

# Appendix 1. Data Processing and Computation to Characterize Hydrology and Compare Water Quality of Four Watersheds in Eastern Puerto Rico

By Robert F. Stallard

## Introduction

This appendix summarizes procedures used to characterize hydrology and to compare water quality among four watersheds studied as part of the U.S. Geological Survey's Water, Energy, and Biogeochemical Budgets (WEBB) research program in Puerto Rico and parallel work in Panama. The four comparative watersheds were the Canóvanas, Cayaguás, Icacos, and Mameyes (fig. 1). An additional small watershed within the Icacos watershed, the Guabá, was examined to study scaling effects; therefore, data from five stream-gaging stations were available (fig. 1). Data from the first 15 years (1991–2005) of the Puerto Rico component were processed for this report, an undertaking that entailed working with several million discharge measurements, thousands of measurements from automated rain gages, and several thousand samples analyzed for some aspect of water quality (from suspended sediment and electrical conductivity to comprehensive chemical analyses). Discharge records from stream-gaging stations and water-quality datasets were processed separately to check for errors and to fill in discharge-data gaps with estimated values. The hydrologic and water-quality datasets were then merged into two types of master files. One contains all primary hydrology data and interpretive information about hydrologic processes, such as calculating trends and tagging recessions and exceptional storms; one hydrologic-data file was compiled for each gage site. The other file type holds primary water-quality data, chemical parameters derived from the primary data, and an abbreviated hydrological characterization of each sample; analyses of samples from all sites and sources were compiled into a single water-quality data file. This water-quality-data file, which contains about 6,400 records for 15 years of data, can be manipulated in conventional spreadsheets.

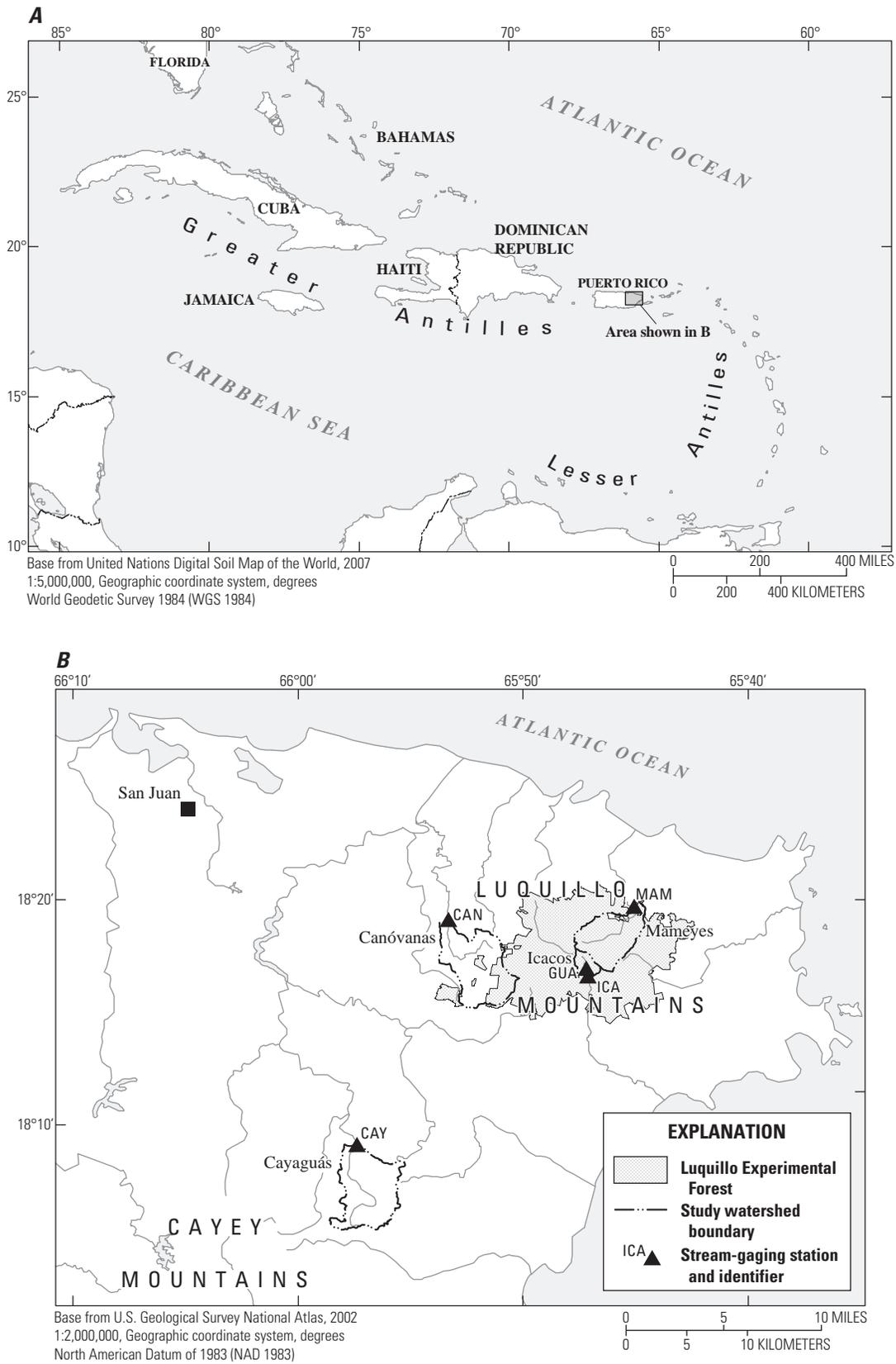
This appendix is divided into seven sections. The first discusses the computational procedures used to handle this large dataset. The second covers hydrological data processing. The third describes processing of the water-quality data, including calculated parameters such as saturation indices, cyclic-salt correction, and dissolved bedrock. The fourth describes the constituent-load computations built around the program LOADEST (Runkel and others, 2004). The fifth details how runoff and yield percentile ranges are calculated. The sixth presents summaries of the dataset. The last compares this dataset with other large datasets assembled by the U.S. Geological Survey, in terms of the range of process rates measured.

## Computational Procedures to Interpret and Merge Hydrology and Chemistry Datasets

This data-processing exercise required the design of computational procedures to handle files that exceed the capacity of many databases and spreadsheets and to link data records on the basis of time stamps and sample identifications. Multimillion-record datasets are well beyond the size normally handled by conventional spreadsheets. Each discharge-data record represents an independent measurement, and subsequent modeling and interpretation is typically based on watershed conditions at or immediately before the time of measurement. A major objective was to prepare the database for models presented in Murphy and Stallard (2012), Stallard (2012b), and Stallard and Murphy (2012).

For much of this work, the program “gawk,” GNU Operating System's implementation of a selection program “awk,” was used (<http://www.gnu.org/software/gawk/gawk.html>) to process data files. Programs that read and process single records at a time are well suited to handling time-series datasets; records are read in chronological order, calculations are made, statistical information is accumulated, and where necessary a short, running history is stored in memory. As processing is completed, the modified records, with computed statistics, are written into a new file. In this manner, complex calculations can be undertaken without having the entire dataset resident in computer memory at once. The “gawk” program has excellent text-parsing capabilities and adequate math and array processing. These features permit the creation of programs that combine data archives in many individual text-based formats into a single standardized text-based format. Capabilities include error recognition, complex renaming, unit conversions, data testing, and more advanced calculations. The “gawk” scripts are compiled when they are called, allowing for fast and effective debugging.

An important feature of an integrated computational package is its ability to reprocess data when appreciable changes are made. The procedures described in this chapter are controlled by a small number of batch-processing scripts. As new data were added or whenever notable changes were made, such as acquisition of better estimates of watershed areas, the entire dataset was rebuilt from primary data sources. In this manner, partial implementation of changes and retention of errors was avoided.



**Figure 1.** Location of Puerto Rico, study watersheds, and stream-gaging stations (CAN, Río Canóvanas near Campo Rico; CAY, Río Cayaguás at Cerro Gordo; GUA, Quebrada Guabá near Naguabo; ICA, Río Icacos near Naguabo; MAM, Río Mameyes near Sabana).

## Noise Reduction, Smoothing, and Derivatives

All hydrologic data contain measurement noise, which complicates calculations such as the rate of change in discharge (the derivative of discharge with respect to time,  $dQ/dt$ ). Time derivatives are used to characterize many hydrologic phenomena, such as the description of rises and recessions (Vogel and Kroll, 1996; Geyer and others, 2008). In general, these characterizations should not make restrictive assumptions about the mathematical functions that are used to describe the data, because such assumptions can bias subsequent modeling. The huge files evaluated here required that simplified calculation tools be used for smoothing data and estimating derivatives. The primary assumption made here is that data, which are expected to match real-world observations, are smoothly varying; step functions and isolated sharp peaks are not physically realistic and indicate faulty data. The filling of intermittent or extended data gaps is discussed in the section "Filling Data Gaps." Spline functions are ideal computational tools for examining smoothly varying data, and two types of spline procedures were used in the data processing here: the relaxed cubic spline and the B-spline convolution (table 1).

The relaxed cubic spline (Reinsch, 1967) is a piecewise, continuous (connecting without gaps) set of cubic polynomials fit through the data, with one polynomial for each data segment. These polynomials are also piecewise continuous in the first and second derivatives. Moreover, in a relaxed cubic spline, the residual error of the regression can be specified beforehand. Data points do not need to be equally spaced along the ordinate axis (time). The time ordinate must increase monotonically, and no two points can have the same ordinate value. The cubic-spline regression uses an array that is seven times as large as the number of data pairs, and when the error is specified, the derivation of spline coefficients is iterative, making it difficult to apply a relaxed cubic spline to an entire time series of millions of measurements.

The B-spline convolution (Jupp, 1976) requires that points be evenly spaced. The advantage of a convolution is that only a few points in the convolution calculation have to be in memory during any computational step. The B-spline both smooths the data and calculates derivatives of the smoothed data at the central point in the convolution. Smoothing is especially useful in reducing measurement noise, which can be recorded during low-flow conditions. Accordingly, smoothed data were used in low-flow estimates, in describing recession curves, and in characterizing data noise. To estimate the derivative of the data during nonlow-flow conditions, a fourth-order function was forced through the central point and two adjacent points on each side, resulting in no smoothing. This procedure was also completed as a convolution.

## Processing Hydrologic Data

Water-stage levels are recorded every 5 minutes at the Río Mameyes stream-gaging station and every 15 minutes for the remaining stations (on Río Canóvanas, Río Cayaguás,

Quebrada Guabá, Río Icacos). Stage levels are subsequently converted into real-time discharge by using a stage-discharge relation. The real-time data are stored in the U.S. Geological Survey's Automated Data Processing System at the Caribbean Water Science Center. For Río Mameyes, the 15-year dataset represented roughly 1.6 million measurements. Real-time-discharge measurements were downloaded from this data processing system and were converted into a more compact format that uses a spreadsheet-style Julian date (days since December 30, 1899) and metric units. Four parameters were generated (table 1): Julian time, raw stage (in meters), corrected stage (in meters), and discharge (in cubic meters per second). Corrected stage refers to documented corrections made in processing the data before acquiring data from the data processing system; normally this category is blank.

The secondary source of discharge data is the official tabulation of daily-mean discharges reported through the U.S. Geological Survey's National Water Information System (NWIS), publicly available on the Web site (<http://waterdata.usgs.gov/nwis>). NWIS data are calculated by using the data in the Automated Data Processing System, and when data are missing in the latter, Water Science Center hydrologists estimate the NWIS data using data from other stream-gaging stations and rainfall. The NWIS data were used to check the efficacy of calculations by the programs described here, and where primary data from the Automated Data Processing System were missing, estimated NWIS daily mean discharge data were used to calculate estimated values to fill those gaps (by using procedures described in the next section).

All data assembly calculations used primary discharge ( $Q$ ) data. Because the drainage areas of the watersheds are different, however, subsequent comparisons used area-normalized properties, specifically runoff rates of water ( $R$ ) and yields of constituents ( $Y_x$ ), which are calculated from water discharge ( $R = Q / \text{area}$ ) and constituent loads ( $Y_x = \text{Load}_x / \text{area}$ ), respectively. Annual runoff also differs among the sites, and discharge-weighted concentrations are used for runoff-normalized comparisons (average concentration of  $X = \text{average yield of } X / \text{average runoff}$ ).

## Site-Specific Issues

A datum shift for the Río Icacos stream-gaging station was identified in 2003 (Matthew Larsen, Caribbean Water Science Center Chief, written commun., January 15, 2003; Diaz and others, 2004), whereby the elevation datum (reference mark (RM)–1) on an old rock stilling well (the original stage measurement recording site, since replaced) shifted downward through subsidence of the manmade structure. From 1979 to 1987, the RM–1 elevation decreased by 0.083 feet (ft); from 1987 to 1995, by 0.111 ft; and from 1995 to 1999, by 0.268 ft. All datums are tied to RM–1, the base reference mark for this station. The Caribbean Water Science Center corrected the record. As a result of this instability, discharge data had been too high (a 0.2-ft change is equal to 10 to 20 cubic feet per second ( $\text{ft}^3 \text{ s}^{-1}$ ) at base-flow conditions). All discharge data for

**Table 1.** Parameters input and calculated in the real-time discharge-data file used to characterize the hydrology of five rivers, eastern Puerto Rico.

Field	Variable	Header	Parameter
Primary data fields			
1	JULIAN	Julian_Time	Julian time in spreadsheet format
2	RAW_ST	Stage_Raw	Raw stage data in meters or coding for missing data. Codes: "--", good discharge, no stage; "==" , estimated discharge. Derivatives and several other derived parameters were not calculated for stream samples when discharges were estimated.
3	STAGE	Stage_Fin	Corrected stage data in meters.
4	QQ	Q_cms	Discharge, $Q$ , in cubic meters per second. If discharge was estimated, the raw stage (RAW_ST) is coded accordingly.
Data fields calculated by the smoothing routine			
5	SMTH_QQ	Smooth_Q	Smoothed discharge, $Q_s$ , in cubic meters per second. Used in low-flow estimates, determining recession curves, and in characterization of data noise.
6	DQ_DT	dQ/dt	The derivative of discharge with respect to time, $dQ/dt$ , in cubic meters per second per hour.
7	SMTH_DQDT	Smooth_dQ/dt	The derivative of smoothed discharge with respect to time, $dQ_s/dt$ , in cubic meters per second per hour.
8	DL10Q_DT	d(log(Q))/dt	The derivative of $\log(Q)$ with respect to time, $d\log(Q)/dt$ , in log(fraction) per hour.
9	SM_DL10QT	Sm_dlogQ/dt	The derivative of $\log(Q_s)$ with respect to time, $d\log(Q_s)/dt$ , in log(fraction) per hour.
Data fields calculated by the antecedent-conditions routine			
10	AVG_Q_1	Avg(Q)_day	Running average discharge for the prior day, in cubic meters per second, for characterization of antecedent conditions.
11	AVG_Q_7	Avg(Q)_week	Running average discharge for the prior week, in cubic meters per second, for characterization of antecedent conditions.
12	AVG_Q14	Avg(Q)_window	Running average discharge for the prior 2-week period, in cubic meters per second, for characterization of antecedent conditions.
13	MAX_Q14	Max(Q)_window	Maximum discharge for the prior 2-week period, in cubic meters per second, for characterization of antecedent conditions.
14	MAX_J14	Max(Q)_julian	Julian time off the maximum discharge for the prior 2-week period for characterization of antecedent conditions.
Baseflow data field calculated by the the baseflow routine			
15	BASEFLOW	Baseflow	Baseflow estimate, in cubic meters per second.
Data fields calculated by the recession and landslide routine			
16	RECESS	Quick_Recess	Quick-recession curve (monotonic decrease of $Q_s$ ), in successive measurements.
17	RECSLO	Slow_Recess	Slow-recession curve (steady drop of AVG_1Q to AVG_1Q to AVG_1Q) in successive measurements. This procedure for calculating recession is similar to most published methods, which are based on mean daily discharge.
18	AVQ_1Q	One-day_Q	Running daily convolution of average flow, in cubic meters per second. For any moment this is an average discharge of one-half day back and one-half day forward in time.
19	OVER_TH	Over-Threshold	If the runoff in the previous 1, 2, or 3 days exceeds the landslide threshold, this runoff is coded in binary: +=1 is 1 day; +=2 is 2 days; +=4 is 3 days.
20	PAST_SD	Slide-day JT	Days since the end of the last period for which runoff in the previous 1, 2, or 3 days exceeds the landslide threshold. If this number is zero, then the observation is within the period where the threshold has been exceeded.

the Río Icacos station were subsequently revised in the Automated Data Processing System and in the NWIS to correct for the datum shift.

The useable record for the Quebrada Guabá stream-gaging station is only 10 years long (1993 through 2002) and ends with gage failure on April 17, 2003, during a large frontal storm that produced several landslides upstream of the gage. A site inspection in August 2010 indicated that pulses of coarse sediment derived from these landslides continue to move through the channel and create a time-varying cross section. Although data for the station were reported by the U.S. Geological Survey after that date, we consider these data to be flawed.

## Filling Data Gaps

Each of the real-time-discharge datasets has gaps ranging from 30 minutes to several months. Commonly, the longest gaps are associated with the loss of a streamgage caused by a huge storm; the length of the gap reflects the time needed to rebuild the gage. These gaps introduce two problems. First, a gap represents a missing water volume that must be estimated to complete annual water and constituent budgets. Second, the computational procedures require evenly spaced, gap-free data. The objective of data interpolation was to calculate estimated discharge in gaps in real-time Automated Data Processing

System data, such that there is either an observed or an estimated discharge for each possible measurement time. Fulfilling this objective makes the smoothing algorithms far more rapid, because convolutions can be used. Gaps were filled by using a cubic spline calculation such that (1) daily mean discharges of the estimated data matched the daily mean discharges in the NWIS, and (2) estimated data were calculated at the same time intervals and are continuous with respect to real data and to first and second derivatives. These discharge estimates are used in daily, monthly, and annual mass balance calculations, but they are flagged as estimated data so that they would not be used to determine discharges of individual samples.

For data gaps of less than a day, discharge values were linearly interpolated from the adjacent data. For gaps too long to be filled by interpolation, both the stage and discharge fields are filled with a no-data code, to ensure that there are no time gaps. For gaps equal to or longer than a day, discharge values were estimated at the same time steps as the real-time data by using NWIS mean-daily discharge; on a few occasions when NWIS estimates were not available, mean-daily discharge values were estimated by using adjacent sites. Daily discharge estimates were downloaded from the NWIS and converted to a simpler format, similar to that described for Automated Data Processing System files.

For gap-filling calculations, mean-daily-discharge files also had to be gap free. The NWIS does not have measured or estimated daily values for the full 15-year record for Río Canóvanas (743 days missing) or Quebrada Guabá (1,541 days missing, including all days after April 17, 2003). For the Canóvanas and Guabá gaps, daily values were estimated by using relations determined from regressions (Bevington and Robinson, 2003) on all combinations of one or more daily-mean-discharge data from nearby gages. For Río Canóvanas, the Río Cayaguás was selected as the independent variable. For Quebrada Guabá, the Río Icacos was the independent variable. The selected regression equations are as follows:

$$Q_{\text{Canóvanas}} = 0.573 \cdot Q_{\text{Cayaguás}} - 0.068 \quad (r^2=0.806)$$

$$Q_{\text{Guabá}} = 0.0298 \cdot Q_{\text{Icacos}} - 0.00102 \quad (r^2=0.872),$$

where  $Q$  is the daily mean discharge at the station indicated by the subscript of the variable, and  $r^2$  is the coefficient of determination and is the fraction of the variance accounted for by the regression (Bevington and Robinson, 2003). A cubic spline was used to fill the gaps. Julian day is the independent variable. For each gapless daily-value file, cumulative total daily discharge in cubic meters is calculated and assigned to the next day; this cumulative total becomes the dependent variable. A nonrelaxed cubic spline (nonrelaxed goes through the data points) is fit through the cumulative data such that at the end of each day, the discharge total calculated by summing the mean-daily discharge data matches the value of the cubic spline. The derivative of this spline is calculated for each time step of the real-time discharge record. This derivative is converted to cubic meters per second and, where real-time discharge data are lacking, these estimates provide a smoothly

varying substitute. The estimated real-time data are used for discharge-related calculations involving time resolution of a day or more, but they are not used to calculate instantaneous-flow statistics or instantaneous mass fluxes.

The final file assembled for all discharge-related calculations was constructed by merging the real-time discharge data from the Automated Data Processing System with the estimated real-time data from the NWIS. These estimates vary smoothly and have the same daily mean-discharge values as the NWIS dataset. The estimated data have no time or discharge gaps, but there are no estimates of stage. The merging process looked for gaps that are greater than a day, because shorter gaps were already filled through interpolation. First, the appropriate discharge estimates were extracted from the smoothed data files derived from daily mean values. A fourth-order polynomial having zero slope at each end and zero area was added to the estimated discharge; this addition adjusted the estimated data to the real data such that all data connect without any abrupt steps or slope breaks. For short gaps of less than 1.5 days, a single adjustment curve is used. For longer gaps, two adjustment curves (of no more than one day in length) were applied, one at each end of the gap. Between the two curves the discharge estimates are not adjusted.

For computational simplicity, the fourth-order polynomial is calculated by assuming that the gap, or the part of the gap that is being adjusted, is normalized such that it starts at zero and has a length of one day. The difference between the observed discharge at the beginning of the gap ( $Q_{O1}$ ) and the estimated discharge ( $Q_{E1}$ ) is  $D_{Q1}$ .  $D_{Q2}$  is calculated the same way at the end of the gap.

$$D_{Q1} = Q_{O1} - Q_{E1}, \text{ and } D_{Q2} = Q_{O2} - Q_{E2}.$$

To simplify the calculation, the duration of the adjustment period was normalized to 1, and a dimensionless time,  $X$ ,  $0 \leq X \leq 1$ , is defined ( $t$  is the time of the adjusted observation, the adjustment start time is  $t_{\text{start}}$ , and the adjusted end is  $t_{\text{end}}$ ):

$$X = (t - t_{\text{start}}) / (t_{\text{end}} - t_{\text{start}}).$$

If the beginning or the end of the polynomial is within the gap, then the respective  $D_{Q1}$  or  $D_{Q2}$  is assigned a zero value. The fourth-order polynomial is given by the following:

$$D_Q(X) = D_{Q1} + c \cdot X^2 + d \cdot X^3 + e \cdot X^4, \quad 0 \leq X \leq 1$$

$$c = -(18 \cdot D_{Q1} + 12 \cdot D_{Q2})$$

$$d = 32 \cdot D_{Q1} + 28 \cdot D_{Q2}$$

$$e = -15 \cdot (D_{Q1} + D_{Q2}).$$

$X$  and  $D_Q(X)$  is calculated for every time step in the adjustment period, and  $D_Q(X)$  is added to the smoothed data.

The resulting gap-filled real-time discharge files are of the following form: Julian time, raw stage (in meters), corrected stage (in meters), and discharge (in cubic meters per second). If the discharge was estimated for gaps of more than a day, a “=” was coded into the raw stage.

## Processing Real-Time-Discharge Data

Sixteen parameters were calculated from gap-filled, real-time discharge data. These parameters, in addition to the four primary measurements, were recorded in a single flat file for subsequent access by other data-characterization programs (table 1). These other programs include graphing programs, a program that links discharges, runoff rates, and rates of change of discharge to individual samples, a program that summarized flow statistics, a program that characterizes recurrence, and a program that does recession analyses. The procedures for estimating these 16 parameters are described below.

### The Smoothing Routine

A variety of processes related to whether samples are collected on the rising or falling limbs of the hydrograph can affect the concentrations of constituents. Whether discharge is increasing or decreasing is determined by the derivative of discharge with respect to time. Measurement errors can produce noise that, if not damped, can make the calculations of derivatives quite difficult. To provide such information for data interpretation, the entire discharge record, including filled gaps, is smoothed by using a B-spline convolution (Jupp, 1976) to calculate a smoothed discharge ( $Q_s$ ) and its derivative ( $dQ_s/dt$ ). Because the errors of  $\log(Q)$  are approximately normally distributed, the calculations used log-transformed data,  $\log(Q)$ . The convolution coefficients (Jupp, 1976, that report's table 6,  $m=4$ ,  $k=2$ ) are as follows:

for the smoothing convolution,  $\log(Q_s)$ : (1, 8, 23, 32, 23, 8, 1)/96

for the derivative,  $d\log(Q_s)/dt$ : (5, 4, 1, 0, -1, -4, -5)/(32•seconds per interval).

Smoothed  $Q_s$  and  $dQ_s/dt$  are calculated from  $\log(Q_s)$  and  $d\log(Q_s)/dt$ , respectively:

$$Q_s = 10^{\log(Q_s)}$$

$$dQ_s/dt = (Q_s \cdot d(\log(Q_s))/dt) / \ln(10)$$

The nonsmoothed derivative of the data is also calculated by using a five-point convolution that is equivalent to fitting a fourth-order polynomial through five consecutive points:

for the derivative,  $d\log(Q)/dt$ : (8, 1, 0, -1, -8)/(12•seconds per interval).

Five fields of data are added after the original four fields (table 1):  $Q_s$ ,  $dQ/dt$ ,  $dQ_s/dt$ ,  $d\log(Q)/dt$ , and  $d\log(Q_s)/dt$  to produce a smoothed-data file.

### Antecedent-Conditions Routine

Antecedent conditions in a watershed are usually assumed to influence the chemistry of streams, and more recent conditions are assumed to be more important than those of the more distant past. To provide information about prior watershed wetness, additional discharge information is extracted from the antecedent data (table 1). Data are

summarized for three time windows: one day, one week, and two weeks. Five new fields of data are appended to the previous nine: a one-day average discharge, a one-week average discharge, a two-week average discharge, the maximum discharge in the previous two weeks, and the Julian time of the maximum discharge in the previous two weeks. In addition, a temporary file of the minimum smoothed discharge,  $Q_s$ , in the previous two weeks, along with the Julian time of this minimum discharge, is generated in order to delineate base flow.

### Base Flow Calculation Routines

When a considerable amount of time had passed after storm events, the water in the channel was assumed to be entirely base flow or else water that was contributed to the river from deeper flow paths (Vogel and Kroll, 1996). A low-flow envelope was defined for the dataset (table 1) of the real-time data (including filled gaps). This envelope was developed iteratively. The table of minimum discharge in the previous two weeks (table 1) and the Julian time of this minimum discharge were used to delineate an envelope of minimum flows. For simplicity the minimums were connected by linear interpolation. This envelope was compared with observed flows,  $Q_s$ , and points below the curve that had been defined by linear interpolation were added to refine the envelope. This procedure was repeated three times, and the base-flow estimate was appended to the data table.

### Recession-Curve and Landslide-Day-Extraction Routines

Two types of phenomena examined in this report, recession curves and landslide-producing storms, rely on characterization of hydrographs of storms. Both of these determinations represent a characterization of antecedent conditions, recession curves for dry periods following storms, and landslide-producing storms, here termed "landslide days," for especially wet events (slide-day in table 1). These routines identify recession curves and landslide days, but do no further calculations.

Recession calculations normally use daily mean discharge (see reviews and model summaries by Hall, 1968; Tallaksen, 1995; Vogel and Kroll, 1996; and Chapman, 1999). Daily mean discharge is used in part because a large body of daily means is available, but it was also used because during low flows, at time scales of less than a day, processes that produce minor fluctuations in discharge (such as transpiration or floating debris) do not substantially affect daily means. Accordingly, two types of recession are identified. In a quick-recession curve, the smoothed discharge,  $Q_s$ , steadily decreases. In a slow-recession curve, daily mean discharge calculated by the antecedent-conditions routine steadily decreases. Each discharge measurement is tagged as being part of neither, one, or both recessions.

Much of the work in the Puerto Rico Water, Energy, and Biogeochemical Budgets (WEBB) program has focused on the landslide process and emphasizes the importance of shallow-soil landslides in physical erosion (Larsen, 2012). Stallard (1999) and Stallard and Kinner (2005) formulated models that include landslide-producing events in characterizations of annual sediment yields; the models are based on an empirical

landslide threshold derived by Larsen and Simon (1993). This landslide threshold relates a rainfall threshold,  $P$  (in millimeters), that must be exceeded for landslides to be initiated, to storm duration,  $D$  (in hours):

$$P(\text{millimeters}) \geq 91.46 \cdot D(\text{hours})^{0.18} . \quad (1)$$

The essence of this relation is that if more than the requisite rain falls in the requisite time, and if the terrain is sufficiently steep ( $>12^\circ$  slope), then landslides may be initiated. For storms with durations of 24, 48, and 72 hours (h), the threshold rainfall is 162, 184, and 197 millimeters (mm) respectively.

Stallard (1999) and Stallard and Kinner (2005) modified this threshold relation to accommodate runoff, rather than rainfall, because many river basins lack adequate rain-gage networks. These reports describe rivers in the Panama Canal watershed for which a decade or more of combined daily-discharge and daily-sediment-load data had been recorded in a terrane with exhumed island-arc geology, topography, and annual rainfall similar to that of eastern Puerto Rico. The daily runoff value,  $R$ , needed to exceed the landslide threshold was calculated from the daily discharge data. A day was defined as a landslide day if the amount of runoff indicates that rainfall during that day exceeded the landslide threshold

$$R(\text{millimeters in } 24, 48, 72 \text{ h}) > F \cdot 91.46 \cdot D^{0.18} \\ (D = 24, 48, 72 \text{ h}), \quad (2)$$

where  $F$  is a catchall empirical coefficient between 0 and 1 that adjusts the threshold for three factors that reduce runoff relative to local rainfall. These three factors are (1) evapotranspiration, (2) infiltration, and (3) rainfall patchiness, whereby parts of a watershed will be over or under the threshold. Because daily sediment data were not available for eastern Puerto Rico, an  $F$  value of 0.85, as determined empirically by Stallard (1999) and Stallard and Kinner (2005) for central Panama watersheds, was used. For storms with a duration of 24, 48, and 72 hours, the threshold runoff is 138, 156, and 167 mm respectively. Calculations were based on whole days, because many sites elsewhere do not have real-time data, and because we could use estimated discharges during real-time data gaps.

Landslide days were identified by comparing runoff totals for the previous 24, 48, and 72 h with the threshold values. The threshold-exceedence state of the river was coded as a binary value. A value of one was added if the threshold is exceeded for the previous day; two was added if the threshold is exceeded for the previous 2 days, and four was added if thresholds were exceeded for the previous 3 days. In addition, the number of days since the landslide threshold was last exceeded was tabulated. This value is zero when the discharge measurement was within a period in which the threshold was exceeded.

## Processing Water-Quality Data

During the eastern Puerto Rico WEBB Program, two principal types of samples were collected: grab samples, which were manually collected on a regular basis from riverbanks

at well-mixed cross sections near each gage site, and event samples, which were collected by an automated sampler during storm events and retrieved when weather conditions permitted. Conductivity, temperature, pH, and dissolved oxygen ( $O_2$ ) of grab samples were measured onsite; only conductivity was measured onsite for event samples, because the other field parameters are unstable if not measured promptly. Depth-integrated samples for suspended sediment were occasionally collected at the same time as a grab sample. All water samples were filtered through 0.2- micrometer ( $\mu\text{m}$ ) filters as soon as practicable in the laboratory. Sample collection and processing are described in detail in appendix 2.

## Measured Water-Quality Parameters

This section describes how water-quality measurements were processed to characterize each water sample and to estimate constituent loads. Constituent measurements include the following:

1. Field properties: sample time, temperature, pH, conductivity, and dissolved oxygen ( $O_2$ ).
2. Major cations: sodium ( $\text{Na}^+$ ), potassium ( $\text{K}^+$ ), magnesium ( $\text{Mg}^{2+}$ ), and calcium ( $\text{Ca}^{2+}$ ).
3. Major anions: chloride ( $\text{Cl}^-$ ) and sulfate ( $\text{SO}_4^{2-}$ ).
4. Alkalinity: (alkalinity =  $\text{HCO}_3^- + 2 \cdot \text{CO}_3^{2-} + \text{OH}^- - \text{H}^+$ ).
5. Dissolved silica: ( $\text{Si}(\text{OH})_4$ ).
6. Dissolved organic carbon (DOC).
7. Nutrient ions: nitrite ( $\text{NO}_2^-$ ), nitrate ( $\text{NO}_3^-$ ), ammonium ion ( $\text{NH}_4^+$ ), and phosphate ( $-\text{PO}_4^{3-}$ : preceding “-” refers to the tendency for phosphate to bind with other constituents).
8. Suspended solids: weight of material filtered onto a 0.2-micrometer ( $\mu\text{m}$ ) filter per unit volume following drying to  $105^\circ\text{C}$ .
9. Suspended bedrock: weight of material filtered onto a 0.2- $\mu\text{m}$  filter per unit volume following ashing at  $550^\circ\text{C}$ . This process is referred to as ashing, and the weight loss is referred to as loss on ignition (LOI).
10. Suspended sand: weight of sand sieved out of the sample using a 65- $\mu\text{m}$  sieve prior to filtration and dried at  $105^\circ\text{C}$ . The weight of this sand is added into measurements of suspended solids. Inspection of sand samples indicated very little clay pellets or organic matter.
11. Trace constituents: strontium ( $\text{Sr}^{2+}$ ), aluminum ( $-\text{Al}^{3+}$ ), manganese ion ( $\text{Mn}^{2+}$ ), reduced iron ( $\text{Fe}^{2+}$ ), total iron ( $\text{Fe}_{\text{total}} = \text{Fe}^{2+} + -\text{Fe}^{3+}$ ), boric acid ( $\text{B}(\text{OH})_3$ ), fluoride ( $\text{F}^-$ ) and bromide ( $\text{Br}^-$ ). Trace constituents were analyzed in a reduced suite of samples.

- Dissolved organic nutrients: dissolved organic nitrogen and dissolved organic phosphorus were analyzed in some samples in the last years of the program (appendix 2).

These constituents were analyzed at several laboratories and by several types of analytical equipment (appendix 2). The data formats from these laboratories differed, and analyses were often reported in different units. For example,  $\text{SO}_4^{2-}$  was reported variously as micromoles per liter, microequivalents per liter, milligrams of sulfur per liter, and milligrams of sulfate per liter. All constituent concentrations were converted to micromoles per liter, except alkalinity, which was converted to microequivalents per liter, and suspended solids, sand, and ashed solids, which were converted to milligrams per liter. Data were standardized iteratively because of typographical errors in transcribing field and laboratory data and ambiguous reporting units. The tracking of anomalous results commonly required consultation of the original field and laboratory notes. Data processing was programmed in the “awk” (“gawk”) language because of its portability and its simplicity in handling both numeric and character data. The final data files were converted into a format suitable for upload into the NWIS.

## Derived Water-Quality Parameters

The interpretation of the analytical results required the calculation of additional, derived, parameters. Table 2 gives constituent properties that were used in deriving some of these parameters. Derived parameters used in developing this report are grouped as follows:

- Error checking: Total cation and total anion charge, net charge, and predicted conductivity (Miller and others, 1988)
- Thermodynamic calculations: Ionic strength, ionic strength correction (gamma), hydrogen ion ( $\text{H}^+$ ), hydroxide ion ( $\text{OH}^-$ ), carbon-dioxide vapor pressure ( $P_{\text{CO}_2}$ ), dissolved carbon dioxide ( $\text{H}_2\text{CO}_3 = \text{CO}_2[\text{d}] + \text{H}_2\text{CO}_3^*[\text{d}]$ ), bicarbonate ion ( $\text{HCO}_3^-$ ), carbonate ion ( $\text{CO}_3^{2-}$ ), the saturation index of calcite ( $\text{SI}_{\text{calcite}}$ ), and oxygen saturation ( $S_{\text{O}_2}$ ).
- Cyclic salt corrections: Large fractions of some constituents are derived from marine solutes blown inland, cyclic salts. Because we assumed that all chloride is cyclic, a marine-derived cyclic salt component was subtracted from various constituents, leaving an estimated bedrock-derived concentration. Following Stallard and Edmond (1981), estimated bedrock-derived constituents are designated with a “\*” —for example,  $\text{Na}^*$ . Cyclic-salt corrections were determined for  $\text{Na}^*$ ,  $\text{K}^*$ ,  $\text{Mg}^*$ ,  $\text{Ca}^*$ ,  $\text{Sr}^*$ ,  $\text{Cl}^*$ , and  $\text{SO}_4^*$ . Other constituents were assumed to have no marine inputs.
- Dissolved bedrock erosion: dissolved bedrock (DBrx).

- Suspended bedrock (SBrx) and particulate organic carbon.

## Error Checking

Two types of error checking were used in quality control of the water-quality data. Charge balance was calculated from the sum of concentrations times valence states (table 2) for the various chemical constituents. To check for errors, total positive and negative charges were compared; major disparities, greater than 5 percent, indicated potential problems. Predicted conductivity was calculated from the equivalent conductivity for each ion (table 2). To check for errors, observed and predicted conductivities were compared (Miller and others, 1988). Because conductivity is a field measurement, this check proved a robust way to track down improperly recorded or handled samples.

## Thermodynamic Calculations

For thermodynamic calculations involving the carbon-dioxide system, equilibrium constants were tabulated as a function of temperature (Stumm and Morgan, 1981). Some of these tabulations were in the form of equations, but many were table entries. Table entries were converted into equations as a function of temperatures and are given below. To correct for the effects of ionic strength, the Davies equation was used to calculate the ionic-strength correction, gamma. Stumm and Morgan (1981) present all equations used in the thermodynamic calculations. Oxygen saturation is given by an equation developed by Elmore and Hayes (1960).

Equilibrium constants (K) are given by the following equations (K refers to temperature in degrees Kelvin; the subscripts of K refer to different reactions, where “W” refers to the deprotonation of water, “H” refers to Henry’s law for total dissolved carbon dioxide, “1” and “2” refer to the first and second deprotonations of carbonic acid, respectively, and “sp” refers to the solubility product of calcite; T refers to temperatures in degrees Celsius; and  $\text{mg L}^{-1}$ , milligrams per liter):

- Reaction  $\text{H}_2\text{O} \leftrightarrow \text{H}^+ + \text{OH}^-$ ,  
 $K_{\text{w}} = -4470.99/\text{K} + 6.0875 - 0.01706 \cdot \text{K}$
- Reaction  $\text{CO}_2[\text{g}] + \text{H}_2\text{O} \leftrightarrow \text{H}_2\text{CO}_3^*$ ,  
 $K_{\text{H}} = 2201.32/\text{K} - 12.60325 + 0.01258 \cdot \text{K}$
- Reaction  $\text{H}_2\text{CO}_3^* \leftrightarrow \text{H}^+ + \text{HCO}_3^-$ ,  
 $K_1 = -3406.15/\text{K} + 14.844 - 0.03277 \cdot \text{K}$
- Reaction  $\text{HCO}_3^- \leftrightarrow \text{H}^+ + \text{CO}_3^{2-}$ ,  
 $K_2 = -2910.82/\text{K} + 6.54767 - 0.02386 \cdot \text{K}$
- Reaction  $\text{CaCO}_3 \leftrightarrow \text{Ca}^{2+} + 2\text{CO}_3^{2-}$ ,  
 $K_{\text{sp}} = -3000/\text{K} + 13.543 - 0.04001 \cdot \text{K}$
- Reaction  $\text{O}_2[\text{g}] \leftrightarrow \text{O}_2[\text{d}]$ ,  
 $\text{O}_2$ , saturation in  $\text{mg L}^{-1} = 14.652 - 0.41022 \cdot \text{T} + 0.007991 \cdot \text{T}^2 - 0.00007777 \cdot \text{T}^3$

**Table 2.** Measured constituent properties used in calculating interpretive parameters and derived constituents in five rivers, eastern Puerto Rico.

[g mol<sup>-1</sup>, grams per mole; μS cm<sup>-1</sup> L mEq<sup>-1</sup>, microsiemens per centimeter per milliequivalent per liter; μmol L<sup>-1</sup>, micromoles per liter; μEq L<sup>-1</sup>, microequivalents per liter; mg L<sup>-1</sup>, milligrams per liter; DOC, dissolved organic carbon; --, not relevant]

Constituent	Database units	Molecular weight (g mol <sup>-1</sup> )	Valence	Equivalent conductivity (μS cm <sup>-1</sup> L mEq <sup>-1</sup> )	Ratio in seawater, constituent to chloride	Cyclic-salt-correction ratio to chloride	Bedrock chemical form	Conversion factor to bedrock (g mol <sup>-1</sup> )
Measured constituents								
Na <sup>+</sup>	μmol L <sup>-1</sup>	22.9898	1	50.1	0.85251	0.85251	Na <sub>2</sub> O	30.99
K <sup>+</sup>	μmol L <sup>-1</sup>	39.102	1	73.5	0.01790	0.01790	K <sub>2</sub> O	47.1
Mg <sup>2+</sup>	μmol L <sup>-1</sup>	24.312	2	53.1	0.10145	0.09689	MgO	40.31
Ca <sup>2+</sup>	μmol L <sup>-1</sup>	40.08	2	59.5	0.01876	0.01876	CaO	56.08
Sr <sup>2+</sup>	μmol L <sup>-1</sup>	87.62	2	59.4	0.00016	0.00016	SrO	103.62
Si(OH) <sub>4</sub>	μmol L <sup>-1</sup>	96.115	0	0	0	0	SiO <sub>2</sub>	60.08
Alkalinity	μEq L <sup>-1</sup>	61.0173	-1	44.5	0.00714	-0.10992	--	--
Cl <sup>-</sup>	μEq L <sup>-1</sup>	35.453	-1	76.4	1	1	--	--
SO <sub>4</sub> <sup>2-</sup>	μmol L <sup>-1</sup>	96.06	-2	80	0.052	0.10597	S	32.06
NO <sub>3</sub> <sup>-</sup>	μmol L <sup>-1</sup>	62.0049	-1	71.5	0	0	--	--
NO <sub>2</sub> <sup>-</sup>	μmol L <sup>-1</sup>	46.0055	-1	71.5	0	0	--	--
NH <sub>4</sub> <sup>+</sup>	μmol L <sup>-1</sup>	18.0388	1	73.4	0	0	--	--
-PO <sub>4</sub> <sup>3-</sup>	μmol L <sup>-1</sup>	94.9714	-2	57.4	0	0	P <sub>2</sub> O <sub>5</sub>	70.97
Oxygen	mg L <sup>-1</sup>	31.9988	0	0	--	--	--	--
DOC	μmol L <sup>-1</sup>	12.0111	-0.07	40.9	0	0	--	--
-Al <sub>3+</sub>	μmol L <sup>-1</sup>	26.9815	0	68.4	0	0	Al <sub>2</sub> O <sub>3</sub>	50.98
-Fe <sub>3+</sub>	μmol L <sup>-1</sup>	55.847	0	68.4	0	0	FeO+Fe <sub>2</sub> O <sub>3</sub>	71.73
B	μmol L <sup>-1</sup>	10.811	0	--	0.00076	0.00076	--	--
F <sup>-</sup>	μmol L <sup>-1</sup>	18.9984	-1	55.4	0.00012	0.00012	--	--
Br <sup>-</sup>	μmol L <sup>-1</sup>	79.909	-1	78.1	0.00153	0.00153	--	--
Derived constituents								
H <sup>+</sup>	μmol L <sup>-1</sup>	1.00797	1	349.8	--	--	--	--
OH <sup>-</sup>	μmol L <sup>-1</sup>	17.0074	-1	198.6	--	--	--	--
H <sub>2</sub> CO <sub>3</sub>	μmol L <sup>-1</sup>	62.0253	0	0	--	--	--	--
HCO <sub>3</sub> <sup>-</sup>	μmol L <sup>-1</sup>	61.0173	-1	44.5	--	--	--	--
CO <sub>3</sub> <sup>2-</sup>	μmol L <sup>-1</sup>	60.0093	-2	80	--	--	--	--

## Cyclic Salt Corrections

In order to estimate bedrock contributions to the dissolved load, the major cations (Na<sup>+</sup>, K<sup>+</sup>, Mg<sup>2+</sup>, Ca<sup>2+</sup>), and sulfate (SO<sub>4</sub><sup>2-</sup>) require corrections for atmospheric inputs. For Na<sup>+</sup>, K<sup>+</sup>, and Mg<sup>2+</sup>, the correction involves subtracting a seasalt contribution based either on the ratio of that ion to chloride (Cl<sup>-</sup>) in rain or on the seasalt ratio (table 2). If we use sodium as an example, the Na<sup>+</sup>:Cl<sup>-</sup> mole ratio for seasalt is 0.8525:

$$\text{Na}^* = \text{Na}^{2+} - (\text{Na}^+:\text{Cl}^-) \cdot \text{Cl}^- \quad (3)$$

Following Stallard and Edmond (1981), the “\*” designates an estimated bedrock-derived constituent. Note that the seasalt correction, based on measured chloride, also effectively removes NaCl derived from human activities. The Mg<sup>2+</sup>:Cl<sup>-</sup> ratio in rain is 0.09689, less than the seasalt ratio of 0.10145 (see discussion in Stallard, 2012a). The seasalt SO<sub>4</sub><sup>2-</sup>:Cl<sup>-</sup> ratio is 0.052, and in some storms, especially high-chloride events, the rain and runoff are near this ratio (Stallard and Murphy, 2012, their fig. 10). However, additional

sulfur comes from pollution and from emissions of organic sulfur compounds (dimethyl sulfide and methyl sulfide) from the marine environment (Stallard and Edmond, 1981). To further complicate matters, forests may also emit organic sulfur compounds back to the atmosphere and sulfur may also be sequestered in vegetation (Stallard and Murphy, 2012). The overall rain SO<sub>4</sub><sup>2-</sup>:Cl<sup>-</sup> ratio is 0.10597, and this ratio was used in calculations. Potassium was handled differently. The K<sup>+</sup>:Cl<sup>-</sup> ratio in rain is 0.0231, whereas the seasalt ratio is 0.0179. Because plant material is so rich in potassium, the greater rain ratio is seen as reflecting local, minor plant contamination, so the seasalt ratio was used. The Ca<sup>2+</sup>:Cl<sup>-</sup> ratio in rain is 0.01876, and seasalt is therefore a minor contribution to the calcium budget. Sahara dust brings in additional soluble calcium (Stallard, 2012a). Because this calcium is also a minor part of the overall calcium budget, and because the dust fallout comes from long-range transport, a uniform dust correction was subtracted from the final calcium budget rather than from individual samples. All silicon, aluminum, and iron in Sahara dust was assumed to be in nonreactive clays and sesquioxides and is therefore insoluble (Stallard, 2012a). Owing

to biological activity, surface seawater has low concentrations of  $\text{Si}(\text{OH})_4$ ,  $\text{NO}_2^-$ ,  $\text{NO}_3^-$ ,  $\text{NH}_4^+$ , and  $-\text{PO}_4^{3-}$ , and essentially zero ratios of these constituents to  $\text{Cl}^-$ .

For individual samples, the atmospheric-correction calculation occasionally yields a negative value. This value may be due to real processes, such as cation exchange in soil that can alter cation proportions, or it may be due to true variations in the ratios to chloride during individual events, as happens with sulfate. With sodium and magnesium, the calculations calculate the differences of large numbers, so analytical errors can produce negative values. This large-number problem is quite severe for high-chloride events (see Stallard, 2012a; Stallard and Murphy, 2012), so high-chloride events are not included in the bedrock-contribution calculations. Regardless of the reason, if the result of the calculation is negative, it was reassigned a zero value, essentially indicating that the sample contained no effective bedrock contribution of that constituent.  $\text{Mg}^*$  and  $\text{SO}_4^*$  calculations result in many zero values;  $\text{Na}^*$  and  $\text{K}^*$  rarely do; and  $\text{Ca}^*$  never.

### Dissolved-Bedrock Erosion

Conventionally, solute denudation rates are calculated by using the concentration of total dissolved solids, defined as the total of the mass concentration of all dissolved constituents. This total is multiplied by discharge to get load, and then divided by drainage area to get the yield of dissolved solids, which is in turn equated with solute denudation. This calculation fails to correct for atmospheric inputs which, in addition to all chloride present and a considerable portion of the major cations, also includes the bicarbonate and carbonate ions that balance most of the bedrock-derived cation charge. Moreover, it fails to recognize that silicate rocks, particularly igneous and metamorphic rocks, are essentially a mixture of complex oxides (table 2), for which the charge of all the soluble cations and silicon was once balanced by oxygen ( $\text{O}^{2-}$ ).

Stallard (1995a,b) uses a different approach that expresses the concentrations of most bedrock-derived constituents as dissolved oxides for mass-balance calculations. Silica and each cation, after correction for atmospheric inputs, are converted into a mass concentration of the corresponding bedrock oxide (table 2). For example, the calculation of dissolved sodium concentration expressed as an oxide ( $\text{Na}_2\text{O}[d]$ ) (table 2) from bedrock-derived sodium,  $\text{Na}^*$ , was performed as follows ( $\text{mg L}^{-1}$ , milligrams per liter;  $\mu\text{mol L}^{-1}$ , micromoles per liter, 30.98, conversion factor in  $\text{g mol}^{-1}$  from table 2):

$$\text{Na}_2\text{O}[d] (\text{mg L}^{-1}) = 0.001 \cdot 30.98 \cdot \text{Na}^* (\mu\text{mol L}^{-1}). \quad (4)$$

Dissolved sulfate is converted into sulfide. The concentrations of the dissolved oxides and sulfide are summed to get total dissolved bedrock (DBrx). For the WEBB watersheds, DBrx estimates were approximately half of the total dissolved solids, an indication of the serious degree of error introduced by estimating bedrock solute yields on the basis of total dissolved solids. Obviously, if the bedrock of a watershed includes substantial carbonates (Stallard,

1995a), for which the charge of calcium and magnesium is balanced by carbonate ions ( $\text{CO}_3^{2-}$ ) rather than by oxygen ( $\text{O}^{2-}$ ), DBrx will underestimate the amount of dissolved bedrock.

In addition to calculating DBrx for samples with complete chemistry (about 870 samples), it was estimated in all samples for which conductivity, chloride, and silica were measured (about 3,780 samples). High-chloride events were excluded, because of the problems of calculating the differences of large numbers. This calculation involves three steps. First, the bedrock contribution to conductivity is estimated by subtracting an atmospheric contribution. The atmospheric contribution is calculated by using the equivalent conductivities of each ion (table 2) to estimate conductivity of a solution that has the chloride concentration of the sample with each ion in the atmospheric-correction ratio to chloride just indicated:

$$\text{conductivity}^* = \text{conductivity} - 0.13028 \cdot \text{Cl}^- \quad (5)$$

The assumption is that the “acid” rain caused by additional sulfate in the rain is mixed with soil and groundwaters that have positive alkalinity. The excess  $\text{H}^+$  in the rain is neutralized by this positive alkalinity, and although a negative alkalinity is used in cyclic-salt correction (table 2), the hydrogen-ion contribution to rainwater conductivity is ignored. Interestingly, this correction is very close to the sea-salt correction (0.13142) without excess sulfate.

The relation between conductivity\* and total bedrock-derived cation charge, ZBrx, is estimated through a regression involving the 1,040 completely analyzed samples. A single set of regression coefficients worked for all Puerto Rico WEBB rivers:

$$\text{ZBrx} = a \cdot (\text{conductivity}^*)^b \quad (6)$$

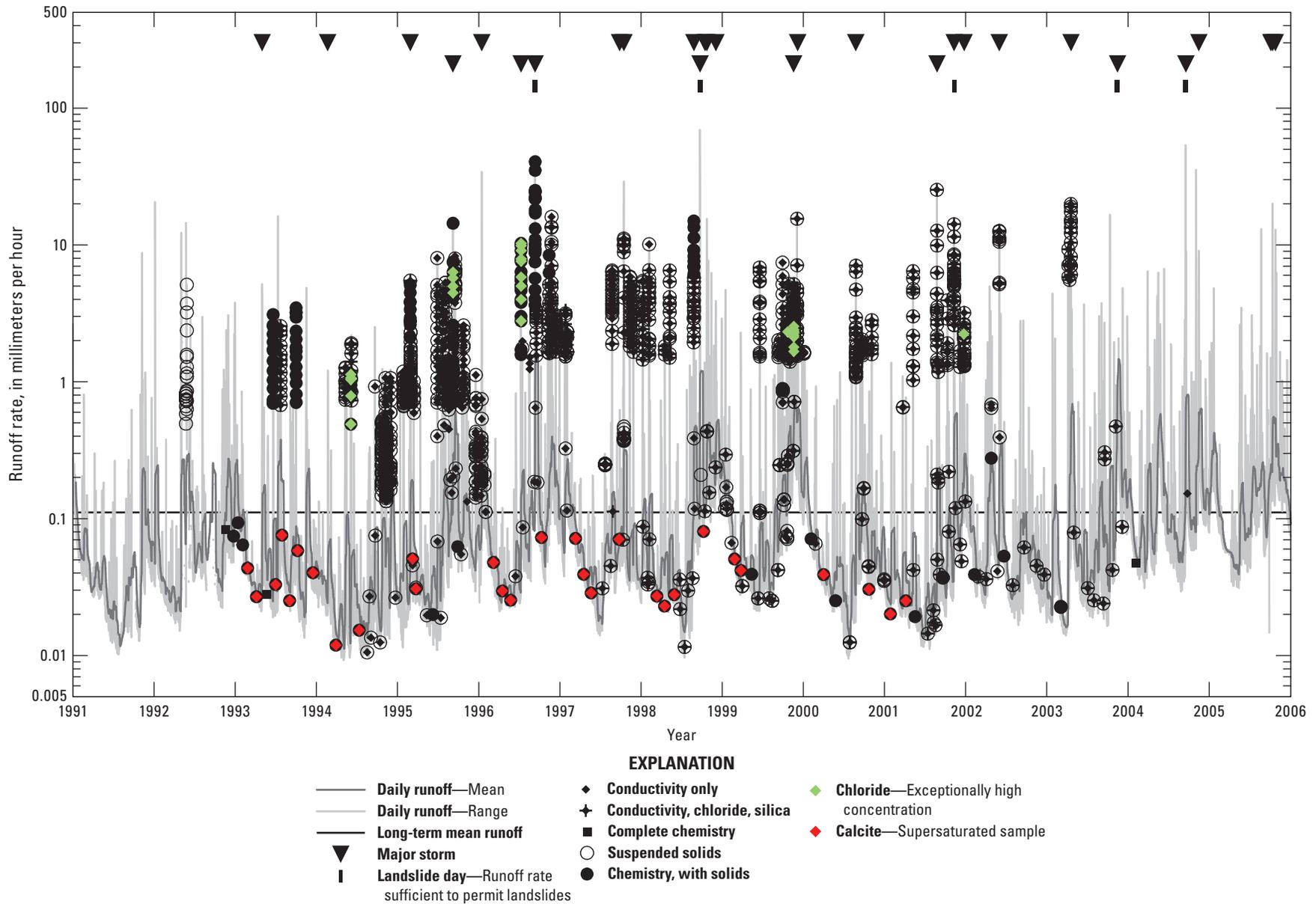
where  $a$  (1.08612) and  $b$  (0.83239) are regression coefficients ( $r^2 = 0.963$ ). Using ZBrx, estimated total dissolved bedrock, DBrx', is determined:

$$\text{DBrx}' = c \cdot (\text{ZBrx})^d + \text{SiO}_2[d] \quad (7)$$

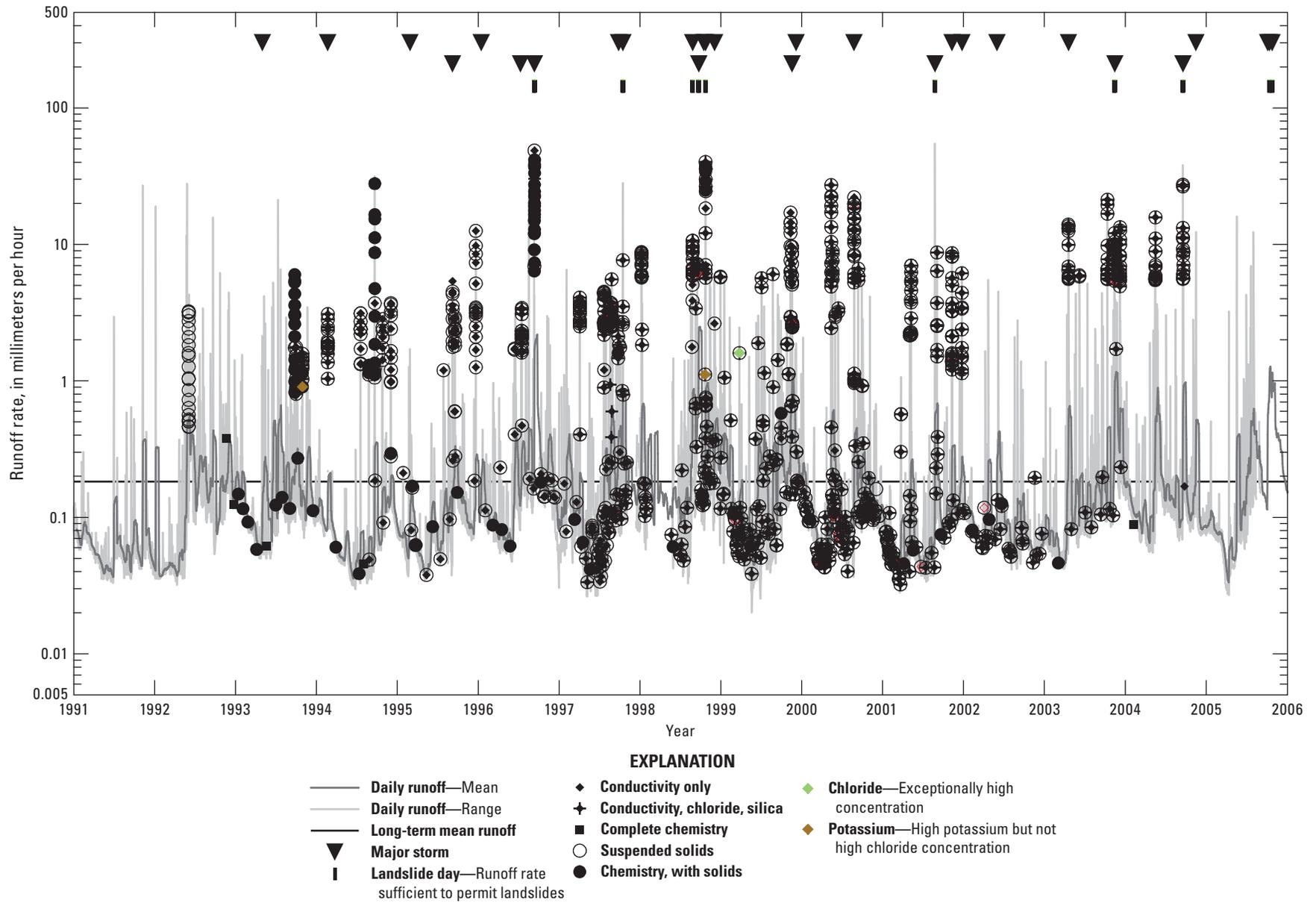
where  $c$  and  $d$  are regression coefficients determined for each river (table 3), and  $\text{SiO}_2[d]$  is the dissolved silica concentration expressed as an oxide:

$$\text{SiO}_2[d] (\text{mg L}^{-1}) = 0.06008 \cdot \text{Si} (\mu\text{mol L}^{-1}). \quad (8)$$

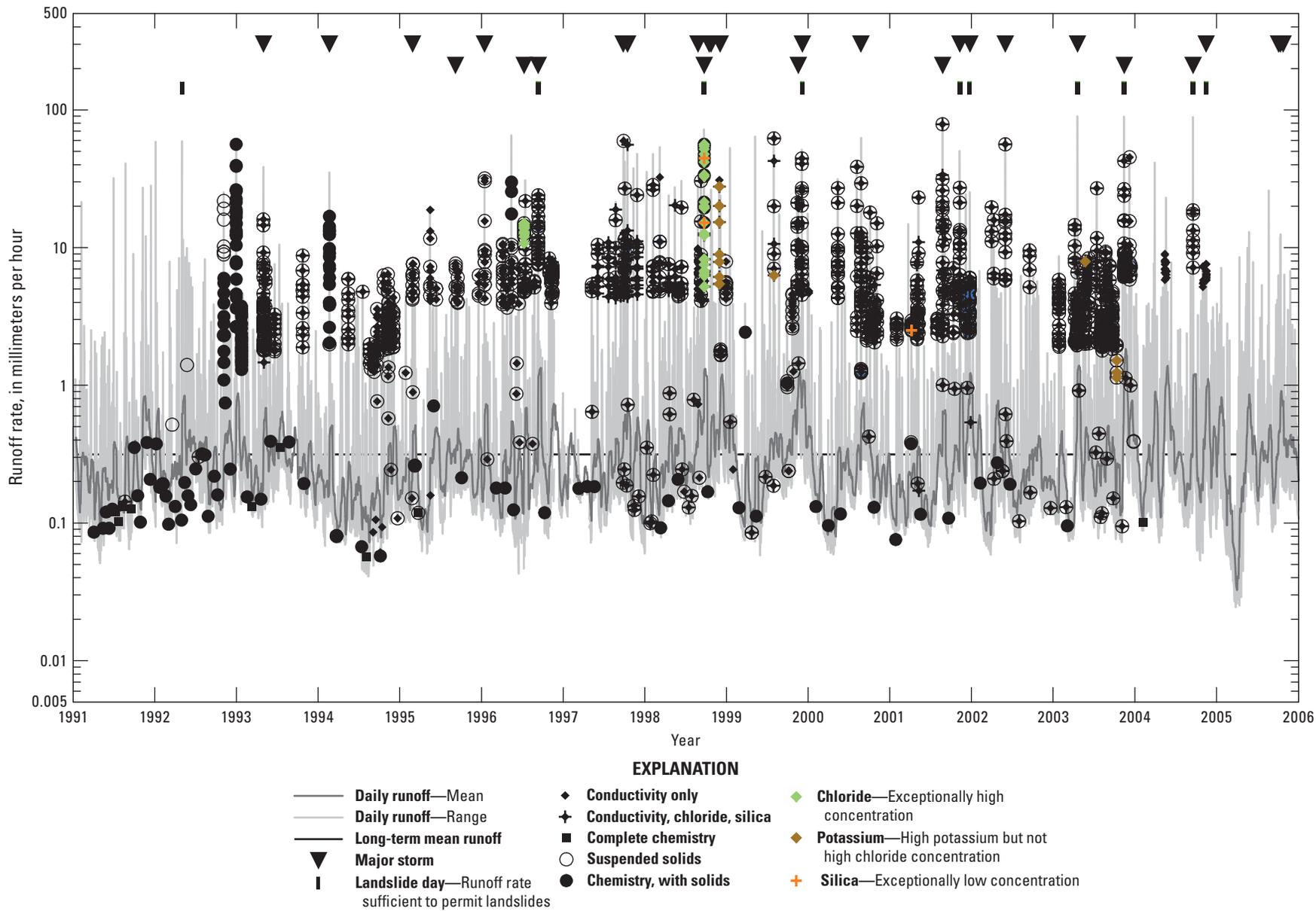
The correlation between DBrx' and DBrx for all 1,040 samples is  $r^2 = 0.990$ , and between  $\log(\text{DBrx}')$  and  $\log(\text{DBrx})$ ,  $r^2 = 0.986$ . Accordingly, both DBrx and DBrx', about 4,650 samples (figs. 2–6), were used in load calculations. Instantaneous dissolved-bedrock erosion rate is the product of either DBrx or DBrx' and instantaneous runoff rate. The large number of determinations of either DBrx or DBrx' gives considerably more rigor to estimates of long-term bedrock erosion through leaching.



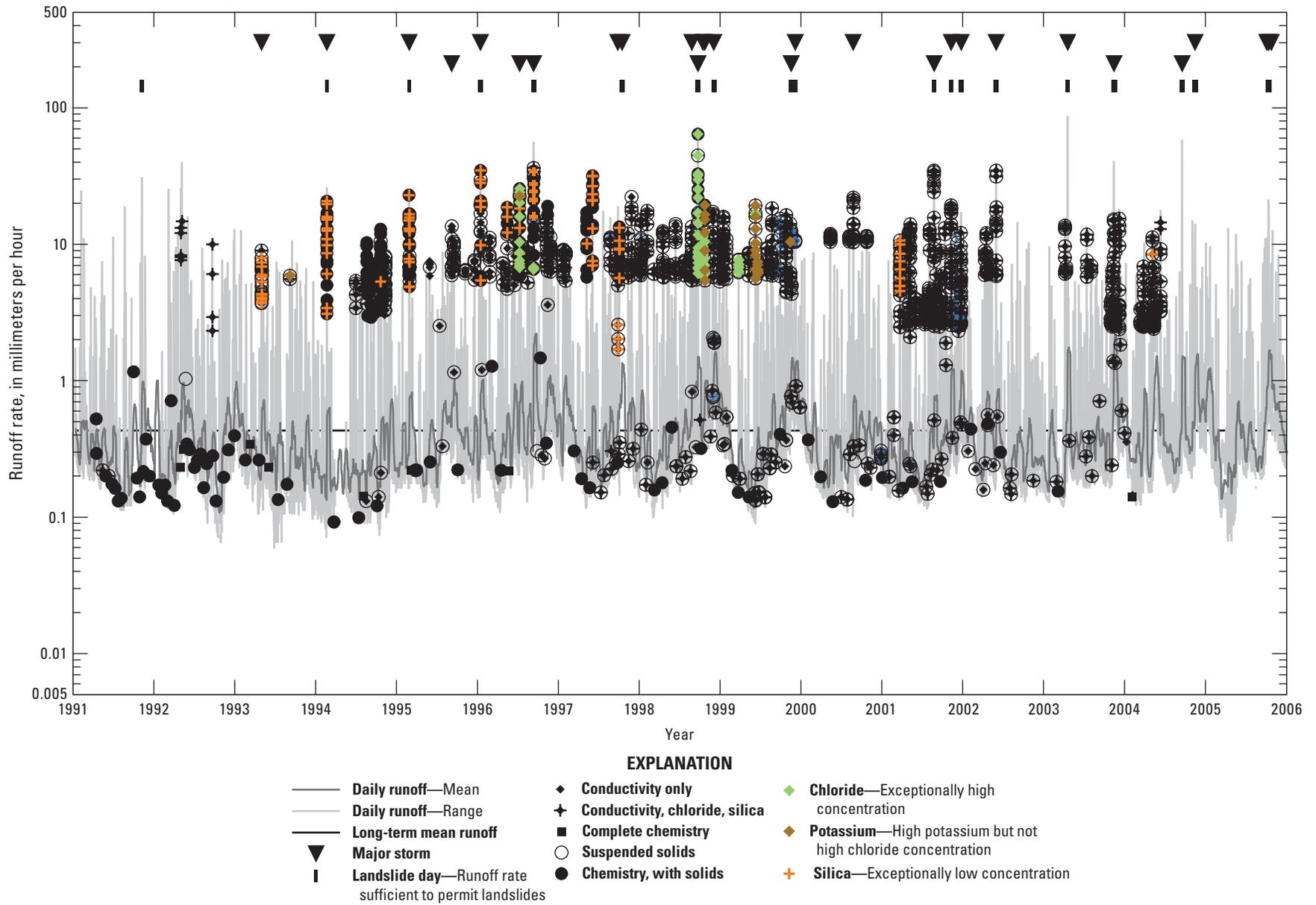
**Figure 2.** Data record for stream-gaging station Río Canóvanas near Campo Rico, Puerto Rico, station number 50061800. Daily range of runoff (real-time data), daily mean runoff, and date on which water sample collected and type of analysis (Stallard and Murphy, 2012).



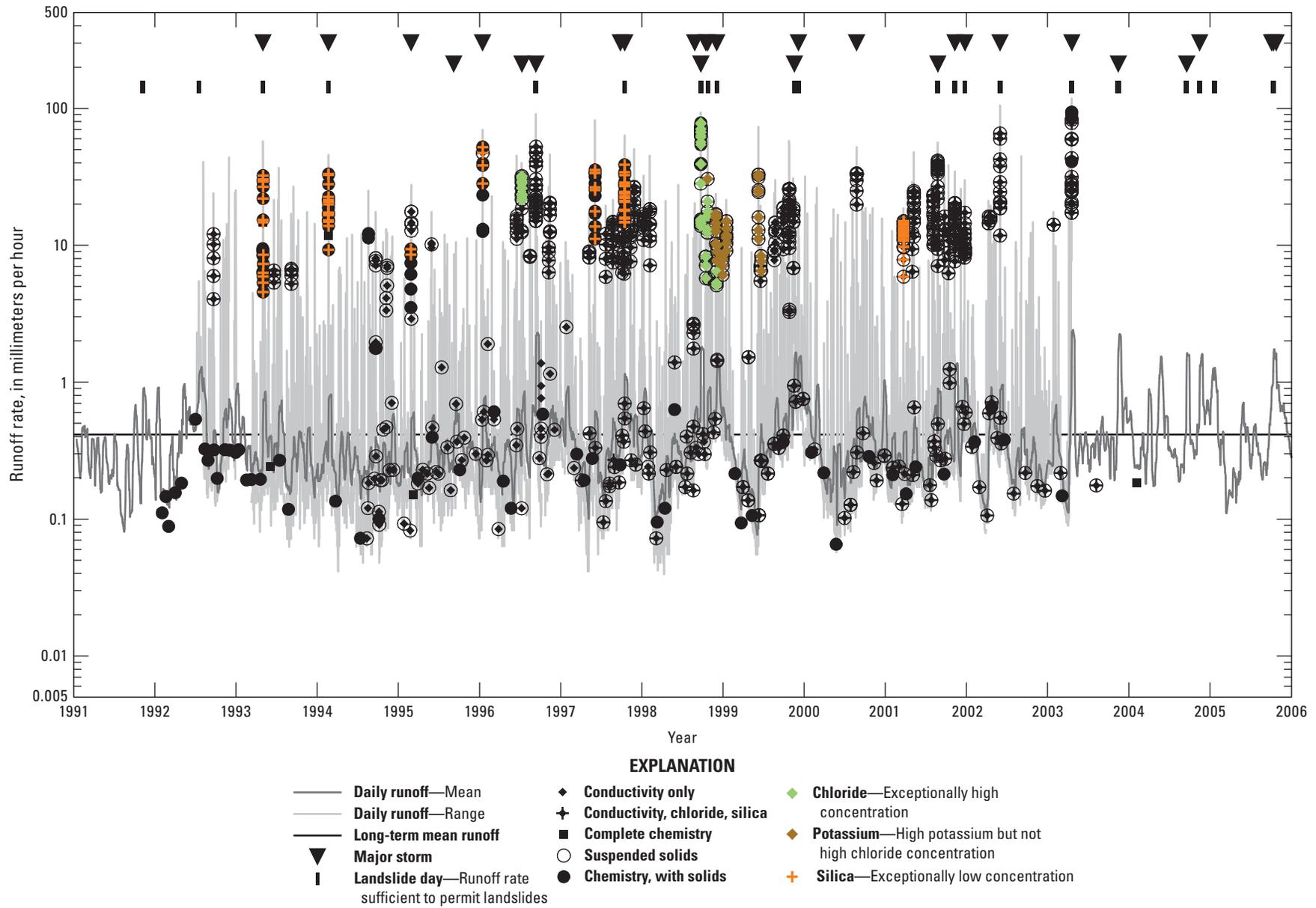
**Figure 3.** Data record for stream-gaging station Río Cayaguás at Cerro Gordo, Puerto Rico, station number 50051310. Daily range of runoff (real-time data), daily mean runoff, and date on which water sample collected and type of analysis (Stallard and Murphy, 2012).



**Figure 4.** Data record for stream-gaging station Río Mameyes near Sabana, Puerto Rico, station number 50065500. Daily range of runoff (real-time data), daily mean runoff, and date on which water sample collected and type of analysis (Stallard and Murphy, 2012).



**Figure 5.** Data record for stream-gaging station Río Icacos near Naguabo, Puerto Rico, station number 50075000. Daily range of runoff (real-time data), daily mean runoff, and date on which water sample collected and type of analysis (Stallard and Murphy, 2012).



**Figure 6.** Data record for stream-gaging station Quebrada Guabá near Naguabo, Puerto Rico, station number 50074950. Daily range of runoff (real-time data), daily mean runoff, and date on which water sample collected and type of analysis (Murphy and Stallard, 2012). First 1.6 and final 2.7 years of daily runoff were estimated using the Río Icacos data record.

**Table 3.** Coefficients of a model relating bedrock-derived cation charge, ZBrx, to dissolved bedrock, DBrx ( $DBrx = c \cdot ZBrx^d + SiO_2$ ) in five watersheds, eastern Puerto Rico.

Watershed	Coefficient		Correlation <sup>1</sup>	Count
	<i>c</i>	<i>d</i>		
Canóvanas	-1.431	0.955	1.000	163
Cayaguás	-1.279	0.914	1.000	127
Mameyes	-1.348	0.958	1.000	177
Icacos	-1.393	0.941	0.999	249
Guabá	-1.398	0.941	0.997	156

<sup>1</sup>Correlation between log(DBrx) and predicted log(DBrx).

### Suspended Bedrock and Particulate Organic Carbon

River-borne sediment is composed of material from three sources: primary and secondary mineral matter derived from the weathering of bedrock, water that has been incorporated by secondary minerals during weathering, and organic matter. For about two-thirds of the measurements of suspended sediment (see Stallard and Murphy, 2012, and appendix 2), we also measured loss on ignition (LOI)—the weight loss of samples that were dried (at 105°C) upon heating up to 550°C. Studies summarized in Mackenzie (1957) and Mackenzie and Cailière (1979) show that in heating to 550°C, all of the clays and sesquioxides lose nonstructural water (most below 200°C). In addition, all sesquioxides, most kaolins, and substantial illites and vermiculites lose lattice water below 550°C; however, smectites, talcs, and chlorites do not. Talc and chlorites are not reported from soils or sediments; smectites are rare. Nonmineral organic carbon and most mineral carbon (coal and graphite) are also oxidized in this range. Carbonates, which are absent in humid-tropical soils and sediment, are generally stable.

In this study, the weight of the sample after heating to 550°C is equated with “suspended bedrock” (SBrx). The sample dried to 105°C is referred to as suspended solids or suspended sediment (SSol). For samples in which suspended bedrock was not measured, regressions for each river relate SBrx to SSol:

$$LOI = SSol - SBrx, \quad (11)$$

and

$$SBrx = SSol \cdot (a + b \cdot \log(SSol)), \quad (12)$$

where *a* and *b* are regression coefficients (table 4).

**Table 4.** Coefficients for a model relating suspended bedrock, SBrx, to suspended solids, SSol ( $SBrx = SSol \cdot (a + b \cdot \log(SSol))$ ) in five watersheds, eastern Puerto Rico.

Watershed	Coefficient		Correlation <sup>1</sup>	Count
	<i>a</i>	<i>b</i>		
Canóvanas	0.809	0.0126	1.000	873
Cayaguás	0.905	-0.0007	1.000	802
Mameyes	0.647	0.0709	1.000	133
Icacos	0.720	0.0477	1.000	646
Guabá	0.651	0.0696	1.000	1,035

<sup>1</sup>Correlation between log(SBrx) and log(SSol).

Particulate organic carbon was not measured on most of the sediment samples; instead, LOI was used as a surrogate measurement. LOI is often equated with organic matter; however, this procedure is not used with samples that are rich in kaolinite, halloysite, or iron and aluminum sesquioxides such as the WEBB samples. Instead, it was assumed here that particulate organic carbon is a fixed fraction of LOI for all the WEBB samples. This fraction was determined such that the 15-year yield of SSol for the Río Icacos consisted of 1 percent carbon, a value that is typical of rivers elsewhere in the world (Stallard, 1998). The result is that LOI for the WEBB rivers consists of 11 percent carbon. Because pure organic matter is about 50 percent carbon by mass (Stallard, 1998), 11 percent of LOI is the noncarbon elements in organic matter, and 78 percent is water. For the forested rivers, at low runoff rates of less than 1 millimeter per hour, the average particulate organic carbon is 2 to 3.5 percent of SSol (Stallard, 2012b, its tables 2, 5).

### Calculation of Discharge, Constituent Loads, Runoff, and Constituent Yields

In order to calculate yields, the real-time discharge and water-quality datasets needed to be merged. This merger matched timestamps in the water-quality dataset with those in the real-time-discharge dataset, and interpolated where necessary. In addition to pairing water quality with discharge, additional properties were appended to each record in the water-quality dataset. These properties included runoff rate (*R*), days since the last landslide day, estimated base flow, and  $d \log(Q_s) / dt$ . Runoff rate was typically used for comparison, because it is area-normalized.

Of the variety of approaches that could be used to estimate constituent loads, the load estimator program LOADEST (Runkel, and others, 2004) was chosen. LOADEST estimates constituent loads in streams and rivers given a time series of streamflow, additional independent variables, and a constituent concentration. LOADEST assists the user in developing a regression model of constituent load (calibration). The explanatory variables normally used in LOADEST are discharge, seasonal cycles (sine and cosine of time), and time. The formulated regression model is then used to estimate loads throughout a user-specified time interval (estimation).

The construction of the input data files used by LOADEST from our two types of datasets was automated. For each river, the LOADEST calibration file is derived from relevant entries in the water-quality file, and the LOADEST estimation file is derived from the real-time discharge-data file. LOADEST requires only minor adaptations of these two datasets. First, the shortest time step allowed for the time series of discharge and other data in the estimation file is one hour. This time step requirement requires subsampling of eastern Puerto Rico real-time discharge data into one-hour intervals (which creates a small (less than 0.1 percent) difference with

the runoff values calculated using all runoff measurements). Second, model run times allowed use of only the simplest option for standard-error calculation (defined by LOADEST parameter SEOPT).

The SEOPT parameter allows selections of increasingly rigorous calibration options. The calibration and estimation procedures within LOADEST are based on three statistical estimation methods: adjusted maximum likelihood estimation, maximum likelihood estimation, and least absolute deviation, an alternative to maximum likelihood estimation that can be used when residuals are not normally distributed. The simplest option for calculating standard error, SEOPT=1, calculates standard error for an adjusted-maximum-likelihood-estimation-based linear approximation; standard error is not calculated for maximum likelihood estimation and least absolute deviation. The data processing for the eastern Puerto Rico dataset required one hour of computer time. The next-level option, SEOPT=2, additionally calculates standard errors for maximum likelihood estimation and least absolute deviation. That option, SEOPT=2, was run for 2 weeks on an otherwise unused computer. On the basis of the number of calculations that had been completed at the end of 2 weeks, more than 150 days of computer time would have been required to calculate standard errors for maximum likelihood estimation and least absolute deviation for all parameters of interest. Therefore, running the model while using SEOPT=2 or 3 was not viable.

To examine the eastern Puerto Rico dataset, a variety of yield models was tested in LOADEST. The user can define the model to be used in LOADEST; alternatively, LOADEST comes with a suite of regression options (defined by LOADEST parameter MODNO). If MODNO is set to zero, then LOADEST automatically selects the best model from a set of nine models (MODNO=1 to 9; Runkel, and others, 2004). This suite of models combines linear and quadratic fits to the natural log of discharge ( $\ln[\text{discharge}]$ ) and time and to the sine or cosine of time ( $\sin[\text{time}]$  or  $\cos[\text{time}]$ ). A regression to time and to time squared implies a systematic drift in chemistry. The effects of increased runoff may be adequately represented by the  $\ln(\text{discharge})$  part of the regression. Accordingly, five model options were chosen. LOADEST was allowed to select the best option (MODNO=0), which could be a model with systematic drift. In addition, all options that did not include systematic drift (MODNO=1, 2, 4, and 6) were run separately and compared with the selected model. In reporting average yields, the results of MODNO=0 and the best driftless model are tabulated.

LOADEST was used to calculate the loads of  $\text{Na}^+$ ,  $\text{K}^+$ ,  $\text{Mg}^{2+}$ ,  $\text{Ca}^{2+}$ ,  $\text{Cl}^-$ ,  $\text{SO}_4^{2-}$ ,  $\text{Si}(\text{OH})_4$ , DOC,  $\text{NO}_3^-$ ,  $\text{NH}_4^+$ ,  $-\text{PO}_4^{3-}$ , and suspended sediment (Stallard and Murphy, 2012). LOADEST was also used to calculate the yields of DBrx, SBrx,  $\text{Na}^*$ , and particulate organic carbon (Stallard, 2012b). Alkalinity was calculated from the other models as described below.

Fundamentally, LOADEST is a regression, and biases produced by the logarithmic transformation of data and non-normality of log-transformed data produce problems that can

invalidate results (Runkel and others, 2004). In particular, zero values complicate LOADEST calculations.  $\text{Mg}^*$  and  $\text{SO}_4^*$  calculations result in many zero values;  $\text{Na}^*$  and  $\text{K}^*$  rarely; and  $\text{Ca}^*$  never. Of these,  $\text{K}^+$  acts as a nutrient, and  $\text{Ca}^{2+}$  is primarily bedrock derived. Accordingly, only  $\text{Na}^*$  yield, derived using LOADEST, is used directly in budget calculations. For all other bedrock constituents, the atmospheric-input corrections are made afterwards on calculated constituent yields, by using the calculated chloride yields, and in the case of  $\text{Ca}^*$ , on additional deposition from Saharan dust (Stallard, 2012b).

Alkalinity loads cannot be estimated directly using LOADEST, because in samples collected at highest discharge values in the Icos and Guabá watersheds, alkalinity is frequently negative (acidity). The primary source of alkalinity is the weathering of silicate and carbonate minerals. This weathering consumes hydrogen ions from the water, releasing nontitratable cations but essentially no nontitratable anions. Organic acids, sulfuric acid, and nitric acid can cause negative alkalinities. In load estimates, alkalinity was calculated from the concentrations of all other constituents (concentrations in mole units):

$$\text{Alkalinity} = (\text{Na}^+ + \text{K}^+ + \text{NH}_4^+ + 2 \cdot (\text{Mg}^{2+} + \text{Ca}^{2+} + \text{Sr}^{2+}) - (\text{Cl}^- + \text{NO}_3^- + 2 \cdot \text{SO}_4^{2-} + Z_{\text{DOC}} \cdot \text{DOC}) \quad (9)$$

where  $Z_{\text{DOC}}$  is the negative charge per carbon (equivalents per mole). For each river,  $Z_{\text{DOC}}$  was calculated by adjusting predicted alkalinity in the water-quality data file to measured alkalinity such that the ratio of averages was 1:1. For the Canóvanas, calculated  $Z_{\text{DOC}} = 0.08$  equivalents per mole ( $\text{eq mol}^{-1}$ ); Cayaguás,  $Z_{\text{DOC}} = 0.06$   $\text{Eq mol}^{-1}$ ; Mameyes,  $Z_{\text{DOC}} = 0.07$   $\text{Eq mol}^{-1}$ ; Icos,  $Z_{\text{DOC}} = 0.05$   $\text{Eq mol}^{-1}$ , and Guabá,  $Z_{\text{DOC}} = 0.05$   $\text{Eq mol}^{-1}$ . For comparison, Tardy and others (2005) estimate an optimum  $Z_{\text{DOC}}$  of 0.064  $\text{Eq mol}^{-1}$  for Amazonian lowland rivers.

Other samples showed anomalous trends on plots of concentration against runoff rate (figs. 2–6). Some of these anomalies were caused by processes that are exceptional, such as high-chloride storms, that are not related to the explanatory variables normally used by LOADEST—discharge, seasonal cycles, and time trends. Statistically these samples are outside a normal distribution, and they were not included in load estimation. High-chloride samples are caused when winds during major storms blow exceptionally large amounts of seasalt inland (Stallard, 2012a). Samples collected during high-chloride events were censored for major seawater constituents ( $\text{Na}^+$ ,  $\text{K}^+$ ,  $\text{Mg}^{2+}$ ,  $\text{Ca}^{2+}$ ,  $\text{Cl}^-$ ,  $\text{SO}_4^{2-}$ ). These same samples also had exceptionally low  $\text{NO}_3^-$ , perhaps reflecting rainwater composition (Stallard and Murphy, 2012), and  $\text{NO}_3^-$  was also censored. Following Hurricane Georges, a high-chloride event, a number of subsequent events demonstrated high potassium, but not high concentrations of chloride and other seasalt constituents. The elevated potassium is likely attributed to the decay of downed vegetation (Stallard and Murphy, 2012). High-potassium samples were censored only for potassium. A few events affecting the

Icos River during the first several years of the program had exceptionally low silica concentrations (Stallard and Murphy, 2012). Because silica concentrations in most event samples during this same time interval were not low, silica concentrations in low-silica samples were also censored.

## Calculation of Runoff and Yield Percentile Ranges

To compare equivalent samples among rivers, runoff rates were ranked by their overall contribution to annual runoff. This ranking allowed equivalent classes of samples to be compared among the five rivers despite differences in annual runoff. All water-quality samples were assigned into runoff-rate percentile classes, similar to the approach used by the U.S. Geological Survey Water Watch Web site (<http://water.usgs.gov/waterwatch/>) in monitoring flood and drought conditions in the United States. Water Watch ranks discharge into percentile classes based on the amount of time a river flows in a given discharge range and compares daily discharge to this ranking. For this report, where the focus is on mass balance, percentile classes (0–10, >10–25, >25–50, >50–75, >75–90, >90–95, >95–99, and >99 percent) are based on runoff and yields. For example, the tenth percentile corresponds to a runoff rate below which 10 percent of the total runoff has been discharged from the basin. Likewise, the tenth percentile in sediment yield corresponds to a runoff rate below which 10 percent of the total sediment mass has been discharged from the basin. When we compared rivers based in percentile classes, our implicit assumption was that each class is hydrologically equivalent in each river (see Stallard and Murphy, 2012). When base flow is referred to in this text, it typically represents the samples in the 0-to-10-percent range.

To calculate percentile classes, the volume of water or mass of a constituent being discharged at a particular runoff was assigned to and totaled in a logarithmic runoff bin, with 20 bins per log cycle. For runoff, the real-time discharge-data files were used, whereas for constituent yields, the individual output load files of LOADEST were used. The bins were cumulatively totaled and normalized by the total volume of water or constituent mass discharged from the watershed for the period 1991–2005. Percentile classes for each constituent were then calculated from the normalized mass-accumulation curve for that constituent. Each sample in the water-quality database was then assigned to a runoff percentile class and rising and falling stages were distinguished on the basis of  $d\log(Q_s)/dt$ . Predicted constituent yields and instantaneous runoff were totaled for each percentile class, and an average concentration for each percentile class was then calculated by dividing total yield for each constituent by total runoff within that class.

## Dataset Features and Characteristics

The combined datasets are presented in two ways: a graphical time series of daily mean runoff (figs. 2–6), simplified sample information (tables 5, 6), and annual summaries of the same information. The simple classification is used to assess the degree of completeness of sample analyses and unusual chemical characteristics. The degree of completeness includes a chemistry characterization and a sediment characterization. In terms of increasing degrees of analysis for chemistry, we distinguished among (1) samples with only conductivity measured, (2) samples with at least conductivity,  $Cl^-$ , and  $Si(OH)_4$  measured, and (3) samples with complete chemistry. We also distinguished the presence or absence of suspended sediment data. Special chemical characteristics are included because of the way several of these were handled in load-estimation calculations and because their biogeochemical significance is discussed in other chapters of this report. These special chemical characteristics are (1) high-chloride samples, (2) normal-chloride, high-potassium samples, (3) low-silica samples, and (4) samples that are supersaturated with respect to calcite.

The sampling design used in the Puerto Rico WEBB program is clearly evident by the distribution of sample-data points in figures 2 through 6. Two primary types of samples were collected, manual “grab” samples and automated “event” samples (appendix 2). Grab samples were collected by researchers or a field crew who visited each site at regular intervals. Most samples collected during low runoff (figs. 2–6) were grab samples, whereas samples collected during high runoff were event samples. At intermediate runoff, relatively fewer samples were collected, an artifact of the sampling protocol. Owing to the large volume of samples collected by automated sampling, the stage threshold for event sampling was adjusted, typically by an increase, first in 1994 and again in 1997, with the objective of sampling the biggest events. This increase is reflected by upward shifts in the event-data distribution evident in figures 2 through 6. Samples in the intermediate zone were typically collected during multiphase events and during falling stage after single events.

## Comparison of Hydrologic Sampling in the Eastern Puerto Rico Large-Scale Study with Hydrologic Sampling in Other Large-Scale Studies

In many settings, rare, high-energy hydrologic events play a major, often dominant, role in the mobilization of river-borne materials from hillslopes and stream channels and in the transport of these materials downriver (Wolman and Miller, 1960). In the WEBB watersheds described in this report, work on water quality emphasized event sampling. From 1991

**Table 5.** Annual hydrologic characteristics, data quality, and sampling effort in five watersheds, eastern Puerto Rico.

[mm, millimeters; mm h<sup>-1</sup>, millimeters per hour; km<sup>2</sup>, square kilometers; -- no samples]

Year	Annual runoff total (mm)	Minimum runoff rate (mm h <sup>-1</sup> )	Minimum daily runoff rate (mm h <sup>-1</sup> )	Average runoff rate (mm h <sup>-1</sup> )	Maximum daily runoff rate (mm h <sup>-1</sup> )	Maximum runoff rate (mm h <sup>-1</sup> )	Number of landslide day periods	Total of landslide day runoff (mm)	Days of gaps	Sediment only	Conductivity only	Conductivity and sediment	Conductivity, chloride, silica	Conductivity, chloride, silica, sediment	Full chemistry	Full chemistry, sediment	Lacking runoff data
Canóvanas (station 50061800, area 25.49 km <sup>2</sup> )																	
1991	471	0.010	0.011	0.054	1.51	8.7	--	--	28.2	--	--	--	--	--	--	--	--
1992	907	0.008	0.013	0.103	1.58	20.6	--	--	86.3	24	--	--	--	--	1	1	--
1993	667	0.014	0.014	0.076	2.79	16.2	--	--	21.6	--	--	--	--	13	1	37	--
1994	371	0.009	0.010	0.042	0.71	2.5	--	--	2.5	--	8	136	2	16	--	8	--
1995	810	0.011	0.012	0.092	3.45	14.7	--	--	--	--	27	258	--	1	--	31	--
1996	1,449	0.021	0.021	0.165	16.98	40.5	1	449	25.7	--	5	38	--	55	--	46	1
1997	697	0.014	0.015	0.080	2.72	29.0	--	--	--	2	1	37	6	74	--	4	--
1998	1,547	0.010	0.010	0.177	8.09	69.0	1	370	--	--	--	4	--	77	--	14	--
1999	1,314	0.021	0.022	0.150	2.12	16.2	--	--	--	7	5	54	--	176	--	4	--
2000	600	0.010	0.011	0.068	1.49	7.1	--	--	1.4	--	--	20	--	46	--	4	--
2001	710	0.011	0.012	0.081	3.85	25.3	--	--	--	--	--	1	1	99	--	4	--
2002	678	0.017	0.018	0.077	2.98	12.7	--	--	--	--	--	5	--	13	--	3	--
2003	1,416	0.014	0.015	0.162	5.99	20.0	1	242	12.9	--	--	--	--	28	--	1	1
2004	1,539	0.030	0.032	0.175	6.76	53.6	1	215	--	--	--	2	--	--	--	1	--
2005	1,448	0.015	0.032	0.165	2.62	20.0	--	--	5.3	--	--	--	--	--	--	--	--
Total	974	0.008	0.010	0.111	16.98	69.0	4	1,276	183.8	33	46	555	9	598	2	158	2
Cayaguás (station 50051310, area 26.42 km <sup>2</sup> )																	
1991	652	0.033	0.035	0.074	3.71	27.0	--	--	--	--	--	--	--	--	--	--	--
1992	1,546	0.032	0.033	0.176	3.25	27.8	--	--	--	22	--	--	--	--	2	--	--
1993	1,603	0.050	0.051	0.183	4.37	21.2	--	--	--	--	--	--	--	19	1	24	--
1994	852	0.030	0.030	0.097	3.55	31.4	--	--	19.3	--	--	35	1	15	1	23	--
1995	939	0.035	0.035	0.107	1.96	13.4	--	--	10.1	--	2	44	--	--	--	4	--
1996	2,240	0.030	0.049	0.255	19.86	48.7	1	551	41.3	1	--	19	7	15	--	23	2
1997	1,759	0.026	0.032	0.201	6.79	28.0	1	196	41.0	--	--	38	4	99	--	4	7
1998	3,297	0.033	0.053	0.376	15.86	43.3	3	1,024	156.5	--	2	19	--	133	--	9	43
1999	2,123	0.020	0.033	0.242	4.03	17.1	--	--	--	--	--	23	1	89	--	4	1
2000	1,339	0.037	0.041	0.152	4.61	27.4	--	--	15.9	3	--	33	--	113	--	4	4
2001	1,061	0.028	0.032	0.121	5.52	54.8	1	42	3.1	--	--	4	--	101	--	4	2
2002	904	0.041	0.042	0.103	2.09	5.5	--	--	18.5	--	--	1	--	24	--	3	--
2003	1,913	0.042	0.043	0.218	5.55	21.2	1	220	--	--	--	2	--	68	--	1	--
2004	1,768	0.055	0.058	0.201	5.71	37.9	1	105	--	2	--	2	--	28	--	1	--
2005	2,237	0.027	0.029	0.255	3.72	16.0	2	78	134.0	--	--	--	--	--	--	--	--
Total	1,615	0.020	0.029	0.184	19.86	54.8	10	2,217	439.7	28	4	220	13	704	4	104	59
Mameyes (station 50065500, area 17.82 km <sup>2</sup> )																	
1991	2,329	0.051	0.053	0.266	3.66	40.8	--	--	1.0	--	--	--	--	2	4	11	2
1992	2,901	0.071	0.091	0.330	4.29	60.6	1	52	12.1	10	--	1	--	1	--	58	--
1993	2,690	0.071	0.095	0.307	3.02	38.5	--	--	45.2	--	--	--	7	74	2	37	--
1994	1,677	0.041	0.043	0.191	2.90	35.1	--	--	4.1	--	17	98	--	17	2	24	--
1995	2,323	0.060	0.065	0.265	1.86	18.8	--	--	--	--	15	42	--	--	1	3	--
1996	2,929	0.043	0.068	0.333	10.72	65.2	1	308	47.5	--	10	53	12	73	--	7	--
1997	2,428	0.060	0.063	0.277	3.01	65.3	--	--	64.0	--	3	1	67	39	--	3	--
1998	3,646	0.079	0.087	0.416	8.42	72.0	1	316	--	--	21	14	35	53	--	18	--
1999	3,247	0.058	0.075	0.371	3.70	63.9	--	--	--	--	16	13	3	37	--	4	--
2000	2,218	0.059	0.062	0.253	4.04	62.5	--	--	--	--	5	6	3	91	--	4	--

**Table 5.** Annual hydrologic characteristics, data quality, and sampling effort in five watersheds, eastern Puerto Rico.—Continued

[mm, millimeters; mm h<sup>-1</sup>, millimeters per hour; km<sup>2</sup>, square kilometers; -- no samples]

Year	Annual runoff total (mm)	Minimum runoff rate (mm h <sup>-1</sup> )	Minimum daily runoff rate (mm h <sup>-1</sup> )	Average runoff rate (mm h <sup>-1</sup> )	Maximum daily runoff rate (mm h <sup>-1</sup> )	Maximum runoff rate (mm h <sup>-1</sup> )	Number of landslide day periods	Total of landslide day runoff (mm)	Days of gaps	Sediment only	Conductivity only	Conductivity and sediment	Conductivity, chloride, silica	Conductivity, chloride, silica, sediment	Full chemistry	Full chemistry, sediment	Lacking runoff data
Mameyes (station 50065500, area 17.82 km <sup>2</sup> )—Continued																	
2001	2,633	0.067	0.072	0.301	4.20	78.8	2	234	--	--	--	29	13	138	--	4	--
2002	2,405	0.059	0.063	0.275	3.41	56.2	--	--	2.3	--	--	--	--	28	--	3	--
2003	3,400	0.073	0.077	0.388	9.57	89.8	2	575	--	3	--	7	--	275	--	1	--
2004	3,553	0.053	0.060	0.405	8.01	88.6	2	397	--	--	--	34	--	7	--	1	--
2005	2,929	0.025	0.027	0.334	2.19	26.0	--	--	--	--	--	--	--	--	--	--	--
Total	2,753	0.025	0.027	0.314	10.72	89.8	9	1,883	176.3	13	87	298	140	835	9	178	2
Icacos (station 50075000, area 3.26 km <sup>2</sup> )																	
1991	3,089	0.113	0.113	0.353	4.76	30.7	1	147	71.6	--	--	--	--	5	--	13	4
1992	3,552	0.066	0.077	0.404	5.04	39.4	--	--	16.5	1	--	--	10	--	2	19	1
1993	2,935	0.059	0.097	0.335	2.13	12.5	--	--	11.2	--	--	6	--	19	2	5	1
1994	2,384	0.071	0.088	0.272	3.65	25.8	--	--	48.3	1	3	57	--	5	1	70	--
1995	3,512	0.119	0.134	0.401	3.95	24.6	--	--	--	--	17	49	--	--	4	21	--
1996	4,989	0.109	0.122	0.568	17.93	55.6	2	674	--	1	5	53	--	46	1	48	--
1997	3,934	0.117	0.130	0.449	5.07	32.7	1	135	37.6	--	1	4	2	84	--	34	26
1998	4,960	0.100	0.114	0.566	7.48	64.1	2	539	12.6	--	11	40	1	151	--	14	--
1999	4,193	0.090	0.099	0.479	5.91	19.7	2	356	--	--	--	37	1	137	--	4	--
2000	3,328	0.101	0.126	0.379	4.53	22.7	--	--	--	1	--	5	--	48	--	4	--
2001	3,395	0.116	0.121	0.388	5.89	37.9	3	483	--	--	--	23	1	217	--	4	--
2002	3,102	0.130	0.142	0.354	6.38	35.1	1	180	3.2	--	--	8	--	36	--	3	--
2003	4,169	0.119	0.131	0.476	13.86	86.0	3	807	--	--	--	1	--	97	--	1	--
2004	4,606	0.123	0.134	0.524	10.97	57.5	2	685	--	--	--	2	--	74	--	1	--
2005	4,295	0.067	0.076	0.490	5.11	21.1	1	301	5.9	--	--	--	--	--	--	--	--
Total	3,762	0.059	0.076	0.429	17.93	86.0	18	4,307	206.8	4	37	285	15	919	10	241	32
Guabá (station 50074950, 0.114 km <sup>2</sup> )																	
1991	2,701	0.074	0.074	0.308	4.44	5.0	1	149	365.0	1	--	--	--	--	--	--	1
1992	4,211	0.068	0.076	0.479	4.71	43.9	1	38	201.7	--	--	--	--	5	--	18	5
1993	2,962	0.063	0.104	0.338	3.50	57.2	1	53	67.7	--	--	1	--	8	2	23	1
1994	2,512	0.041	0.079	0.287	5.36	45.6	1	206	9.1	1	--	24	--	--	1	26	--
1995	2,764	0.063	0.096	0.315	2.89	27.6	--	--	4.2	--	2	24	--	--	--	9	--
1996	4,114	0.063	0.090	0.468	20.26	91.0	3	735	18.5	--	5	31	--	59	--	12	7
1997	3,484	0.041	0.074	0.398	4.52	82.1	1	107	--	--	2	11	1	64	--	38	--
1998	4,247	0.040	0.069	0.485	12.60	93.3	3	714	--	--	1	10	--	94	--	11	--
1999	4,117	0.055	0.055	0.470	7.10	73.2	2	475	--	--	--	26	--	78	--	4	--
2000	3,487	0.057	0.061	0.397	4.14	52.0	--	--	--	--	--	2	--	17	--	4	--
2001	3,523	0.094	0.115	0.402	6.79	56.2	3	637	15.9	--	--	2	--	152	--	4	3
2002	2,963	0.079	0.084	0.338	7.89	105.1	1	232	20.0	--	--	--	--	30	--	3	--
2003	4,453	0.090	0.096	0.508	18.57	118.8	3	1052	258.0	--	--	1	--	60	--	6	38
2004	4,588	0.107	0.110	0.522	11.36	13.1	2	769	366.0	--	--	11	--	8	--	1	18
2005	4,255	0.047	0.060	0.486	5.27	5.8	2	422	365.0	--	--	--	--	--	--	--	--
Total	3,624	0.040	0.055	0.414	20.26	118.8	24	5588	1691.2	2	10	143	1	575	4	159	73

**Table 6.** Number and types of samples collected in different discharge classes and special sample types in five rivers, eastern Puerto Rico.

[--, no samples]

Discharge classes and special sample types	Number of samples	Sediment only	Conductivity only	Conductivity and sediment	Conductivity, chloride, silica	Conductivity, chloride, silica, sediment	Full chemistry	Full chemistry, sediment
Canóvanas								
0–10 percent, rising	26	--	--	2	--	15	--	9
0–10 percent, falling	45	--	--	9	--	19	1	16
>10–25 percent, rising	16	--	--	3	--	5	--	8
>10–25 percent, falling	33	1	--	9	--	13	1	9
>25–50 percent, rising	53	3	3	38	--	9	--	--
>25–50 percent, falling	100	3	3	65	1	26	--	2
>50–75 percent, rising	136	10	11	83	--	28	--	4
>50–75 percent, falling	384	12	22	215	2	110	--	23
>75–90 percent, rising	223	2	2	47	2	146	--	24
>75–90 percent, falling	307	2	5	78	3	189	--	30
>90–95 percent, rising	24	--	--	2	1	13	--	8
>90–95 percent, falling	19	--	--	1	--	12	--	6
>95–100 percent, rising	3	--	--	--	--	1	--	2
>95–100 percent, falling	1	--	--	--	--	--	--	1
High chloride	29	--	--	2	--	11	--	16
High potassium	--	--	--	--	--	--	--	--
Low silica	--	--	--	--	--	--	--	--
No runoff	2	--	--	1	--	1	--	--
Cayaguás								
0–10 percent, rising	60	--	--	14	--	40	1	5
0–10 percent, falling	87	--	--	18	1	56	1	11
>10–25 percent, rising	50	2	--	11	--	32	--	5
>10–25 percent, falling	68	--	--	17	1	40	--	10
>25–50 percent, rising	37	--	--	9	--	24	--	4
>25–50 percent, falling	80	2	--	21	--	50	1	6
>50–75 percent, rising	55	4	--	8	1	32	1	9
>50–75 percent, falling	119	11	--	23	5	69	--	11
>75–90 percent, rising	130	3	2	32	1	88	--	4
>75–90 percent, falling	192	5	--	41	1	136	--	9
>90–95 percent, rising	37	--	1	5	--	26	--	5
>90–95 percent, falling	39	--	1	8	--	27	--	3
>95–100 percent, rising	24	--	--	5	--	5	--	14
>95–100 percent, falling	23	1	--	4	3	10	--	5
High chloride	1	--	--	--	--	1	--	--
High potassium	2	--	--	--	--	1	--	1
Low silica	--	--	--	--	--	--	--	--
No runoff	59	--	--	4	--	53	--	2
Mameyes								
0–10 percent, rising	6	--	3	--	--	2	--	1
0–10 percent, falling	50	--	1	4	--	11	7	27
>10–25 percent, rising	8	--	1	1	--	2	--	4
>10–25 percent, falling	32	--	--	3	1	9	--	19
>25–50 percent, rising	3	--	--	2	--	1	--	--
>25–50 percent, falling	33	2	1	6	--	11	1	12
>50–75 percent, rising	10	--	--	1	1	7	--	1
>50–75 percent, falling	39	2	4	19	--	9	--	5
>75–90 percent, rising	161	1	10	32	6	93	--	19
>75–90 percent, falling	446	--	15	98	17	285	--	31
>90–95 percent, rising	202	1	22	37	23	113	1	5
>90–95 percent, falling	344	4	20	80	52	179	--	9
>95–100 percent, rising	86	1	3	7	8	52	--	15
>95–100 percent, falling	87	2	7	7	9	47	--	15
High chloride	34	--	--	1	14	7	--	12
High potassium	13	--	--	--	9	4	--	--
Low silica	4	--	--	--	--	2	--	2
No runoff	3	--	1	--	--	1	--	1

**Table 6.** Number and types of samples collected in different discharge classes and special sample types in five rivers, eastern Puerto Rico.—Continued

[--, no samples]

Discharge classes and special sample types	Number of samples	Sediment only	Conductivity only	Conductivity and sediment	Conductivity, chloride, silica	Conductivity, chloride, silica, sediment	Full chemistry	Full chemistry, sediment
Icacos								
0 to 10 percent, rising	19	--	--	1	--	9	--	9
0 to 10 percent, falling	55	--	--	5	1	19	1	28
>10 to 25 percent, rising	12	--	--	2	--	9	--	1
>10 to 25 percent, falling	49	1	--	10	--	22	4	12
>25 to 50 percent, rising	12	--	--	2	--	5	1	4
>25 to 50 percent, falling	47	1	--	14	2	16	1	13
>50 to 75 percent, rising	9	1	--	--	--	7	--	1
>50 to 75 percent, falling	22	--	--	6	--	13	--	3
>75 to 90 percent, rising	171	--	1	44	1	112	--	13
>75 to 90 percent, falling	330	1	11	65	2	229	--	22
>90 to 95 percent, rising	222	--	16	53	3	134	--	16
>90 to 95 percent, falling	196	--	5	49	3	126	--	13
>95 to 100 percent, rising	83	--	2	17	3	45	1	15
>95 to 100 percent, falling	67	--	2	13	--	44	1	7
High chloride	62	--	--	--	--	42	--	20
High potassium	24	--	--	1	--	20	--	3
Low silica	89	--	--	1	--	28	1	59
No runoff	32	--	1	2	--	27	--	2
Guabá								
0–10 percent, rising	16	--	--	1	--	13	1	1
0–10 percent, falling	43	--	--	12	--	13	--	18
>10–25 percent, rising	9	--	--	1	--	6	--	2
>10–25 percent, falling	49	--	1	15	1	18	1	13
>25–50 percent, rising	19	--	--	8	--	8	--	3
>25–50 percent, falling	43	--	--	12	--	20	--	11
>50–75 percent, rising	18	--	4	3	--	9	--	2
>50–75 percent, falling	23	--	--	4	--	16	--	3
>75–90 percent, rising	63	--	1	16	--	42	--	4
>75–90 percent, falling	22	1	--	5	--	14	--	2
>90–95 percent, rising	104	--	2	22	--	74	1	5
>90–95 percent, falling	126	--	2	22	--	93	--	9
>95–100 percent, rising	57	--	--	3	--	44	--	10
>95–100 percent, falling	56	--	--	1	--	50	--	5
High chloride	50	--	--	--	--	42	--	8
High potassium	40	--	--	--	--	40	--	--
Low silica	83	--	--	--	--	25	1	57
No runoff	72	--	--	18	--	48	--	6

through 2005, 507 hydrologic events were sampled by automated samplers in five rivers during 263 storms, including all major hurricanes during that period (figs. 2–6). Although the climate of eastern Puerto Rico is weakly seasonal (more rain in May and in August through December) large storms occur throughout the year in response to cold fronts from November to May and tropical disturbances from June to October. In this study, 332 samples that were analyzed for suspended sediment and chemistry were collected at runoff rates greater than 20 millimeters per hour ( $\text{mm h}^{-1}$ ), and some chemical samples were even collected at runoff rates of almost  $100 \text{ mm h}^{-1}$ .

Without comparison with other sampling programs, it is difficult for the reader to assess sampling effort necessitated

by the need to accurately describe storm-related processes at the eastern Puerto Rico WEBB site. For such a comparison, a large data collection assembled by Alexander and others (1996a, 1996b, 1998) was chosen. This collection contains data describing rivers from the U.S. Geological Survey Hydrologic Benchmark Network for 1962 to 1995 and from the U.S. Geological Survey National Stream Quality Accounting Network (NASQAN) for 1973 to 1995. This combined database is referred to here as Water Quality Network–1996 (WQN96). The average record duration per site in the WQN96 database is 18.36 years, comparable to the 15 years of the WEBB dataset. Watersheds from these networks represent a broad cross section of rivers in the

United States, including 8 in Hawaii, 6 in Puerto Rico, and 1 in Guam. Most of the rivers are described by a set of chemical and physical measurements that match those sampled in the Puerto Rico WEBB rivers. The programs that produced the WQN96 database sampled extreme flows in many of the rivers. To compare sampling efforts with the eastern Puerto Rico WEBB program, all sites lacking discharge measurements or having ambiguous drainage areas (lakes, canals, and diversion—total of 24) and all sites without a body of comparable chemical and physical measurements (less than 4 suspended-sediment measurements or less than 10 complete chemical analyses for a total of 20) were eliminated from comparison. A total of 53 rivers were excluded from the comparison, leaving 626 rivers.

When compared with the 626 rivers in the WQN96 database, the eastern Puerto Rico rivers are among those that are wettest and have the greatest runoff. The watersheds with the greatest mean-annual runoff in our study are the Mameyes, Icacos, and Guabá, which are forested. Only three rivers in the WQN96 database had comparable mean-annual runoff (the Queets River and North Fork of the Quinault, both in Washington; and Honolii Stream, Hawaii). Mean annual runoff of the Cayaguás is exceeded by 11 rivers from California, Hawaii, and Washington. Mean annual runoff for the Río Canóvanas is exceeded by 26 other rivers.

A runoff rate of 1 mm h<sup>-1</sup> was set as the dividing line between low- and high-runoff samples in the eastern Puerto Rico WEBB and the WQN96 databases. In the WEBB rivers, 30 to 37 percent of the total annual runoff occurs when runoff rates exceed 1 mm h<sup>-1</sup> (0.2778 cubic meters per second per square kilometer (m<sup>3</sup> s<sup>-1</sup> km<sup>-2</sup>), or 25.41 cubic feet per second per square mile (ft<sup>3</sup> s<sup>-1</sup> mi<sup>-2</sup>). A higher runoff rate might have been chosen, but very few samples in the combined WQN96 database were collected at high runoff rates. Only samples that had enough chemical analyses to calculate bedrock denudation (table 7) or a charge balance (comprehensive chemistry, table 8) were included. For samples with a full charge balance, the WEBB dataset also recorded suspended solids; the WQN96 samples usually did not. The sampling effort for comprehensive chemistry, expressed as samples per river per year (table 8), was about five times as great for eastern Puerto Rico WEBB rivers as compared with the overall combined WQN96 database. For tropical rivers, the WEBB effort was less than three-fold as great. For samples collected at rates less than 1 millimeter per hour (mm h<sup>-1</sup>), there was a comparable effort; however, for runoff rates greater than 1 mm h<sup>-1</sup>, the eastern Puerto Rico effort was 50-fold as great; 543 samples were collected as compared with 149 for the entire combined database or 38 for the tropical rivers in the database. The maximum runoff rates sampled for chemistry in the combined WQN96 database were 12.79 mm h<sup>-1</sup> for Biscuit Brook, New York, and 9.77 mm h<sup>-1</sup> for Honolii Stream, Hawaii. The eastern Puerto Rico database has 125 samples with runoff rates that exceed 20 mm h<sup>-1</sup> and 3 that even exceed 90 mm h<sup>-1</sup>. Basin area might be a factor in this comparison, and in terms of basin

**Table 7.** Number of samples in three river-hydrology datasets that have sufficient data to calculate total denudation rates.

[WQN96, database that combines data of the Hydrologic Benchmark and National Assessment of the Quality of Water Networks (Alexander and others, 1996b); WEBB, Water, Energy, and Biogeochemical Budgets; <, less than; >, more than; mm h<sup>-1</sup>, millimeters per hour]

Runoff rate of sample	Number of rivers in class	Number of samples	Samples per river	Samples per river per year
All rivers in WQN96 with chemistry or sediment data				
Total	679	211,281	311	16.73
All rivers in WQN96 with sufficient data to calculate total denudation rates				
<1 mm h <sup>-1</sup>	548	30,187	55	2.96
>1 mm h <sup>-1</sup>	43	96	2	0.12
>10–20 mm h <sup>-1</sup>	0	0	–	–
>20 mm h <sup>-1</sup>	0	0	–	–
Total	548	30,283	55	2.97
All tropical rivers in WQN96 with sufficient data to calculate total denudation rates				
<1 mm h <sup>-1</sup>	15	980	65	3.51
>1 mm h <sup>-1</sup>	11	35	3	0.17
>10–20 mm h <sup>-1</sup>	0	0	–	–
>20 mm h <sup>-1</sup>	0	0	–	–
Total	15	1,015	68	3.64
Eastern Puerto Rico WEBB samples with sufficient data to calculate total denudation rates				
<1 mm h <sup>-1</sup>	5	922	184	13
>1 mm h <sup>-1</sup>	5	3,398	680	49
>10–20 mm h <sup>-1</sup>	5	932	186	13
>20–30 mm h <sup>-1</sup>	5	303	61	4.3
>30–40 mm h <sup>-1</sup>	5	124	25	1.8
>40–50 mm h <sup>-1</sup>	5	53	11	0.76
>50–60 mm h <sup>-1</sup>	3	32	11	0.76
>60–70 mm h <sup>-1</sup>	3	20	7	0.48
>70–80 mm h <sup>-1</sup>	2	11	6	0.39
>80–90 mm h <sup>-1</sup>	1	5	5	0.36
>90 mm h <sup>-1</sup>	1	3	3	0.21
Total	5	4,320	864	62

area, the WQN96 rivers overlap all the WEBB rivers, except the Guabá which has a 0.114-km<sup>2</sup> watershed at the stream-gaging station.

Globally, few other event samples of streams have been collected and analyzed for sediment and chemistry at runoff rates of greater than 5 mm h<sup>-1</sup>. In the tropics, two samples were collected at about 16 mm h<sup>-1</sup> by the U.S. Geological Survey in Puerto Rico in 1970 during a tropical depression (Haire, 1970). Goldsmith and others (2008) collected samples in Taiwan during Typhoon Mindulle at runoff rates of as much as 7.3 mm h<sup>-1</sup>, and Hicks and others (2000) collected samples in New Zealand during Cyclone Bola at runoff rates of as much as 4.8 mm h<sup>-1</sup>.

The eastern Puerto Rico WEBB sampling program represents the opportunity to assess how physical and biogeochemical processes affect small watersheds at extremely high runoff conditions. Because of the extensive sampling of

**Table 8.** Number of samples in three river-hydrology datasets that have comprehensive chemistry.

[WQN96, combined dataset of the Hydrologic Benchmark and National Assessment of the Quality of Water Networks (Alexander and others, 1996b); WEBB, Water, Energy, and Biogeochemical Budgets program; <, less than; >, greater than; mm h<sup>-1</sup>, millimeters per hour]

Runoff rate of sample	Number of rivers in class	Number of samples	Samples per river	Samples per river per year
All rivers in WQN96 with chemistry or sediment data				
All	679	211,281	311	16.73
All rivers in WQN96 with runoff and comprehensive chemistry				
<1 mm h <sup>-1</sup>	548	26,568	48	2.61
>1 mm h <sup>-1</sup>	54	149	3	0.15
>10–20 mm h <sup>-1</sup>	1	1	1	0.05
>20 mm h <sup>-1</sup>	0	0	–	–
All	548	26,717	49	2.62
All tropical rivers in WQN96 with runoff and comprehensive chemistry				
<1 mm h <sup>-1</sup>	15	1,167	78	4.18
>1 mm h <sup>-1</sup>	13	38	3	0.16
>10–20 mm h <sup>-1</sup>	0	0	–	–
>20 mm h <sup>-1</sup>	0	0	–	–
All	15	1,205	80	4.32
Eastern Puerto Rico WEBB samples with runoff and comprehensive chemistry				
<1 mm h <sup>-1</sup>	5	314	63	4.5
>1 mm h <sup>-1</sup>	5	543	109	7.8
>10–20 mm h <sup>-1</sup>	5	256	51	3.7
>20–30 mm h <sup>-1</sup>	5	125	25	1.8
>30–40 mm h <sup>-1</sup>	5	56	11	0.80
>40–50 mm h <sup>-1</sup>	5	24	5	0.34
>50–60 mm h <sup>-1</sup>	3	18	6	0.43
>60–70 mm h <sup>-1</sup>	3	9	3	0.21
>70–80 mm h <sup>-1</sup>	2	6	3	0.21
>80–90 mm h <sup>-1</sup>	1	3	3	0.21
>90 mm h <sup>-1</sup>	1	3	3	0.21
All	5	857	171	12

high-runoff-rate events, these data are unique: they represent a seldom-examined realm of process-related water-quality studies. The extensive sampling for suspended sediment, combined with the use of surrogate parameters to calculate dissolved bedrock, means that the WEBB database is much better poised than most other long-term studies for evaluating denudation processes (table 7).

## References

- Alexander, R.B., Slack, J.R., Ludtke, A.S., Fitzgerald, K.K., and Shertz, T.L., 1996a, Data from selected U.S. Geological Survey national stream water quality monitoring networks (WQN): U.S. Geological Survey Open-File Report 96–337, 79 p.; CD-ROM.
- Alexander, R.B., Slack, J.R., Ludtke, A.S., Fitzgerald, K.K., and Shertz, T.L., 1996b, Data from selected U.S. Geological Survey national stream water quality monitoring networks: U.S. Geological Survey Digital Data Series DDS–37, 2 CD-ROMs.
- Alexander, R.B., Slack, J.R., Ludtke, A.S., Fitzgerald, K.K., and Shertz, T.L., 1998, Data from selected U.S. Geological Survey national stream water quality monitoring networks: Water Resources Research, v. 34, p. 2401–2405.
- Bevington, P.R., and Robinson, K.D., 2003, Data reduction and error analysis for the physical sciences—2: New York, McGraw Hill, 320 p.
- Chapman, T.B., 1999, A comparison of algorithms for stream flow recession and baseflow separation: Hydrological Processes, v. 13, p. 701–714.
- Díaz, P.L., Aquino, Z., Figueroa-Alamo, C., Garcia, R., and Sanchez, A.V., 2004, Water resources data, Puerto Rico and the U.S. Virgin Islands, water year 2002: U.S. Geological Survey Water-Data Report PR–02–1, 652 p.
- Elmore, H.L., and Hayes, T.W., 1960, Solubility of atmospheric oxygen in water: Journal of Sanitary Engineering Division, American Society of Civil Engineers, v. 86, SA4, p. 41–53.
- Geyer, Tobias, Birk, S., Liedl, R., and Sauter, M., 2008, Quantification of temporal distribution of recharge in karst systems from spring hydrographs: Journal of Hydrology, v. 348, p. 452–463.
- Goldsmith, S.T., Carey, A.E., Lyons, W.B., Kao, S.-J., Lee, T.-Y., and Chen, J., 2008, Extreme storm events, landscape denudation, and carbon sequestration: Typhoon Mindulle, Choshui River, Taiwan: Geology, v. 36, no. 6, p. 483–486.
- Haire, W.J., 1972, Flood of October 5–10, 1970, in Puerto Rico: Commonwealth of Puerto Rico Water-Resources Bulletin 12, 42 p.
- Hall, F.R., 1968, Base-flow recessions—A review: Water Resources Research, v. 4, no. 5, p. 973–983.
- Hicks, D.M., Gomez, B., and Trustrum, M.A., 2000, Erosion thresholds and suspended sediment yields, Waipaoa River Basin, New Zealand: Water Resources Research, v. 36, no. 4, p. 1129–1142.

- Jupp, D.L., 1976, B-splines for smoothing and differentiating data sequences: *Mathematical Geology*, v. 8, no. 3, p. 243–266.
- Larsen, M.C., 2012, Landslides and sediment budgets in four watersheds in eastern Puerto Rico, ch. F in Murphy, S.F., and Stallard, R.F., eds., *Water quality and landscape processes of four watersheds in eastern Puerto Rico*: U.S. Geological Survey Professional Paper 1789, p. 153–178.
- Larsen, M.C., and Simon, A., 1993, A rainfall intensity-duration threshold for landslides in a humid-tropical environment, Puerto Rico: *Geografiska Annaler*, v. 75A, no. 1–2, p. 13–23.
- Mackenzie, R.C., 1957, Differential thermal analysis of clays: *The Mineralogical Society [London]*, 389 p.
- Mackenzie, R.C., and Caillère, S., 1979, Thermal analysis, DTA, TG, DTG, in Van Olphen, Hendrik, and Fripiat, J.J., eds., *Data handbook for clay material and other nonmetallic Minerals*: New York, Pergamon, p. 243–284.
- Miller, R.L., Bradford, W.L., and Peters, N.E., 1988, Specific conductance—Theoretical considerations and application to analytical quality control: U.S. Geological Survey Water-Supply Paper 2311, 16 p.
- Murphy, S.F., and Stallard, R.F., 2012, Hydrology and climate of four watersheds in eastern Puerto Rico, ch. C in Murphy, S.F., and Stallard, R.F., eds., *Water quality and landscape processes of four watersheds in eastern Puerto Rico*: U.S. Geological Survey Professional Paper 1789, p. 43–84.
- Reinsch, C.H., 1967, Smoothing by spline functions: *Numerische Mathematik*, v. 10, p. 177–183.
- Runkel, R.L., Crawford, C.G., and Cohn, T.A., 2004, Load estimator (LOADEST): A FORTRAN program for estimating constituent loads in streams and rivers: U.S. Geological Survey Techniques and Methods Book 4, 69 p., program and test files.
- Stallard, R.F., 1995a, Relating chemical and physical erosion, in White, A.F., and Brantley, S.L., eds., *Chemical weathering rates of silicate minerals*: Washington, D.C., Mineralogical Society of America Reviews in Mineralogy, v. 31, p. 543–564.
- Stallard, R.F., 1995b, Tectonic, environmental, and human aspects of weathering and erosion—A global review using a steady-state perspective: *Annual Review of Earth and Planetary Sciences*, v. 12, p. 11–39.
- Stallard, R.F., 1998, Terrestrial sedimentation and the carbon cycle—Coupling weathering and erosion to carbon burial: *Global Biogeochemical Cycles*, v. 12, no. 2, p. 231–252.
- Stallard, R.F., 1999, Erosion and the effects of deforestation in the Panama Banal Basin, in *Panama Canal Watershed Monitoring Project: Report of the Panama Canal Watershed Monitoring Project*, ch. II.8, 8 volumes, 21 CD-ROMs.
- Stallard, R.F., 2012a, Atmospheric inputs to watersheds of the Luquillo Mountains of eastern Puerto Rico: ch. D in Murphy, S.F., and Stallard, R.F., eds., *Water quality and landscape processes of four watersheds in eastern Puerto Rico*: U.S. Geological Survey Professional Paper 1789, p. 85–112.
- Stallard, R.F., 2012b, Weathering, landscape equilibrium, and carbon in four watersheds in eastern Puerto Rico, ch. H in Murphy, S.F., and Stallard, R.F., eds., *Water quality and landscape processes of four watersheds in eastern Puerto Rico*: U.S. Geological Survey Professional Paper 1789, p. 199–248.
- Stallard, R.F., and Edmond, J.M., 1981, Geochemistry of the Amazon—I. Precipitation chemistry and the marine contribution to the dissolved load at the time of peak discharge: *Journal of Geophysical Research—Oceans and Atmospheres*, v. 86, p. 9844–9858.
- Stallard, R.F., and Kinner, D.A., 2005, Estimation of landslide importance in hillslope erosion within the Panama Canal watershed, in Harmon, R.S., ed., *The Río Chagres, Panama—A multidisciplinary profile of a tropical watershed*: Dordrecht, The Netherlands, Springer Water Science and Technology Library, v. 52, p. 281–295.
- Stallard, R.F., and Murphy, S.F., 2012, Water quality and mass transport in four watersheds in eastern Puerto Rico, ch. E in Murphy, S.F., and Stallard, R.F., eds., *Water quality and mass transport in research watersheds in eastern Puerto Rico*: U.S. Geological Survey Professional Paper 1789, p. 113–152.
- Stumm, Werner, and Morgan, J.J., 1981 *Aquatic chemistry*: New York, John Wiley, 780 p.
- Tallaksen, L.M., 1995, A review of baseflow recession analysis: *Journal of Hydrology*, v. 165, p. 349–370.
- Tardy, Yves, Bustillo, V., Roquin, C., Mortatti, J., and Reynaldo, R., 2005, The Amazon—Bio-geochemistry applied to river basin management, pt. I. Hydro-climatology, hydrograph separation, mass transfer balances, stable isotopes, and modelling: *Applied Geochemistry*, v. 20, p. 1746–1829.
- Vogel, R.M., and Kroll, C.N., 1996, Estimation of baseflow recession constants: *Water Resources Management*, v. 10, p. 303–320.
- Wolman, M.G., and Miller, J.P., 1960, Magnitude and frequency of forces in geomorphic processes: *Journal of Geology*, v. 68, p. 54–74.