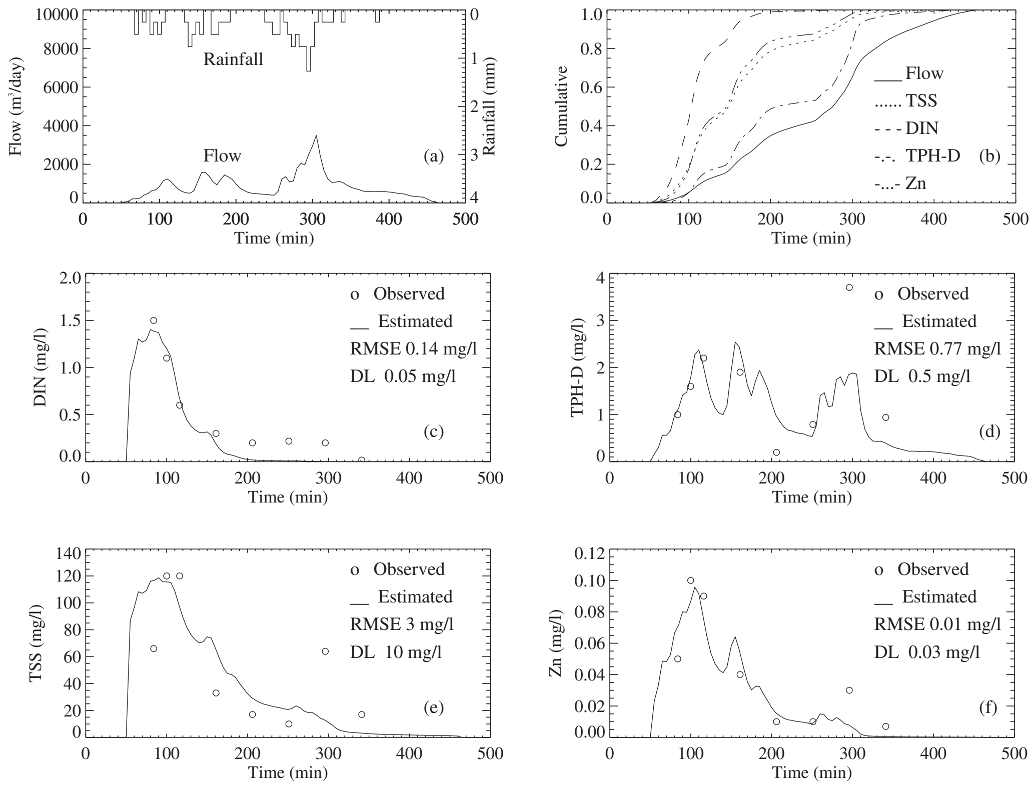

Fig. 4. Results for the 01/11/2006 storm. Optimized parameter values are presented in Table 6.

toward the end of the event, the highest concentration values were measured at the beginning of the storm. The parameter w ranged from 1.2 to 1.9 for the different contaminants, which agrees with the nonlinear trend observed in the data.

Simulated Annealing Algorithm

Fig. 5 shows the changes in the objective function [Eq. (5)] when the parameters were calibrated for Zn on the 03/13/2006 event. During the first iterations, high values for the objective function were obtained due to the fact that the SA algorithm was randomly exploring the parameters space. After 1×10^4 iterations the objective function seemed to have reached its minimum value. The optimization process was stopped at 1.2×10^4 iterations. Note that a sensitivity analysis was not performed for this study. For all the storms used during the calibration stage, the same optimal parameter values were obtained when rerunning the model (several times) with a different cooling schedule. A comparison of the success and failure of the SA algorithm and other optimization techniques is beyond the scope of this paper.

Statistical Analysis

Fig. 6 shows the CDFs obtained for the different model parameters (optimized values) and pollutants. A normal distribution was fitted to these CDFs and a 95% confidence interval ($\alpha = 0.05$ level of significance) was computed to measure the goodness-of-fit of the distribution. For this purpose, the KS test was used. The mean $\bar{x}$ and the standard deviation (S) of the fitted normal distributions are shown in Table 7. The distribution of the parameters k_w (for DIN) and k_w (for TSS) did not pass the KS test. Gaume et al. (1998) found fitting parameter values after calibrating a similar runoff water quality model for suspended solids. In their study,

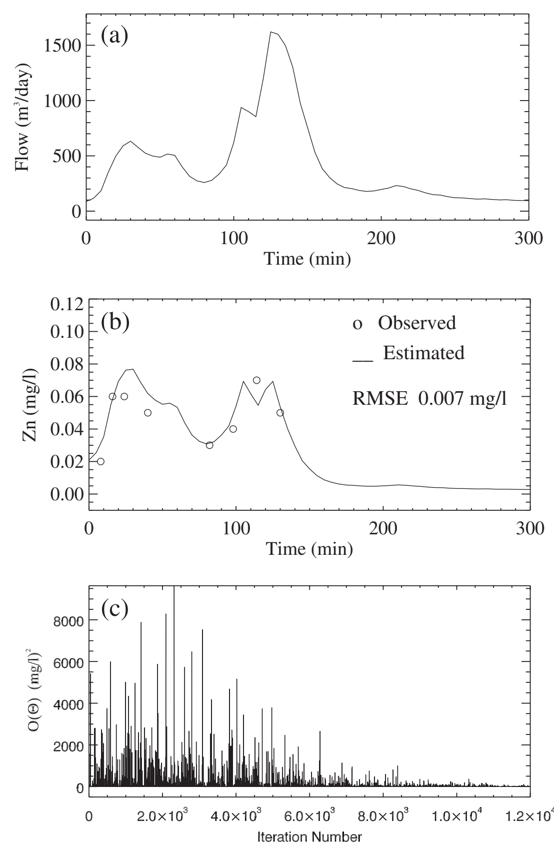

Fig. 5. Optimization process and calibration results for the 03/13/2006 storm. Optimized parameter values are shown for zinc. Optimized parameters: $M_m = 0.13$ g/m^2; $k_b = 0.0004$ day^{-1}; $k_w = 0.0002$ day^{w-1}/m^{3w}; $w = 1.71$; and $b = 2.50$ days (antecedent dry period).

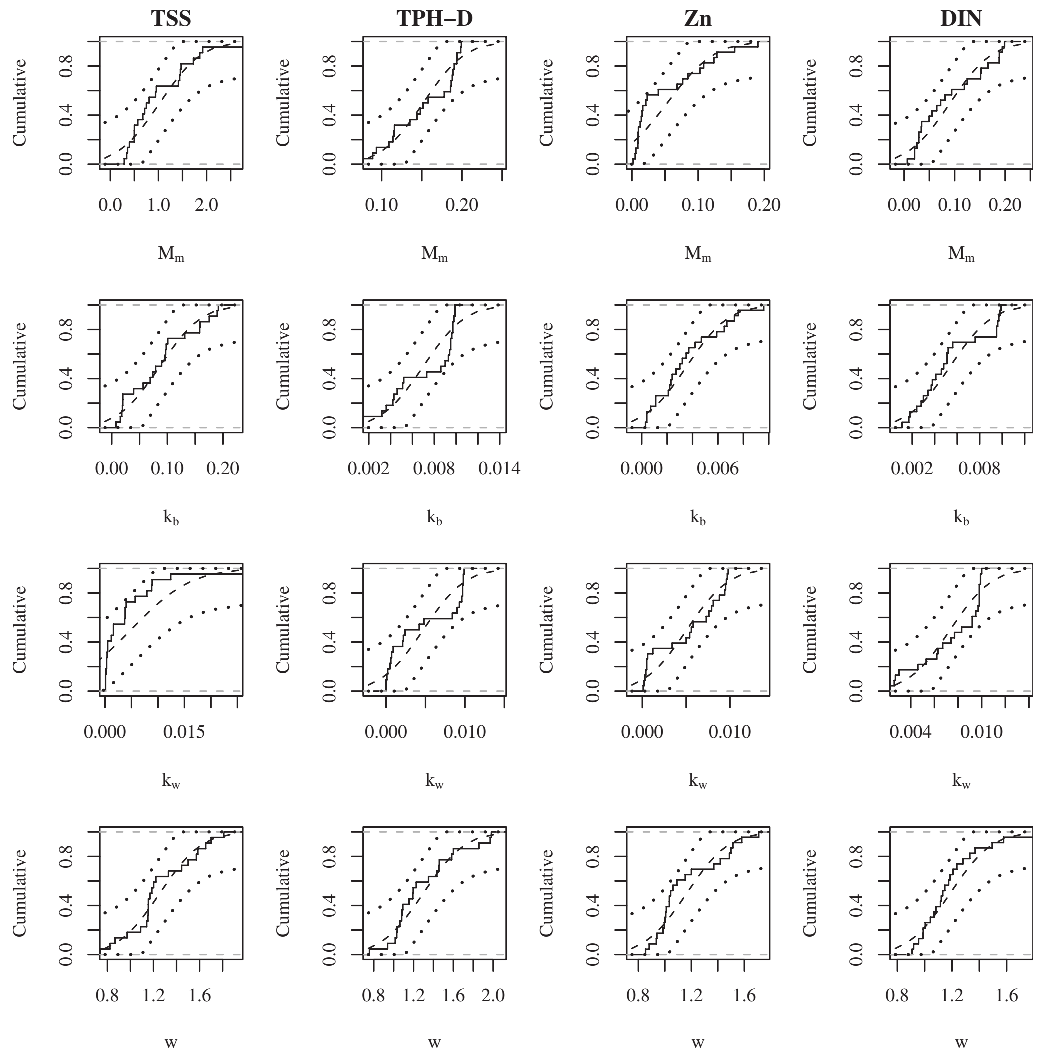

Fig. 6. CDFs of optimized model parameter values (solid line) and fitted normal CDFs (dashed line). A 95% confidence interval is reported (dotted line). Units: $M_m(\mathrm{g/m^2})$; $k_b(\mathrm{day^{-1}})$; $k_w(\mathrm{day^{w-1}/m^{3w}})$; and $w(-)$.

the parameter values ranged as follows: 40–120 for M_m, 0.2–0.8 for k_b, 0.005–0.030 for k_w, and 1.0–2.0 for w. Note that their model was calibrated for an urban storm-water sewer system, so this could explain why some parameter values are different from those provided in this paper. A mean M_m value of 1.029 $\mathrm{g/m^2}$ was obtained for TSS. This mean value follows within the range of variation ($0.018 \sim 25$ $\mathrm{g/m^2}$) found by Chen and Adams (2006) for a similar storm-water modeling study. The watershed used in their study was composed of residential and commercial land use. Chen and Adams (2006) also provided calibrated parameter values for total phosphorus, copper, and total Kjeldahl nitrogen).

A correlation analysis was performed to study the degree of linear dependencies among the model parameters. Table 8 shows the correlation coefficients for the different parameter combinations and pollutants. The mathematical structure of the accumulation and wash-off model can explain some values of the correlation matrix. For example, it was found that there was a negative correlation between the wash-off coefficient k_w and the pollutant buildup M_m parameter; however, note that these parameters are inversely proportional [Eqs. (3) and (4)]. A negative correlation (for all contaminants) between k_w and w can be explained by the fact that these parameters are affecting the variable flow in the same equation. It also appears that a positive amount of pollutant buildup M_m correlates to a positive removal rate k_b during pollutant accumulation. One could interpret that a quantity proportional to the pollutant buildup can be removed and transported by wind or traffic during dry periods. However, this positive correlation was only found for TSS, TPH-D, and DIN. These results are consistent with other studies (Gaume et al. 1998; Kanso et al. 2003; Kanso et al. 2005), where a similar correlation structure was reported for TSS.

Monte Carlo Simulation

A Monte Carlo simulation was performed to evaluate the ability of the model to predict pollutant concentrations at the catchment's outlet. Three storms were selected for validation: 06/21/2006, 07/22/2006, and 9/6/2006. Note that the results for the 9/6/2006 storm are not shown due to space constraints. The fitted normal distributions to the optimized parameter values (Table 7 and Fig. 6) were used to randomly generate M_m, k_b, k_w, and w. The parameters were assumed to be independent and noncorrelated for modeling purposes. Note that a multivariate test of independence was not performed for this study. Further research should explore the use of statistical tools such as the generalized likelihood uncertainty estimator (GLUE) when analyzing the uncertainties of the

Table 7. Parameters of the Normal Distribution Fitted to the Optimized Model Parameter Values

Parameter	Statistic	TSS	TPH-D	Zn	DIN
$M_m(\mathrm{g/m^2})$	$\bar{x}$	1.029	0.152	0.053	0.091
	S	0.693	0.042	0.056	0.066
	n	22	22	23	23
	D_{critical}	0.281	0.281	0.275	0.275
	$D_{n,\alpha}$	0.178	0.202	0.270	0.153
$k_b(\mathrm{day^{-1}})$	$\bar{x}$	0.086	0.007	0.004	0.005
	S	0.060	0.003	0.003	0.003
	n	22	22	23	23
	D_{critical}	0.281	0.281	0.275	0.275
	$D_{n,\alpha}$	0.139	0.246	0.129	0.179
$k_w(\mathrm{day}^{w-1}/\mathrm{m}^{3w})$	$\bar{x}$	0.005	0.005	0.005	0.007
	s	0.009	0.004	0.004	0.003
	n	22	22	23	23
	D_{critical}	0.281	0.281	0.275	0.275
	$D_{n,\alpha}$	0.297	0.217	0.194	0.428
$w(-)$	$\bar{x}$	1.255	1.298	1.166	1.187
	s	0.290	0.335	0.252	0.244
	n	22	22	23	23
	D_{critical}	0.281	0.281	0.275	0.275
	$D_{n,\alpha}$	0.184	0.175	0.226	0.163

Note: The media ($\bar{x}$) and the standard deviation (S) follow the units of each model parameter. Other statistical parameters are dimensionless.

wash-off model. The GLUE procedure (Beven and Binley 1992) implicitly incorporates the correlation structure of the model since it evaluates a set of fitting parameters rather than individual values.

Simulation results are shown in Figs. 7 and 8. The central solid line indicates the expected concentrations and will be used in this study as a measure of the model behavior. The dark shaded region indicates the 30 and 70% uncertainty limits. The light gray shaded region indicates the 10 and 90% uncertainty limits. For this study, the estimates of the 10th, 30th, 70th, and 90th percentiles of the predictions were used to define the uncertainty limits. Observed concentration values and the detection limit are also plotted. The number of simulations was set to 5,000 for each contaminant. Fig. 7 presents results for the 06/21/2006 storm. Results for this storm suggest an acceptable performance of the model since most of the observed concentration values fell within the uncertainty limits. The maximum observed concentration for TPH-D, TSS, and Zn was obtained for the first sample; however, the model was not able to adequately estimate these concentrations for TPH-D and TSS. DIN observed concentrations fell within the 30% and 70% uncertainty limits.

Table 8. Correlation Coefficients for Different Model Parameter Combination

Contaminant	Parameter	M_m	k_b	k_w	w
TSS	M_m	1			
	k_b	0.38	1		
	k_w	0.03	−0.18	1	
	W	−0.12	−0.08	−0.56	1
TPH-D	M_m	1			
	k_b	0.19	1		
	k_w	−0.27	−0.53	1	
	w	0.32	0.02	−0.55	1
Zn	M_m	1			
	k_b	−0.39	1		
	k_w	−0.51	0.16	1	
	w	0.31	−0.51	−0.66	1
DIN	M_m	1			
	k_b	0.10	1		
	k_w	−0.31	−0.21	1	
	w	−0.59	−0.12	−0.38	1

Note: Units: $M_m(\mathrm{g/m^2})$; $k_b(\mathrm{day^{-1}})$; $k_w(\mathrm{day}^{w-1}/\mathrm{m}^{3w})$; and $w(-)$.

Table 9. Observed and Estimated Total Mass for the Storms Used for Validation

		Total Mass (g)			
Storm	Contaminant	Observed	Estimated (q_{30})	Estimated (q_{50})	Estimated (q_{70})
6/21/2006	TSS	5,259	2,548	5,354	9,378
	TPH-D	39	32	60	100
	Zn	4.0	3.3	6.4	10.6
	DIN	34	23	44	71
7/22/2006	TSS	4,887	3,543	6,858	11,249
	TPH-D	97	40	71	112
	Zn	4.5	3.7	6.7	10.3
	DIN	41	33	59	90

Monte Carlo simulations for the 07/22/2006 event are displayed in Fig. 8. For DIN, the first three observed concentrations fell outside the plotted uncertainty limits [Fig. 8(a)]. This indicates poor performance of the model. However, the remaining observed concentrations are much closer to the expected concentrations. Results for TPH-D fell within the 30 and 70% uncertainty limits and the expected concentration line predicts very well some observed concentrations [Fig. 8(b)]. Results for the TSS tend to overestimate pollutant concentrations since there are six observed values that fell below the 30% uncertainty limit [Fig. 8(c)]. For Zn, four samples fell within the 30 and 70 % uncertainty limits and the remaining values below the 30% uncertainty limit [Fig. 8(d)].

The total (observed and estimated) washed off mass was computed for the 06/21/2006 and 07/22/2006 storm events. The total mass was computed as the sum of the products between concentration and runoff volume over the number of samples collected during the storm. Table 9 shows results for each contaminant and storm event. The three estimated values correspond to the 30%(q_{30}), 50%(q_{50}), and 70%(q_{70}) uncertainty limits. For each validation exercise, the observed mass fell within the 30% and 70% uncertainty limits obtained from the model (as percentiles of the predicted values). However, the model seemed to overestimate the total wash-off mass since almost for all cases the median estimated values were higher than those observed. Note that the wash-off model is empirical and perhaps cannot represent precisely the transport processes for every single storm. Other studies (Haiping and Yamada 1998) have suggested that the wash-off model does not estimate accurately pollutant concentrations during the falling limb of a pollutograph. Figs. 7(a) and 8(a, c, and d) indicate that the model estimates poorly pollutant concentrations during the falling limb, which will affect the total estimated washed off mass. Soonthornnonda et al. (2008) recommended the wash-off model be applied to the near constant portion of the hydrographs at peak flows. The study reported a significant improvement of fits by omitting points in the early and late portions of the hydrographs. Additionally, it may be possible that some of

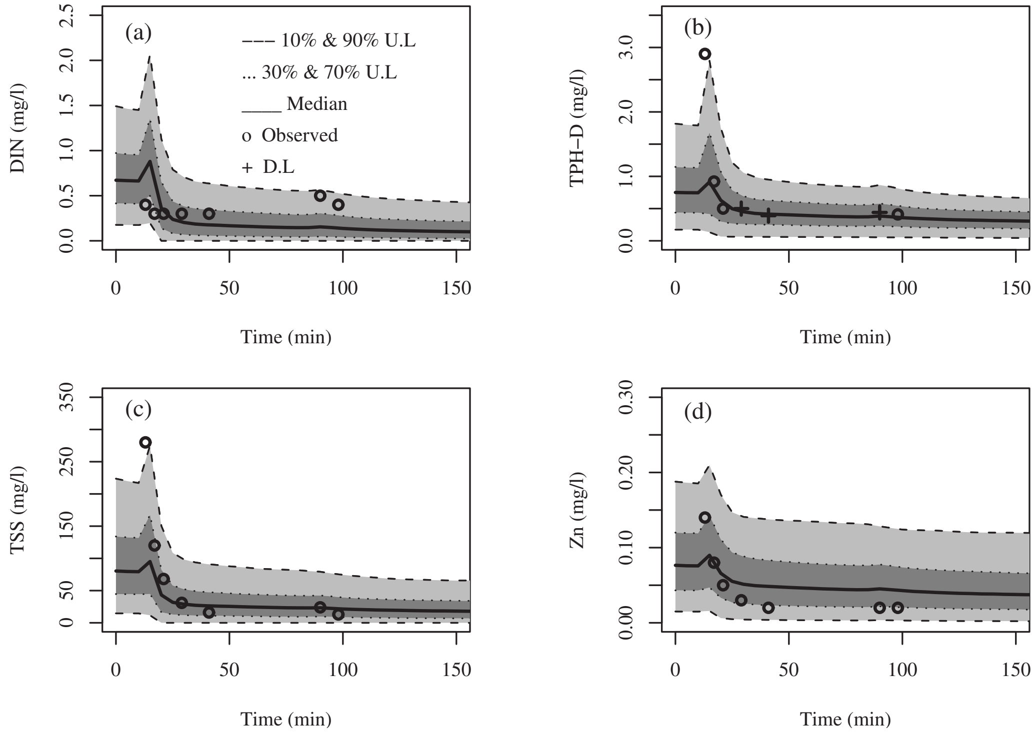

Fig. 7. Monte Carlo simulations of predicted pollutant concentrations at the catchment's outlet for the 06/21/2006 storm event. U.L. =uncertainty limits (as percentiles of the predicted values). The dark shaded region indicates the 30 and 70% U.L. The light gray shaded region indicates the 10 and 90% U.L. Observed concentrations values and the reported detection limits were plotted as discrete points.

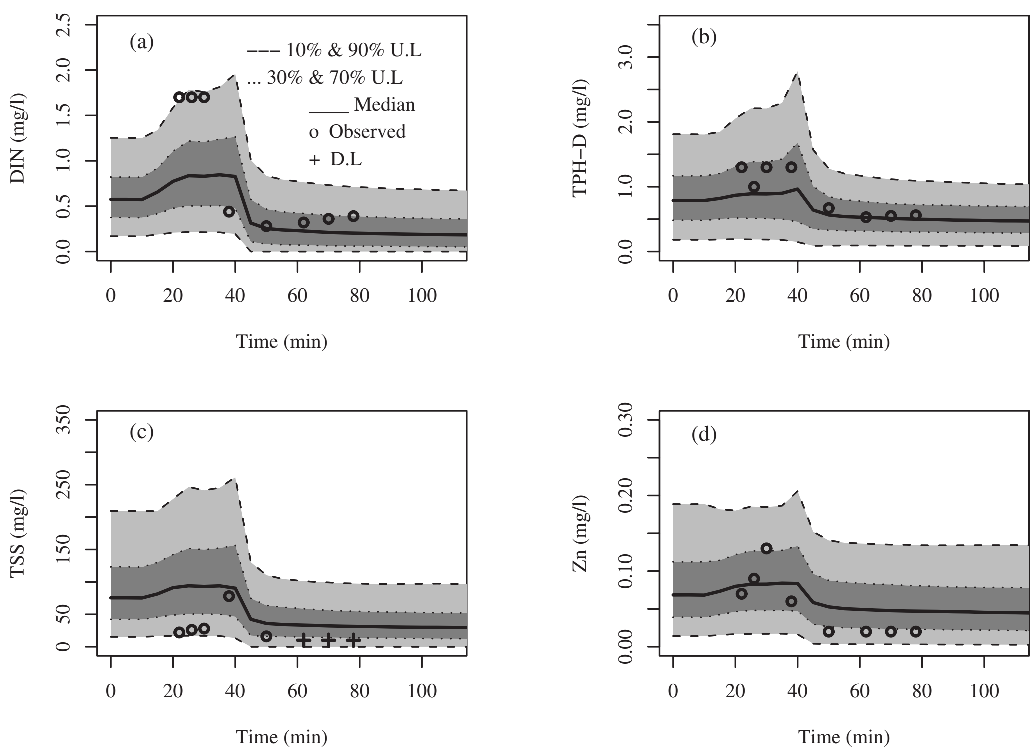

Fig. 8. Monte Carlo simulations of predicted pollutant concentrations at the catchment's outlet for the 07/22/2006 storm event. U.L. =uncertainty limits (as percentiles of the predicted values). The dark shaded region indicates the 30 and 70% U.L. The light gray shaded region indicates the 10 and 90% U.L. Observed concentrations values and the reported detection limits were plotted as discrete points.

the real distributions fitted to the optimized parameter values deviate from normality. Nonparametric tests as the KS test can be unreliable with small data sets.

Conclusions

An existing accumulation and wash-off model was calibrated for a range of contaminants that have little documentation (Zn, TPH-D, and DIN). The model was also calibrated for TSS to allow for a comparison with other studies. The model was calibrated on a storm by storm basis and the variability of the optimal parameters described by normal probability distributions. Parameter values were found for each storm event assuming a null residual mass approach. The ranges of parameter values for TSS were higher than those for Zn, DIN, and TPH-D. This is possibly due to the fact that TSS concentrations are higher in magnitude when compared to the other pollutants. This may be avoided by normalizing the concentration values before running the model.

Concentration values were drawn from a fitted probability distribution when a BDL value was reported. A gamma probability distribution seemed to represent concentration values above the reporting limit. In general, BDL values were reported at the decreasing (falling limb) portion of the hydrograph so that the total washed off mass was not significantly affected. However, a more detailed analysis on how this approach affects the final results is recommended.

Other contribution of this research was the usage of the SA algorithm as an optimization technique when simulating stormwater quality. An annealing schedule ($T=20{,}000$, $a=0.9$, and $m=100$) was identified for the characteristics of the objective function. Computation time was not a concern and the algorithm converged quickly for the majority of the cases.

The model was calibrated for 24 storms and only two rainfall events were left for validation: 06/21/2006 and 07/22/2006. Mixed results were obtained for the storms used during the validation stage. In some instances, observed concentrations fell within the uncertainty limits, defined as the 10th and 90th percentiles of the predictions. However, some individual observed concentrations fell outside the uncertainty limits, indicating poor model performance. Additionally, the total observed washed off mass was compared against the total estimated washed off mass. Although median estimated values fell within the 30th and 70th percentiles, some results indicate that the model could be overestimating total washed off mass values. Calibration of the model can be performed by using storms with similar characteristics (for example, high flow storms versus low flow storms). This approach may narrow the uncertainties of the parameter values, which will possibly make the validation stage more robust.

The mathematical formulation of the wash-off model can be improved. Chen and Adams (2006) proposed a similar wash-off model that incorporates a rainfall-runoff transformation. In their model, the washed off mass M_w is a function of rainfall volume, a dimensionless runoff coefficient, and storage volume. However, Chen and Adams (2006) applied the model to estimate CDF of pollutant loads rather than individual concentrations. Parameter values obtained in this study could be used in more complex applications such as Bayesian based models where an a priori probability distribution is necessary. The application of the model results presented in this paper should be limited to watersheds with similar characteristics and weather conditions. The transferability of results to other watersheds is still a challenge when using accumulation/wash-off models for planning purposes.

Acknowledgments

The UNH Stormwater Center is housed within the Environmental Research Group (ERG) at the University of New Hampshire (UNH) in Durham, N.H. Funding for the program was and continues to be provided by the Cooperative Institute for Coastal and Estuarine Environmental Technology (CICEET) and the National Oceanic and Atmospheric Administration (NOAA). The writers acknowledge the significant contributions made by anonymous reviewers.

References

Ahyerre, M., Chebbo, G., Tassin, B., and Gaume, E. (1998). "Storm water quality modeling, an ambitious objective?" *Water Sci. Technol.*, 37(1), 205–213.

Alley, W. M., and Smith, P. E. (1981). "Estimation of accumulation parameters for urban runoff quality modeling." *Water Resour. Res.*, 17(6), 1657–1664.

APHA, AWWA, and WEF. (2005). *Standard methods for the examination of water and wastewater*, 21st Ed., American Public Association, Washington, D.C.

Beven, K., and Binley, A. (1992). "The future of distributed models: Model calibration and uncertainty prediction." *Hydrolog. Process.*, 6, 279–298.

Chen, J., and Adams, B. J. (2006). "A framework for urban storm water modeling and control analysis with analytical models." *Water Resour. Res.*, 42, W06419. doi:

DeCoursey, D. G. (1985). "Mathematical models for nonpoint water pollution control." *J. Soil Water Conserv.*, 5, 408–413.

EPA. (2006). "Guidance on systematic planning using the data quality objectives process." *Rep. No. EPA QA/G-4*, Environmental Protection Agency, Washington, D.C.

Gaume, E., Villanueve, J.P., and Desbordes, M. (1998). "Uncertainty assessment and analysis of the calibrated parameter values of an urban storm water quality model." *J. Hydrol.*, 210, 38–50.

Haiping, Z., and Yamada, K. (1996). "Estimation for urban runoff quality modeling." *Water Sci. Technol.*, 34(3–4), 49–54.

Haiping, Z., and Yamada, K. (1998). "Simulation of nonpoint source pollutant loadings from urban area during rainfall: An application of a physically-based distributed model." *Water Sci. Technol.*, 38(10), 199–206.

Han, Y., Lau, S., Kayhanian, M., and Stentrom, M. (2006). "Characteristics of highway stormwater runoff." *Water Environ. Res.*, 78, 2377–2388.

Helsel, D. R., and Hirsch, R. M. (2002). *Statistical methods in water resources techniques of water resources investigations*, Book 4, Chap. A3, U.S. Geological Survey.

Hopgood, A. A. (2001). *Intelligent systems for engineers and scientists*, 2nd Ed., CRC, Boca Raton, Fla.

House, M. A., et al. (1993). "Urban drainage—Impacts on receiving water quality." *Water Sci. Technol.*, 27(12), 117–158.

Kanso, A., Gromaire, M. C., Gaume, E., Tassin, B., and Chebbo, G. (2003). "Bayesian approach for the calibration of models: Application to an urban stormwater pollution model." *Water Sci. Technol.*, 47(4), 77–84.

Kanso, A., Tassin, M., and Chebbo, G. (2005). "A benchmark methodology for managing uncertainties in urban runoff quality models." *Water Sci. Technol.*, 51(2), 163–170.

Kirkpatrick, S., Gelatt, C. D., Jr., and Vecchi, M. P. (1983). "Optimization by simulated annealing." *Science*, 220(4598), 671–680.

Millar, R. G. (1999). "Analytical determination of pollutant wash-off parameters." *J. Environ. Eng.*, 125(10), 989–992.

Minton, G. (2002). *Stormwater treatment: Biological, chemical, and engineering principles*, Resource Planning Associates, Washington, D.C.

Pardo-Iguzquiza, E. (1998). "Optimal selection of number and location of rainfall gages for areal rainfall estimation using geostatistics and simulated annealing." *J. Hydrol.*, 210(1–4), 206–220.

Pitt, R., Field, R., Lalor, M., and Brown, M. (1995). "Urban stormwater toxic pollutants: assessment, sources, and treatability." *Water Environ. Res.*, 67, 260–275.

Roseen, R. M., Ballestero, T. P., Houle, J. H., Avellaneda, P. M., Wildey, R., and Briggs, J. (2006). "Storm water low-impact development, conventional structural and manufactured treatment strategies for parking lot runoff." *Transp. Res. Rec.*, 1984, 135–147.

Rossman, L. A. (2005). "Stormwater management model—User's manual, Version 5." *EPA/600/R-05/040*, Water Supply and Water Resources Division, National Risk Management Research Laboratory, EPA, Cincinnati.

Sansalone, J., and Cristina, C. (2004). "First flush concepts for suspended and dissolved solids in small impervious watersheds." *J. Environ. Eng.*, 130(11), 1301–1314.

Sartor, J. D., Boyd, G. B., and Agardy, F. (1974). "Water pollution aspects of street surface contaminants." *J. Water Pollut. Control Fed.*, 46(3), 458–467.

Shaheen, D. G. (1975). "Contributions of urban roadway to water pollution." *Rep. No. EPA-600/2–75–004*, Environmental Protection Agency, Washington, D.C.

Soonthornnonda, P., Christensen, E., Liu, Y., and Li, J. (2008). "A washoff model for stormwater pollutants." *Sci. Total Environ.*, 402, 248–256.

Sumner, N. R., Fleming, P. M., and Bates, B. C. (1997). "Calibration of a modified SFB model for twenty-five Australian catchments using simulated annealing." *J. Hydrol.*, 197(1–4), 166–188.

Tsihrintzis, V. A., and Hamid, R. (1997). "Modeling and management of urban stormwater runoff quality: A review." *Water Resour. Manage.*, 11, 137–164.

Zoppou, C. (2001). "Review of urban storm water models." *Environ. Modell. Software*, 16, 195–231.