Use of Total Suspended Sediment Data in Watershed Model Calibration

John Doherty[1*] and John M. Johnston[2]

[1]Department of Civil Engineering, University of Queensland, St Lucia, Australia 4068,

Watermark Numerical Computing, Brisbane, Australia.

[2] USEPA Office of Research and Development, Ecosystems Research Division, 960

College Station Rd., Athens, GA, USA 30605

Key Words: modelling, suspended sediment, HSPF, PEST, model calibration, uncertainty

* Contact author: phone int + 61 7 3379 1664, email: jdoherty@gil.com.au .

Abstract

The incorporation of suspended sediment data into the calibration process of a watershed model is often fraught with difficulty. Difficulties arise from the often "noisy" nature of such data and the fact that TSS measurements are often made only sporadically. Furthermore they are often made during periods of low flow when least sediment is carried by a stream. As a result of this, rarely is there sufficient data available to allow unique estimation of the parameters associated with the erosion and sediment transport functionality of such models. To make matters worse, in many cases only limited assistance can be provided to the parameterization process through the assignment of values to model parameters based on knowledge of the physical processes which they represent, due to the "lumped" nature of these processes when simulated by the model.

The study documented herein examines some of these problems in the context of calibrating a HSPF model of a watershed. It demonstrates the use of nonlinear parameter estimation techniques (using the PEST software) in allowing TSS data to be incorporated into the model calibration process, in spite of the difficulties surrounding the use of this data. Recognizing that no parameter set estimated through the calibration process is unique, the focus is then shifted to exploring model predictive uncertainty arising from parameter uncertainty using PEST's predictive analyzer.

Introduction

The use of advanced nonlinear parameter estimation techniques in the calibration and predictive analysis of watershed models is documented in Doherty and Johnston (2002). In the work described in that paper, use of these techniques was restricted solely to the processing of streamflow data, and to the estimation of parameters which govern the model's hydrologic outputs.

The present paper documents the use of nonlinear parameter estimation methods in estimating parameters associated with the erosion and sediment transport components of a watershed model. The sporadic and noisy nature of sediment data makes the estimation of these parameters a much more difficult procedure than the estimation of hydrologic parameters. This difficulty is exacerbated by the insensitivity of model outputs to some of these parameters over at least part of their allowable range, and the sometimes extremely nonlinear nature of the relationship between these parameters and model outputs. Parameter correlation is also a problem, this term describing the fact that it is often possible to vary two or more parameters in harmony with very little effect on model outputs, this making estimation of the individual parameters virtually impossible.

As a result of low parameter sensitivity and high parameter correlation, a great deal of nonuniqueness is often associated with parameters estimated through the calibration process, even after "reality checks" have been placed on their values through the application of "outside knowledge" based on an understanding of the physics or chemistry of the processes being simulated. Uncertainty in the estimated values of model parameters can then lead to uncertainty in the values of predictions made by the model.

This, in turn, leads to the necessity to analyze the uncertainty associated with model predictions. This paper also addresses the issue of model predictive uncertainty analysis in the erosion and sediment transport context.

The Watershed

The Contentnea Creek Basin, a coastal plain watershed, is located in the Neuse River Basin in North Carolina (Figure 1). Rainfall in the area averages 127 cm per year (Giese et al, 1997). The mean annual maximum temperature is approximately 10 degC, while the mean monthly minimum temperature is 30 degC. The physiography is relatively uniform throughout the basin, with relatively low relief. The soils are well-drained sands and sandy loams developed on sediments of marine origin. The primary land covers within the basin are forest, agriculture, grassland and urban, with the first two land use types accounting for nearly 70% of the area of the basin.

As is described in Doherty and Johnston (2002) four models were built for four neighboring watersheds situated within this basin. These models were built as part of a study dedicated to predicting alterations to water quality within the Contentnea Creek Basin as a result of increasing urbanization, altered farming practices and climatic change. The present investigation focuses on one of these models, viz. the model built to simulate hydrologic and sediment processes in the Contentnea Creek above Hookerton; this is the basin labeled "Contentnea in Figure 1, but the model will henceforth be referred to as "the Hookerton model" in order to retain the same terminology as that employed by Doherty and Johnston. The area of this watershed is about 100000 acres.

Software

Simulation of hydrologic and sediment erosion/transport processes was undertaken using version 12 of HSPF (Bicknell et al, 2001). The watershed was simulated using four HSPF PERLNDs, one IMPLND and a RCHRES (a PERLND is a pervious land segment, an IMPLND is an impervious land segment and a RCHRES is a free-flowing reach or mixed reservoir). The four PERLNDs were used to represent the four major land use types mentioned above. The IMPLND was used for the simulation of urban impervious areas (this comprising less than 2% of the total area of the watershed). The RCHRES was used to simulate flow of water in the creek system draining the watershed.

Model calibration was undertaken using PEST (Doherty, 2001a) in conjunction with a suite of utility software written to support the use of PEST in the surface water modeling context (Doherty, 2001b); the principal member of this suite is TSPROC, a time-series processor optimized for use in the calibration context. PEST is a model-independent parameter estimator with advanced predictive analysis and regularization features. Its model-independence rests on the fact that it is able to communicate with a model through the latter's own input and output files, thus allowing easy calibration setup with an arbitrary model. Such a "model" can be encapsulated in a batch or script file if desired. Hence model pre-and postprocessing software (such as TSPROC) can be used as part of the calibration process.

PEST implements a particularly robust variant of the Gauss-Marquardt-Levenberg method of parameter estimation. While this method requires that a continuous relationship exist between model parameters and model outputs, it can normally find the minimum of the objective function in fewer model runs than any other parameter estimation method. This is important where model run times are lengthy, or even moderate.

TSPROC is able to read time-series data from a variety of sources including ASCII files and USGS Watershed Data Management (ie. WDM) files. It can undertake temporal interpolation of one time series to another, carry out mathematical manipulations of arbitrary complexity between one or more time series, compute time series statistics, and calculate various quantities derived from time series including exceedence times, and volumetric/mass accumulation between one or many arbitrary dates and times. It also facilitates the use of both raw and processed time series data in the calibration process by automatically generating PEST input files for calibration runs involving some or all of these quantities.

Use of PEST and TSPROC in calibrating the hydrologic component of the Hookerton model (and its three neighboring watershed models) is fully documented in Doherty and Johnston.

The Data

Total suspended sediment (TSS) samples have been collected at irregular intervals at the Hookerton Gauging Station since 1975. In the present study, data gathered after the end of 1995 was ignored so that the time period used for calibration of the sediment component of the model would coincide with that used for calibration of the hydrologic component of the model (see Doherty and Johnston). Unfortunately, partitioning of TSS samples into sediment size classes was not undertaken; hence only total sediment data was available for use in model calibration.

Figure 2 shows the TSS data plotted to linear and logarithmic scales. In Figure 3 TSS data are compared with flow data. In the first part of this latter figure TSS measurements are superimposed on flow measurements. A close inspection of this figure reveals that while there were some occasions when TSS readings appear to have been made during periods of high flow in order to “capture” the sediment characteristics of the stream during periods of high load, many of the TSS measurements were taken during periods of comparatively low flow. Thus the dataset taken as a whole does not provide a very good basis for model calibration during those periods when erosion and sediment movement are most active. Unfortunately, this is true of many suspended sediment datasets.

The lower part of Figure 3 depicts the so-called “sediment rating curve”, ie. a plot of TSS against stream discharge. The increase of TSS with flow rate is readily apparent from this figure. However it is also apparent that there would probably be a high degree of scatter about any regression line that was fitted to this data using software such as the USGS program ESTIMATOR (see for example Cohn et al., 1989 and Cohn and Gilroy, 1991).

Where there is a strong correlation between stream discharge and sediment load, the sediment rating curve can be a powerful tool for the analysis of stream sediment transport. Its most significant contribution to analyses of this type is that it allows discharge to act as a surrogate for sediment load over those periods for which TSS measurements are not available (which, in most cases, is most of the time). Thus, for example, if a rating curve can be determined with sufficient accuracy, the total sediment transported from a watershed over a given period of time can be evaluated by first calculating daily sediment concentrations from daily stream discharges using the sediment rating curve, and then summing daily sediment concentrations times daily flows over the period of interest. If it is further assumed that the amount of sediment in the bed of a stream is the same at the end of the period as it was at the beginning of the period, then this total transported sediment can be equated to the total amount of sediment eroded or washed from the watershed over that time period; this, in turn, can lead to an estimate of long-term erosion rate. However while such a calculation is conceptually possible, in practice there is a considerable amount of uncertainty associated with calculations of this type. One contributor to this uncertainty is the uncertainty associated with parameters which describe the sediment rating curve. A second source of uncertainty is the assumption that streambed sediment storage does not change over the analysis period, for this may often be very difficult to verify.

An alternative means of calculating the total amount of sediment exported from a watershed is that afforded by the use of a calibrated model. Use of a model has the advantage that it can be applied to those sediment types (viz. silt and clay) to which the sediment rating curve concept is not as directly applicable as it is for sediment types such as sand – see below. Nevertheless, as will be discussed below, estimates of sediment export made using a model will also be subject to a large amount of uncertainty. An attempt to quantify this uncertainty will be made. It is perhaps this ability to at least partially quantify the degree of predictive uncertainty associated with calculations of these and other types made by a model, which makes the use of a model, rather than a more empirical relationship, more attractive in many environmental data processing contexts.

HSPF and In-Stream Sediment Transport

In HSPF sediment eroded from a PERLND is directed to a RCHRES. There, the delivery of sediment downstream (or to storage within the bed of a stream) is simulated using the SEDTRN group of the RCHRES block. No attempt is made herein to evaluate the erosion and sediment transport algorithms employed by this group. Nevertheless a few comments will be made on those aspects of its algorithm that have a bearing on the present investigation.

Sand

The amount of sand in suspension in a flowing stream is calculated by HSPF in a different manner to that in which the amount of suspended silt and clay is calculated. Three options are provided by HSPF for suspended sand calculation, viz. , the Toffaleti equation, the Colby method, and the power function method. In all of these cases HSPF first calculates the *potential* suspended sand concentration based on the velocity of the stream. If the current suspended sand concentration exceeds this potential, sand is deposited; if it is less than this potential, sand is scoured from the bed of the stream (if bed sand is available). In the present study the power function method was employed, though in a slightly modified form, as is explained below.

HSPF's characterization of suspended sand concentration has a number of important repercussions for the calibration of a HSPF model on the basis of in-stream sediment data. The first of these is that as long as there is sand in the bed of a stream, no direct inference can be drawn from the amount of suspended sand in the stream about the erosional characteristics of any PERLND or IMPLND which feeds the stream (except in the very long term – see below). This is because the amount of suspended sand is a function purely of the velocity (and hence current discharge rate) of the stream. Any sand that is delivered to a stream that is in excess of its sand-carrying capacity will be immediately deposited on its bed. Similarly, any shortfall in the amount of sand carried by the stream will be immediately filled by streambed sand. Hence, under these circumstances, measurements of suspended sand can only provide information pertaining to the estimation of those parameters which govern the relationship between stream discharge and stream sediment carrying capacity. That is, the calibration process can only be used to infer the sediment rating curve (or rather the "sand rating curve") of the stream.

In contrast, if there is no sand in the bed of a stream, any suspended sand carried by the stream will be the direct result of erosion taking place within the PERLNDs and IMPLNDs which feed the stream. Thus, in these circumstances, measurements of suspended sand concentration can provide some information by which estimation of parameters governing watershed erosion can take place. However these conditions are more likely to prevail in upland watersheds drained by young streams, than in lowlands drained by more mature streams.

Silt and Clay

Suspended transport of silt and clay is simulated by HSPF using a different algorithm from that which is used for the simulation of sand transport. For these finer sediments no carrying capacity is defined. Instead sediment can be scoured from the streambed bottom if the shear stress exceeds the "critical shear stress for scouring" (HSPF parameter TAUCS), and deposited if the shear stress is less than the "critical shear stress for deposition" (HSPF parameter TAUCD). Shear stress is calculated from a number of internal quantities that depend on stream discharge, slope and geometry.

Sediment eroded from PERLNDs and IMPLNDs which drain to a RCHRES is placed directly into the RCHRES's suspended storage. It is then deposited at a rate determined by the sediment's settling velocity if the shear stress is below TAUCD. Alternatively, especially during periods of high flow when the RCHRES receives most of its suspended sediment, the sediment can be quickly transported from the system. Because of this, and the fact that there is no means available to achieve an "equilibrium sediment level" at any flow rate (as is assumed when using the sediment rating curve concept), HSPF's simulation of suspended silt and clay normally results in large variations of these quantities over short periods of time. Suspended silt and clay concentration quickly rises when flows are high, scouring is active and sediment inflow is high; it then quickly dies away as suspended sediment settles, or is transported from the system.

Normally TAUCS is set above TAUCD. Hence there is a "dead area" when neither deposition nor scouring take place. If it happens that streamflow is within this dead area much of the time, then these parameters become insensitive when an attempt is made to estimate them through the calibration process. The possibility of this inadvertently occurring is heightened by the fact that unless a user specifically requests that HSPF generate a shear stress output time series, he/she has no knowledge of this important quantity relative to his/her TAUCS and TAUCD settings. Furthermore, the absence of scouring is disguised by the fact that silt and clay can enter the system during periods of high flow through erosion and washoff of connected PERLNDs and IMPLNDs. Under these circumstances TAUCS and its associated parameter M (erodability coefficient) can become totally insensitive. With little difficulty TAUCD and W (settling velocity) can meet the same fate, for transportation of suspended sediment out of the system, rather than deposition, can reduce the impact of deposition on suspended sediment concentration to almost zero.

Thus the estimation of RCHRES silt and clay transport parameters can be extremely difficult because of their high correlation with PERLND/IMPLND erosion parameters, and because the sensitivity of these parameters is so highly dependent on their values (and the values of other parameters). In some instances these problems can be overcome by supplying values for these parameters "from outside" of the calibration process. However if a value thus supplied results in rapid scouring or deposition of streambed sand/clay (as can easily happen), then there is no alternative but to adjust its value during the calibration process.

Similar considerations apply to the amount of sand and silt stored in bed sediments as those that were discussed above with respect to sand bed storage. That is, it is not too difficult to find a set of parameters which cause HSPF to winnow all silt and clay from the bed of a stream in a short amount of time, or add an unacceptable amount of silt/clay to these sediments during and shortly after large rainfall events. One way to prevent this from happening is to set TAUCD very low and TAUCS very high so that virtually no interaction between the stream and its bed takes place. While this has the benefit of ensuring that bed silt/clay storage remains unchanged during the calibration time, and thus that measurements of stream silt/clay loads can be used to infer PERLND/IMPLND sediment supply parameters, it may not result in a realistic simulation of system behavior. However the alternative option of trying to estimate a set of parameters which ensure that, in spite of the often rapid buildup and decay of suspended silt and clay during a rain event, the amount of bed clay and silt is approximately the same at the end of a long simulation period as it was at the beginning, can be a difficult matter, especially if calibrating a model by hand.

Lumped Parameters

The algorithms used by HSPF to compute suspended sediment concentration for both the sand and silt/clay fractions rely on the calculation of in-stream variables such as shear stress and stream velocity. Calculation of such quantities depends, to some extent, on the geometry of the stream as supplied through various HSPF parameters, including the RCHRES FTABLE which provides the relationship between discharge, surface area, depth and volume. The information supplied in this table is necessarily "lumped", and hence does not represent exact stream characteristics. Hence quantities derived from this and other parameters that are used in the calculation of sediment transport are also "lumped", applying more to the RCHRES as an abstraction of reality than to any particular part of the stream system which the RCHRES represents.

While the use of algorithms which are based on the physics and chemistry of environmental processes in lumped-parameter models such as HSPF can allow them to better simulate real-world system behavior, caution must nevertheless be exercised when providing values for the parameters on which they rely. For example, the shear stress calculated by HSPF for a RCHRES will exist nowhere in particular within a stream reach due to the variable geometry of the reach along its length (and probably other reasons as well). Furthermore, it is very easy to supply a set of FTABLE entries which "look right" but which can surprise the user if he/she inspects HSPF-calculated quantities which are based on the entries of this FTABLE such as stream depth and shear stress, a problem that is exacerbated by the piecewise linear nature of the FTABLE. In fact HSPF calculates a constant river velocity over the entire first segment of an FTABLE, this being an artifact of the piecewise linear nature of this data structure. This in turn, has serious repercussions for the calculation of suspended sediment concentrations – both for sand and for silt/clay. It follows therefore, that parameter values supplied "from outside" the parameter estimation process, based on the physics of sediment scour, transport and deposition, may not result in a good replication of historical suspended sediment measurements by the model. Hence model parameters must be estimated at least partly through the calibration process. Unfortunately, as has already been discussed, such estimates (especially those pertaining to sediment transport) are prone to a high degree of nonuniqueness. This, in turn, may lead to a large amount of predictive nonuniqueness, a matter which will shortly be addressed.

A Small Alteration to HSPF

A small alteration was made to the algorithm which describes sand transport in a HSPF RCHRES. In the "power function" option, potential sand carrying capacity (PSAND) of the stream is calculated using the equation:

PSAND = KSAND*AVVELE**EXPSND (1)

where AVVELE is the average streambed velocity over the RCHRES during a particular time step, and KSAND and EXPSND are parameters to be determined during the calibration process. This was replaced by the following equation:-

PSAND = KSAND*ROM**EXPSND (2)

in which ROM is the average stream discharge over the time step. Use of this second equation has the advantage that it eliminates the "constant velocity problem" over the first FTABLE segment mentioned above. It also resembles the sediment rating curve description of stream sand content. Perhaps higher order terms should have been used in the relationship between PSAND and ROM in accordance with similar terms commonly employed in sediment rating curves; see, for example, Cohn et al (1989). However use of the above simple equation required minimal alterations to HSPF as no new parameters were required; furthermore, an inspection of Figure 3 suggests that it is unlikely that the scatter about the rating curve of best fit would be substantially reduced by the introduction of higher order terms.

Calibration Procedure

Estimated Parameters

As was discussed above, the Hookerton Model is comprised of 5 PERLNDs and an IMPLND linked to a single RCHRES. In order to reduce the number of parameters requiring estimation, all four PERLNDs were initially assigned the same hydrologic parameters (these parameters belong to the PWATER group of the HSPF PERLND module), except for the FOREST parameter which governs the amount of evapotranspiration taking place during winter. However parameters related to the dimensions of the system (for example land use areas, lengths of overland flow paths, average slopes etc) were assigned in accordance with watershed geometry and topography. PWATER parameters estimated for the PERLNDs through the calibration process are listed in Table 1. Note that the values for IMPLND parameters were assumed rather than estimated; this did not degrade the calibration process due to the small size of the IMPLND relative to the PERLNDs. See Doherty and Johnston (2002) for full details of the calibration process.

A similar strategy was adopted for the estimation of PERLND sediment parameters (ie. parameters belonging to the SEDMNT group) to that which was employed for the estimation of PERLND hydrologic parameters; that is, PERLNDs were collected into groups for the purposes of parameter assignment. SEDMNT parameters deemed as "adjustable" are listed in Table 2, along with a brief description of the role of each. During the calibration process KRER, JRER, JSER and JGER were assumed to be the same for all PERLNDS, except for the forest PERLND where KRER was assumed to be zero. KSER and KGER were assumed to be the same for agricultural and urban PERLNDs, and to be a fifth of this (for KSER) and a fourth of this (for KGER) for the grassland and forest PERLNDs; note that it is the agricultural/urban parameter values that are reported with the calibration results in Table 4 – see below. Once again, sediment parameters for the IMPLND were simply assigned, and hence not estimated through the calibration process.

Table 3 lists the RCHRES transport parameters estimated through the calibration process; these all belong to the SEDTRN group. The first two of these parameters (viz. KSAND and EXPSAND) pertain to the transport of suspended sand. As is discussed above, they do not have their usual HSPF role as they now pertain to equation (2) above rather than to equation (1). In order to reduce the number of parameters requiring estimation, M for clay was assumed to be equal to M for silt while TAUCD for clay was assumed to be 0.8 times that of silt. However TAUCS for clay was estimated separately from that of silt. To ensure that TAUCS is always greater than TAUCD, the ratio of these two parameters (named TAUCRAT) was actually estimated, instead of TAUCS itself; for each sediment type a lower bound of 1 was placed on this ratio.

Thus a total of 13 parameters pertaining to watershed sediment erosion and transport were declared as adjustable. Obviously unique estimation of all of these parameters on the basis of the limited dataset displayed in Figures 2 and 3 is impossible, even with the help of "outside knowledge" of these parameters based on the mechanics of the operative processes. As is discussed in Doherty and Johnston, if environmental models are to be used correctly, the idea that a unique parameter set can be estimated in many modeling contexts must be abandoned. Instead, it must be realized that the calibration process can do no more than impose a set of complex constraints on parameter values, the purpose of these constraints being to ensure that parameters that are used to make model predictions are such that the model is able to replicate historical system behavior as well as possible using these same parameters. As parameter nonuniqueness will probably lead to predictive nonuniqueness, assessment of the uncertainty associated with model predictions, arising from the uncertainty with which model parameters can be estimated, then becomes an important part of the model deployment process. This is the approach taken herein.

The Observation Dataset

In carrying out the parameter estimation process, PEST minimizes an "objective function" comprised of the sum of squared weighted deviations (ie. residuals) between model outputs and corresponding field measurements; see Doherty (2001a) for more details. When estimating the parameters listed in Tables 2 and 3, the parameter estimation problem was set up in such a way that three types of "observations" contributed to the objective function. These are now discussed in detail.

TSS Measurements

The 103 TSS measurements illustrated in Figures 2 and 3 comprised one subgroup of the observation dataset used in the parameter estimation process. The weight assigned to each of these measurement was calculated as the inverse of the measurement itself, thereby preventing the handful of very large TSS measurements from dominating the inversion process. Suspended sediment concentrations calculated by HSPF were time-interpolated to measurement dates and times (using the TSPROC postprocessor described above) in order to allow a direct comparison to be made between field TSS measurements and their model-generated counterparts.

TSS Statistics

Whether a model such as the present one is calibrated by hand, or whether it is calibrated with the help of nonlinear parameter estimation software such as PEST, it cannot be expected that a set of parameters can be derived which will result in a good fit between each individual TSS measurement and its model-generated counterpart. Often, the best that can be hoped for is the estimation of a set of parameters which reproduce the *character* of a TSS dataset, if not the details. To facilitate PEST's search for such a set of parameters, two "statistical observations" were included in the observation dataset used by the parameter estimation process, these being the mean and standard deviation of the TSS measurements. The model outputs corresponding to these measurement statistics were calculated on the basis of model-generated sediment concentrations *time-interpolated to the dates and times of sediment observations*; that is, they were calculated on the basis of the 103 model-generated counterparts to field TSS measurements. As such, they allow a direct comparison to be made between two aspects of the "character" of the respective TSS datasets, with the modeled dataset undergoing a "selection process" identical to that to which field TSS dataset was subjected.

In formulating the objective function (whose task it is for PEST to minimize), the mean and standard deviation "observations" were assigned equal weights. These weights were chosen such that, at the beginning of the parameter estimation process (where the model uses initial parameter values selected by the user) the contribution made to the overall objective function by the residuals pertaining to these two observations together was equal to the contribution made to the objective function by all of the TSS residuals. This strategy ensured that neither the statistical observations nor the native TSS observations dominated the parameter estimation process. Thus PEST was able to take both of these observation types into account, reducing the residuals associated with each of them if possible when upgrading parameter values.

RCHRES Bed Composition

Three extra "observations" were included in the calibration dataset, all of which were provided with a "measured" value of zero. The first was the difference between the amount of sand in the bed of the RCHRES at the beginning of the calibration period and that at the end of the calibration period. The second and third "observations" pertained to similar differences taken for silt and clay. Inclusion of these as components of the calibration dataset prevented the occurrence of large amounts of scouring or deposition by the model over the calibration period, this being in accord with direct observations of the condition of the watershed. Each of these "observations" was provided with the same weight. The weight was such that the contribution made to the overall objective function by the residuals associated with these three "bed sediment difference" observations was roughly the same as that contributed by native TSS data on the one hand, and the statistics pertaining to TSS data on the other hand, at the commencement of the parameter estimation process.

Outcomes of the Calibration Process

Simultaneous Calibration against Head and Flow

An attempt was made to estimate both flow and transport parameters as part of the same calibration process by including discharges (and postprocessed discharges as discussed by Doherty and Johnston), as well as TSS measurements (and postprocessed TSS measurements as discussed above) in the calibration dataset, and estimating all of the parameters listed in Tables 1, 2 and 3 simultaneously. As is documented in Doherty and Johnson, calibration of the hydrologic parameters listed in Table 1 against a single discharge time series leads to nonunique estimates of these parameters. Joint estimation of flow and sediment parameters on the basis of both flow and discharge data was undertaken to test whether inclusion of sediment data in the calibration process would reduce the range of uncertainty of at least some of the hydrologic parameters.

It was found that PEST's performance was somewhat disappointing during runs of this type due to the deleterious effects of low sensitivity and high correlation of some parameters. The adverse effects of parameter insensitivity and correlation are always worse when there are a large number of parameters to estimate than when there are only a few. In the present case these problems were overcome through judicious use of PEST's user-intervention functionality, in which troublesome parameters are temporarily held at their current values at critical stages of the parameter estimation process, leaving PEST free to adjust the other parameters. However this can be a labor-intensive process. Hence it was decided to estimate sediment parameters using a model for which hydrologic parameters had already been estimated using the methodology discussed in Doherty and Johnson. The hydrologic parameter values used in the present study are listed in the third column of Table 1.

Sediment Parameter Values

Sediment parameters estimated by PEST using the methodology outlined above are listed in the first column of Table 4. Convergence to this set of parameter values took place within 5 optimization iterations; no numerical difficulties were encountered by PEST.

As is discussed in Doherty (2001a), as a by-product of the Gauss-Marquardt-Levenberg method of parameter estimation, PEST is able to calculate the uncertainty associated with each estimated parameter. While uncertainty calculation by this means is based on a linearity assumption that is grossly violated in most modeling contexts, the uncertainty values achieved as a result of this process do serve to indicate the confidence levels that can be placed on parameters determined through model calibration. However in the present instance uncertainty calculation was impossible due to singularity of the parameter covariance matrix resulting from parameter nonuniqueness. The fact that the parameters listed in Tables 2 and 3 could not be estimated uniquely on the basis of the TSS data depicted in Figures 1 and 2 comes as no surprise. If desired, other sets of "calibration-constrained" sediment parameters, different from those in the first column of Table 4, but which calibrate the model just as well as these parameters, could have been estimated in the same manner as that in which multiple hydrologic parameter sets were calculated by Doherty and Johnston. This was not done in the present study; nevertheless, as is documented in the next section, the effects of sediment parameter nonuniqueness on model predictive nonuniqueness were explored using PEST.

In the course of undertaking the parameter estimation process, PEST calculates the sensitivity of each model output to which there is a corresponding field measurement to each adjustable parameter. It also calculates the "composite sensitivity" of each adjustable parameter, this being the sensitivity of that parameter to the model-generated counterparts to observations taken as a whole. If the composite sensitivity of a parameter is very low or zero, that parameter cannot be estimated through the inversion process. In a highly nonlinear parameter estimation problem such as that documented herein, some parameters can be "locally insensitive"; unfortunately, even "local insensitivity" makes estimation of the pertinent parameters very difficult.

The composite sensitivities calculated by PEST of parameters KRER and JRER were both zero. These parameters describe the ability of rain to detach sediment from the soil matrix. Detached sediment is then transported to a stream by overland flow if the sediment carrying capacity of overland flow is sufficient. This capacity is determined by parameters KSER and JSER. If these latter parameters are such that all detached soil cannot be transported overland, then the detachment parameters become insensitive, for it is then KSER and JSER which determine sediment export, rather then KRER and JRER. This was the case for the current PEST run. However if another set of initial parameter values had been chosen to begin the parameter estimation process, the opposite may have been the case as sediment export would then have been limited by the capacity of rainfall to detach sediment, rather then by the capacity of overland flow to transport it. The situation becomes even more complicated when it is considered that, on the basis of in-stream TSS measurements alone, it is impossible to distinguish detachment from scouring as the mechanism for sediment production. Hence estimation of the scour parameters KGER and JGER at the same time as the other sediment parameters mentioned above is virtually impossible.

It is thus apparent that, even without the problems incurred by the necessity to simultaneously estimate RCHRES SEDTRN parameters, estimates of PERLND SEDMNT parameters will always be accompanied by a large margin of uncertainty.

Comparison of Model Outputs with Measurements

In undertaking the parameter estimation process, PEST had little difficulty in reducing the discrepancies between TSS statistics (ie. mean and standard deviation as discussed above) and their model-calculated counterparts to almost zero. Similarly, PEST was able to ensure that the amounts of sand, silt and clay stored in the stream bed were unchanged over the calibration period. However, not surprisingly, a perfect fit could not be obtained between individual TSS measurements and the corresponding model outputs.

The top part of Figure 4 shows measured TSS values joined by straight line segments (dark lines). Model-calculated TSS values interpolated to measurement dates and times are joined by grey lines. This connection of measurements using linear segments is not meant to imply linearity of TSS concentrations between measurement times; it is simply a graphical means of conveying the "character" of the dataset, and of allowing a comparison to be made with the "character" of corresponding model outputs. It is apparent from Figure 4 that, as expected, the point-by-point matching of the two datasets is far from excellent. However, as was specifically sought through appropriate formulation of the objective function, the mean and standard deviation of the two datasets are very close, thus ensuring that modeled TSS values, when interpolated to the same dates and times as measured TSS values, have the same "look" when plotted and inspected.

In the bottom part of Figure 4, measured TSS values are superimposed on the complete model-generated TSS time series. Though far from perfect, the fit is easily as good as that which could have been achieved by manual calibration. Furthermore, the inclusion of bed storage information in the objective function ensured that this fit was not achieved at the cost of unnatural erosion or deposition of the stream bottom.

Predictive Analysis

General

Given the "lumped" nature of the parameters employed by a model such as HSPF, and given the fact that these parameters can be estimated with only a high degree of nonuniqueness through the calibration process, determination and documentation of a unique set of parameter values that purport to represent the erosion and transport characteristics of a watershed is a questionable activity. A more fruitful way to use a model such as HSPF in the investigation of sediment erosion and transport processes is to dispense with the idea of parameter uniqueness altogether. Instead, it is better to acknowledge that there is a (possibly large) range of parameter values that can result in acceptable fits between model outputs and field data (especially when the best fit that can be achieved is not very good), and that are in accord with "outside knowledge" of these values forthcoming from an understanding of the processes that they represent. It follows that there is also a (possibly large) range of parameter values that should be used when the model is deployed to make a prediction, and that there is thus a high potential for predictive nonuniqueness. Hence no prediction should be made by a model without some attempt being made to quantify the magnitude of uncertainty associated with that prediction. Such predictive uncertainty analysis can be undertaken with the help of PEST's predictive analyzer.

As is documented in Doherty (2001a), PEST's predictive analyzer calculates the maximum and minimum value that a model prediction can take, while ensuring that parameters used by the model are such as to maintain that model in a calibrated state. Thus in calculating the range of model predictive uncertainty, the calibration and prediction processes are combined. The user supplies a "limiting objective function" above which the model is deemed to be uncalibrated. PEST then adjusts parameter values in order to maximize or minimize the user-specified model prediction, while ensuring that estimated parameter values are such that the calibration criterion is not violated; the use of PEST's parameter bounds functionality ensures that parameters remain within acceptable ranges during this process.

The value selected for the limiting objective function depends on the types of observations used in the calibration process and the weights assigned to them. On the basis of the calibration strategy discussed in the previous section, PEST was able to lower the objective function to a value of 4.2×10^4. For the purpose of analyzing model predictive uncertainty, the limiting objective function was set at 4.8×10^4, this resulting in a model-to-measurement fit which is only slightly different from that achieved at the objective function minimum. Given the "tightness" of this limit, the extent of predictive uncertainty may have been underestimated in the process described below.

The prediction

The "prediction" in the present example is the total amount of sediment exported from the system over the period spanning 1975 to 1995, ie. over the total calibration period. Used in this way, HSPF acts as a "temporal interpolator" of the sporadic TSS measurements taken over the study period, thus assuming a role not too different from that of a sediment rating curve in performing calculations of this type. However, as has already been discussed, the advantage of using a model rather than a regression line to undertake such interpolation is that the model incorporates, at least to some extent, the mechanics of the operative processes. This, in turn, should enhance a modeler's ability to undertake predictive uncertainty analysis through using a tool such as PEST's predictive analyzer in conjunction with the model, for a model has the capacity to perform calculations for conditions that are different from those occurring during the calibration period using equations based on physical principals to perform extrapolation to the new conditions. Nevertheless the model, too, relies on "curve fitting" for the assignment of parameters through the calibration process; furthermore some of these parameters occur in equations which employ power functions of discharge (or quantities related to discharge). This could result in the calculation of inappropriately high uncertainty ranges when the model is used to predict sediment concentrations at flows that are much higher than those at which TSS measurements were made.

The total mass of sediment exported from the watershed over the model calibration period calculated using the "best-fit" parameters listed in the first column of Table 4 was 1.13×10^6 tonnes. Maximized and minimized sediment masses calculated using PEST's predictive analysis functionality in the manner discussed above, were 1.5×10^6 tonnes and 7.9×10^5 tonnes respectively. Parameters giving rise to these predictions are listed in columns 2 and 3 of Table 4. Visually, the fit between model outputs and field measurements over the calibration period for the "maximization" and "minimization" parameters is not too different from that depicted in the top part of Figure 4 for the "best-fit" parameters. The major differences between the respective model-calculated TSS time series, however, occur at extreme flow events where no TSS measurements were made. Figure 5 compares TSS measurements with the model-generated TSS time series for minimized (top picture) and maximized (bottom picture) total sediment export.

Conclusions

The use of nonlinear parameter estimation and predictive analysis methods in conjunction with a watershed simulation model to process suspended sediment data has been demonstrated. In common with most studies of this type, the data available for processing was sparse and unrepresentative of extreme system conditions. It was also "noisy" in the sense that it spanned a large range of measurement magnitudes, and was not directly amenable to fitting with the outputs of a process-based model. On the other hand, the data was not of such poor quality that its information content was zero. Thus environmental management of the watershed in which the data was gathered demands that it be processed, and that the results of this processing be incorporated into any predictions made of future watershed behavior under the same or altered land use practices.

Unlike many investigations based on computer simulation of environmental processes, use of a model in the present study was not based on the premise that a unique parameter set could be established which could then be used by the model to make all future predictions. Rather it was freely acknowledge that for a variety of reasons, including improper knowledge of watershed sediment processes and the availability of only a noisy and inadequate dataset, it would not be possible to ascribe to the model a set of parameters that would allow it to make precise predictions of sediment-related quantities. Hence the calibration process was seen as a means of imposing a complex set of constraints on parameter values used by the model; that is, no parameter set could be used by the model to make a prediction unless the parameters comprising that set were reasonable (while accepting the fact that the lumped nature of these parameters may broaden the bounds of what is considered reasonable), and unless that parameter set results in a satisfactory fit between model outputs and field measurements under historical conditions.

Once parameter nonuniqueness is accepted as a fact of life, use of a model to make predictions of system behavior, or to process data in order to derive secondary quantities of interest to watershed managers (as was done in the present investigation) must include an analysis of the uncertainty associated with model outputs. A further, perhaps more subtle, outcome of the acceptance of parameter nonuniqueness, is recognition of the fact that the model prediction process cannot be entirely separated from the model calibration process. This is because, in attempting to ascertain the uncertainty associated with key model predictions, the modeler must, with the help of tools such as PEST's predictive analyzer, vary parameters in such as way as to establish the range of uncertainty of those predictions while simultaneously ensuring that constraints imposed by the calibration process are respected.

We conclude this paper by reminding the reader that it was not our purpose to present the results of a detailed study of sediment erosion and transport processes operating in the Contentnea Creek system, for it is readily accepted that parameters presented herein are in need of further refinement. Rather, the purpose of this study was to explore, and then document, the use of nonlinear and predictive analysis methods in processing TSS data of the type depicted in Figures 2 and 3 to exemplify the type of processing methodology that is now readily available to all modelers. It is hoped that use of the techniques described herein will free the modeler from the heavy burden (often thrust upon him/her by those with a poor understanding of environmental modeling) of having to make a definitive prediction of some aspect of watershed behavior. Rather, the use of software such as PEST in combination with complex, process-based models such as HSPF, allows the modeler to process all available data to the maximum possible extent and, in the course of doing this, quantify the limits with which it is possible to predict system behavior. This represents a new, and much needed, addition to contemporary modeling practice.

Acknowledgements

J. D. wrote much of the time series analysis software that underpins the work documented herein while employed as a Visiting Research Scientist at the University of Idaho, Idaho Falls. He wishes to acknowledge the funding and resources made available to him from that institution that allowed the development of this software to take place.

Both PEST and the utility software used to carry out the analyses documented herein are available from the first author, or from the EPA CEAM web site at:

http://www.epa.gov/ceampubl/pest.htm

References

Bicknell, B.R., J.C. Imhoff, J.L. Kittle, T.H. Jobes and A.S. Donigian, 2001. HSPF User's Manual. Aqua Terra Consultants, Mountain View, California.

Cohn, T.A., DeLong, L.L., Gilroy, E.J., Hirsch, R.M., and Wells, D.K., 1989, Estimating constituent loads: Water Resources Research, v. 25, no. 5, p. 937-942.

Cohn, T.A., and Gilroy, E.J., 1991, Estimating loads from periodic records: U.S. Geological Survey Branch of Systems Analysis Technical Memo 91.01, 81 p.

Doherty J., 2001a. PEST-ASP User's Manual. Watermark Numerical Computing.

Doherty, J., 2001b. PEST Surface Water Utilities User's Manual. Watermark Numerical Computing and University of Idaho.

Doherty, J. and Johnston, J., 2002. Methodologies for Calibration and Predictive Analysis of a Watershed Model.

Giese, G.L., Eimers, J.L., and Coble, R.W., 1997. Simulation of Ground-water Flow in the Coastal Plain Aquifer System of North Carolina, in Regional Aquifer-System Analysis—Northern Atlantic Coastal Plain: U.S. Geological Survey Professional Paper 1404–M, 142 p.

Table 1. HSPF PWATER parameters estimated during the calibration process. Other parameters were assigned values independently of the calibration process. See Doherty and Johnston (2002) for details.

Parameter name	Parameter function	One set of estimated values from Doherty and Johnston (2002)
LZSN	Lower zone nominal storage	2.0 in
UZSN	Upper zone nominal storage	2.0 in
INFILT	Related to the infiltration capacity of the soil	0.0526 in/hr
BASETP	The fraction of potential ET that can be sought from baseflow.	0.20
AGWETP	Fraction of remaining potential ET which can be satisfied from active groundwater storage	0.00108
LZETP	Lower zone ET parameter - an index to the density of deep-rooted vegetation.	0.50
INTFW	Interflow inflow parameter	10.0
IRC	Interflow recession parameter	0.677 day^{-1}
AGWRC	Groundwater recession parameter	0.983 day^{-1}

Table 2. PERLND SEDMNT parameters estimated during the calibration process.

Parameter name	**Parameter function**
KRER	Coefficient in the sediment detachment equation
JRER	Exponent in the sediment detachment equation
KSER	Coefficient in the sediment removal equation
JSER	Exponent in the sediment removal equation
KGER	Coefficient in the sediment scour equation
JGER	Exponent in the sediment scour equation

Table 3. RCHRES SEDTRN parameters estimated during the calibration process.

Parameter Name	**Parameter function**
KSAND	Coefficient in equation (2) for sand carrying capacity
EXPSND	Exponent in equation (2) for sand carrying capacity
TAUCD (silt)	Initial shear stress for deposition of silt
TAUCRAT (silt)	Ratio of TAUCS to TAUCD for silt
M (silt)	Erodibility coefficient of silt
TAUCD (clay)	Initial shear stress for deposition of clay
TAUCRAT (clay)	Ratio of TAUCS to TAUCD for clay
M (clay)	Erodibility coefficient of clay

Table 4. Sets of estimated parameter values.

Parameter name	Best-fit parameter set	Parameter set for minimized prediction	Parameter set for maximized prediction
KRER	35.0	35.0	35.0
JRER	1.0	1.0	1.0
KSER	1.01	0.5	1.93
JSER	3.005	2.73	3.73
KGER	0.33	0.27	0.40
JGER	4.49	5.00	3.74
KSAND	3.58	3.93	3.32
EXPSND	0.49	0.43	0.55
TAUCD (silt)	0.103 kg/m^2	0.102 lb/ft^2	0.106 lb/ft^2
TAUCRAT (silt)	2.29	2.30	2.27
M (silt)	0.0037 kg/m^2/hr	0.0039 lb/ft^2/day	0.00416 lb/ft^2/day
TAUCD (clay)	0.083 kg/m^2	0.082 lb/ft^2	0.085 lb/ft^2
TAUCRAT (clay)	3.045	3.038	3.02
M (clay)	0.0037 kg/m^2/hr	0.0038 lb/ft^2/day	0.00416 lb/ft^2/day

Units for many of these parameters are complex, due to the exponential term in the equations which contain them.

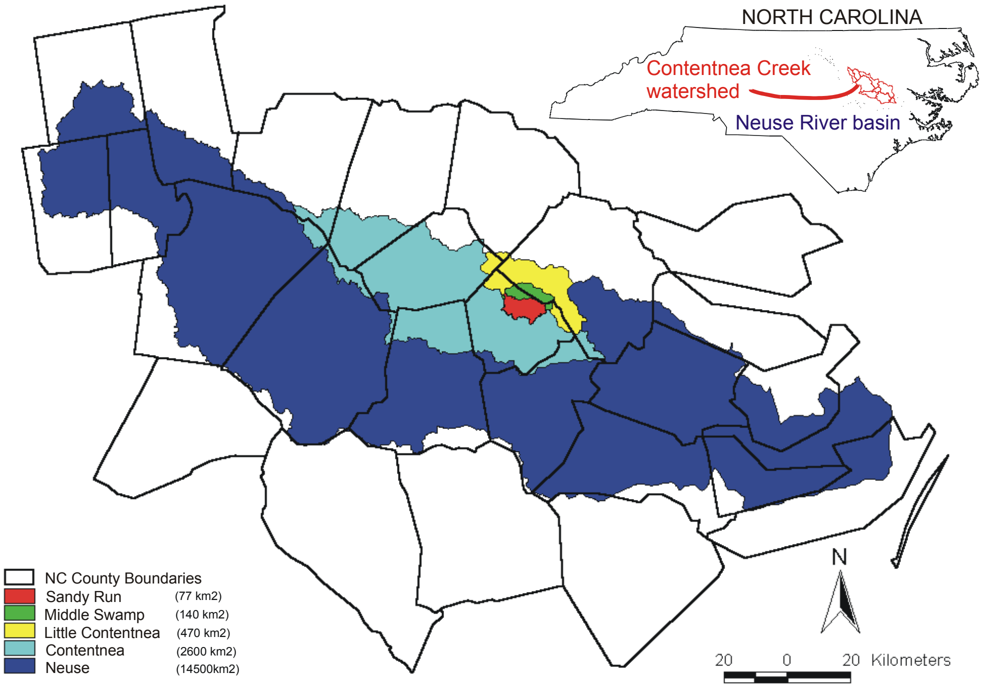

Figure 1. Contentnea Creek watershed study area and surroundings. The Contentnea subwatershed is located within the Neuse River Basin of North Carolina.

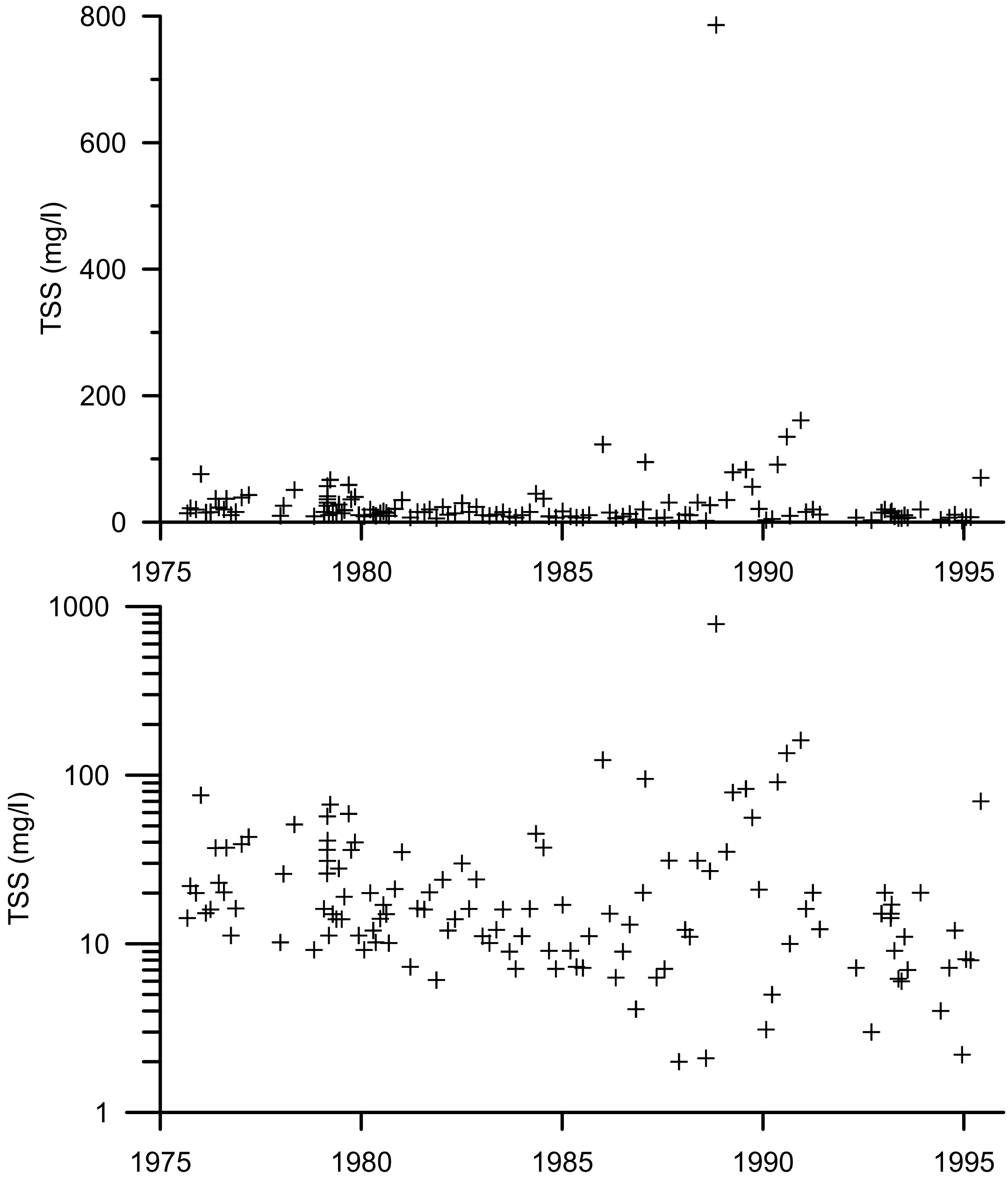

Figure 2. TSS data gathered over the period 1975 to 1995 at Hookerton Gauging Station.

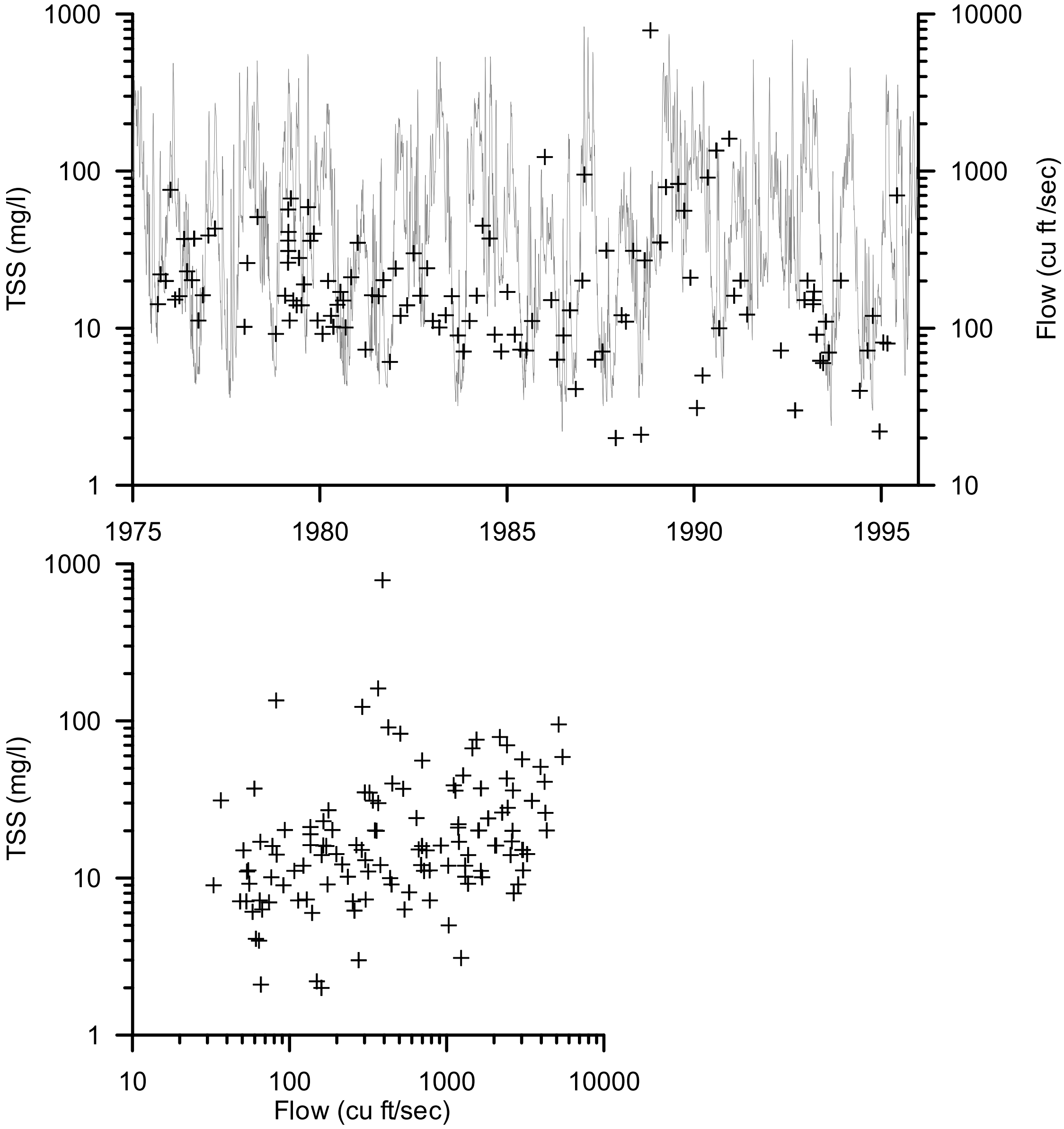

Figure 3. The top part of this figure shows TSS measurements superimposed on stream flow measurements. TSS is plotted against flow in the lower part of the figure

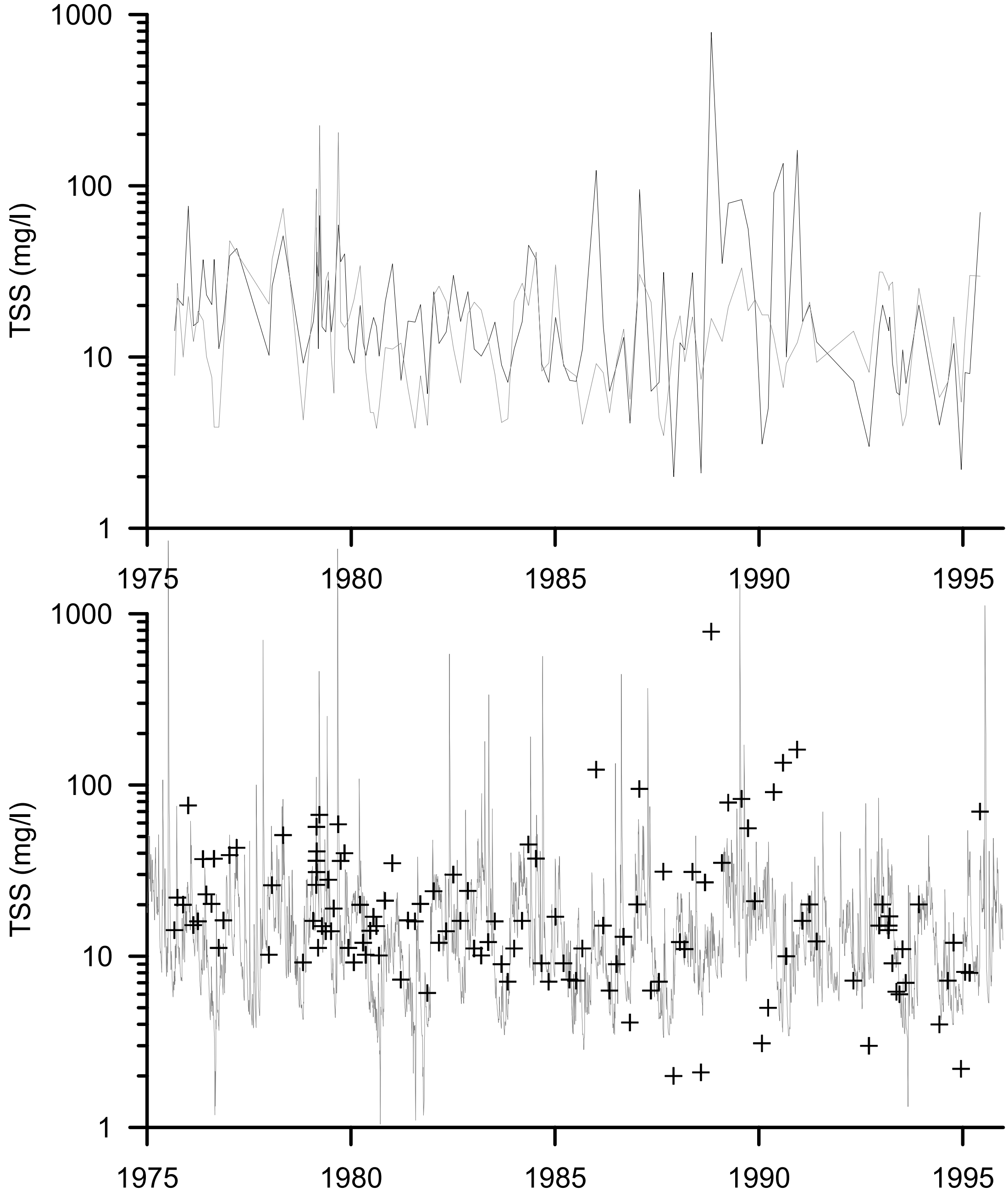

Figure 4. The top graph allows a comparison between TSS measurements and model outputs to be made on a point-by-point basis. In the bottom graph TSS measurements are superimposed on the model-generated TSS time series. In both of these graphs the model output is depicted in grey.

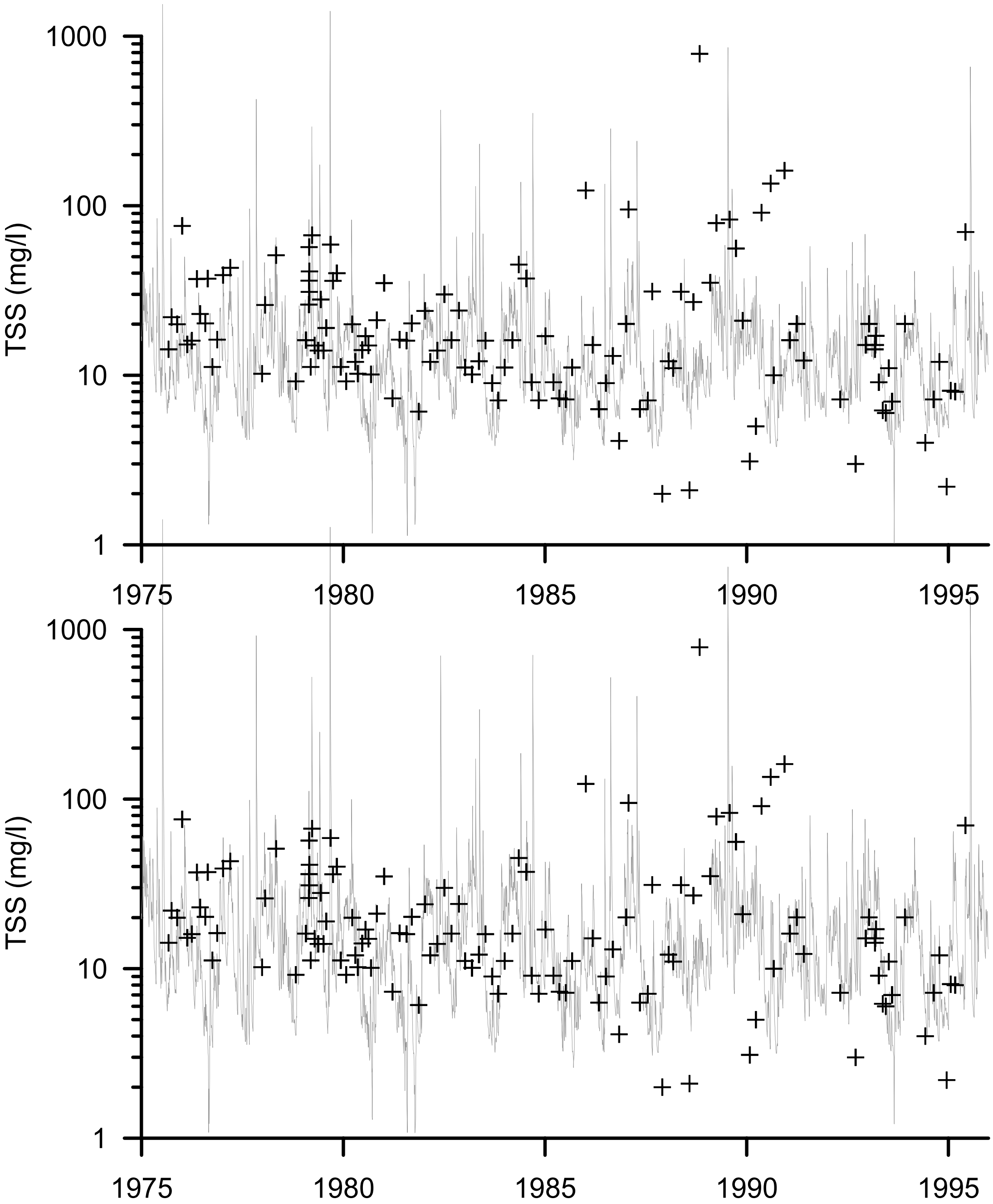

Figure 5. Observed and model-generated TSS values over the calibration period. The total exported suspended mass over the calibration period is minimized in the top graph and maximized in the bottom graph.