Uncertainty Analysis and Hydrologic and Hydraulic Model Linkage in the Watershed Modeling System

Christopher M. Smemoe, Research Associate, Environmental Modeling Research Laboratory, 242-B CB, Brigham Young University, Provo, UT 84604; phone (801) 422-5713; smemoe@byu.edu
Ahmad Salah, Graduate Student Brigham Young University, 242 CB, Brigham Young University, Provo, UT 84604, asalah@byu.edu
E. James Nelson, Associate Professor, 242-D CB, Brigham Young University, Provo, UT 84604; phone (801) 422-7632; jimn@byu.edu

Abstract

While the process of modeling and mapping a 100-year floodplain is inherently full of uncertainty, engineers are still required to determine a single line boundary for flood insurance rate and other floodplain studies. This is despite the fact that a good engineer knows that, within reasonable limits, different modeling parameters can be chosen that result in shifting the floodplain boundary either in or out. Because of the uncertainty associated with modeling input parameters it is not difficult to come up with multiple answers that are reasonable in approach, and yet quite different in results. The Environmental Modeling Research Laboratory (EMRL) at Brigham Young University (BYU), in conjunction with the Army Corps of Engineers (USACE), has developed modeling tools that allow for Monte Carlo type simulations to be run on developed hydrologic (HEC-1), hydraulic (HEC-RAS), and floodplain delineation models in order to determine a probabilistic flood plain boundary. By linking the models so that the result of one model can be used automatically as input to the next, several simulations can be run which explore the distribution of probable values on such important parameters as rainfall, watershed losses, discharge, and channel roughness. The map resulting from these simulations shows the probability of flooding within certain regions. The 10, 50, 100, or any other percent probability of flooding can easily be identified. This allows the engineer to create a visualization of the floodplain that is "honest" in terms of indicating the inherent uncertainty in the processes that are used to develop a floodplain map.

By taking this process one step further and examining, through Monte Carlo type simulation, the entire space of probable values for each return period, a more certain boundary for the 100-year (or any return period) floodplain can be determined. The process of exploring the range of possible values over all probabilities is done in a fashion similar to the Corps of Engineers Flood Damage Assessment (HEC-FDA) program. The result of such a process results in a spatial map of annual exceedance probabilities. This map is determined by dividing the number of times a particular point is flooded from a simulation by the total number of simulations. Contouring the 0.01 exceedance probability results in a more certain 100-year floodplain that incorporates the uncertainty inherent in the modeling parameters. Such a line is not the result of a single

set of modeling parameters and could not be reproduced by a single simulation; rather it is the composite of all the simulations.

This paper and presentation will focus on the latest developments of the Watershed Modeling System in addressing uncertainty for hydrologic and hydraulic modeling.

Introduction

Anybody who has created a hydrologic or hydraulic model of any sort realizes that the modeling process is fraught with uncertainty. For example, if a hydrologist has a set of rain gages, how can she determine the precipitation that occurred in a watershed during an intense storm event? Here are three approaches:

1. Draw Thiessen polygons between each gage station and determine the area-weighted contribution to the watershed's storm from each rain gage,
2. Draw precipitation contours (isohyets) from the gage data and determine the area-weighted contribution from each contour interval, or
3. Use the arithmetic mean of the gages.

All three precipitation approaches provide different precipitation values, yet all are valid. There is another factor of precipitation uncertainty. The amount of precipitation is not the same at each point in a watershed, and large watersheds may have areas of heavy rainfall and areas with no rain at all. An areal reduction factor can be applied to rainfall values in large watersheds, though this value is uncertain. Radar data have helped to quantify the distribution of rainfall in a watershed, but even these datasets can be prone to inaccuracy and calibration errors. Some researchers, such as Veneziano and Langousis (2005), have more accurately defined the areal reduction factor, but an element of uncertainty remains.

This element of uncertainty exists in each parameter used in a hydrologic, hydraulic, or water quality model. From loss coefficients to Manning's roughness values, the physical characteristics of a watershed, floodplain, or lake are difficult to pinpoint. Watershed characteristics change from storm to storm and they even change during a storm. Floodplain and lake characteristics also change over time.

How can one quantify this uncertainty? This paper will discuss some procedures that have already been developed and others that are under development for defining the uncertainty of parameters in hydrologic, hydraulic and water quality models. This paper will also show how this input uncertainty changes how the results of hydrologic, hydraulic, and water quality models are displayed.

Methods

Creating a detailed floodplain map requires one to obtain flooded locations and water depths (within 0.1 feet (0.03 M)) at every point in a floodplain. This is done by running a hydrologic model to determine the 100-year storm flow (or any other return period), using this flow as input to a hydraulic model to determine the flood stage at this 100-year flow, and then delineating the floodplain boundary at this flood stage (see Figure 1). This whole process is a difficult, if not impossible, task since every input value to a hydrologic and hydraulic model has some amount of uncertainty.

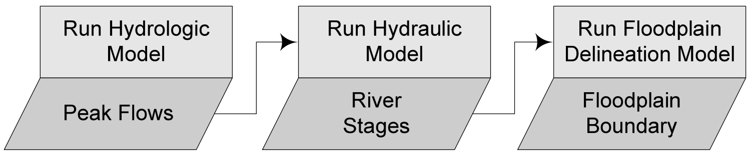

Figure 1: Creating a floodplain map

This paper will show how an automated process can be used to run hydrologic, hydraulic, and floodplain delineation models repeatedly to create two types of maps: flood probability and annual exceedance probability maps. A flood probability map is a contour map showing the probability of flooding at each point in a floodplain during a certain recurrence interval (such as the 100-year interval). An annual exceedance probability map is created by simulating several hundred yearly peak floods. This map shows the probability of surface water flooding occurring at any point in the floodplain during any single year.

In the US, two dimensional water quality models have been used for several years. The CE-QUAL-W2 model has been widely used for hydrodynamic and water quality analysis. On the other hand, two-dimensional distributed hydrologic models are becoming more accepted as tools for hydrologic analysis. The Gridded Surface Subsurface Hydrologic Analysis (GSSHA) model has been widely used for hydrologic modeling. It includes two-dimensional overland flow coupled with a one-dimensional stream flow, a one-dimensional infiltration model, and a two-dimensional groundwater model. This paper discusses how GSSHA can be combined with CE-QUAL-W2 to address uncertainty in integrated water resources/water quality management of watersheds. Temporal and spatial uncertainty of water quality parameters can be estimated by integrating GSSHA and CE-QUAL-W2 in a stochastic approach. This can be accomplished by running GSSHA stochastically and by generating multiple output files which CE-QUAL-W2 would take as multiple input files for various runs. The output of these runs is then used to report on the temporal and spatial uncertainty.

All the capabilities presented in this paper have been developed in the Watershed Modeling System (WMS). WMS is a computer program used as a pre- and post-processor for hydrologic, hydraulic, floodplain, and water quality models. It has been developed by the EMRL at BYU in conjunction with the USACE and other government and private entities. A person generates a model and computes all its parameters in WMS using GIS data. The model is then saved to a model input file, and the model is run from the software. WMS then reads and displays the results from the simulation.

Flood Probability Mapping

A flood probability map is a contour map showing the probability of flooding at each point in a floodplain during a certain recurrence interval. For example, the 100-year recurrence interval is frequently used when performing a floodplain study. A 100-year flood probability map shows the probability of any point on the floodplain being flooded during a 100-year flood event.

How is a flood probability map created? Two different scenarios can exist when creating this type of map. In one scenario, one has a mean value for the 100-year flow rate and its standard deviation. From this mean and standard deviation, a non-skewed, normally-distributed probability density function (PDF) can be generated. A person simulates this flow in a hydraulic model to get river stages and then determines the floodplain boundary from these stages (see Figure 2). The process of running the hydraulic model and the floodplain delineation model is repeated a sufficient number of times to obtain a wide range of floodplain boundaries for the 100-year storm. Each time the hydraulic model is run, a stochastic sampling of the probability distribution of one or more input parameters (such as the 100-year flow rate) is used to determine the value for that parameter. A flood probability map is created by dividing the number of times each point in the floodplain is flooded by the total number of simulations. This operation gives the probability of flooding at any point in the floodplain.

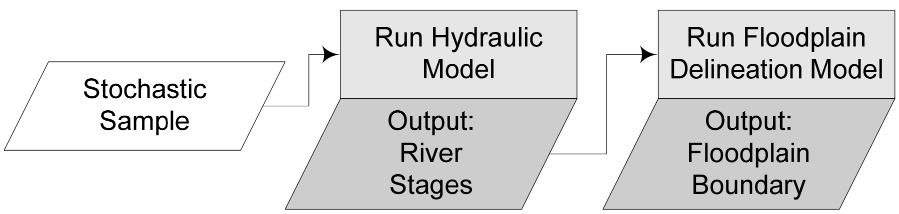

Figure 2: Creating a flood probability map with known flow rates

In another scenario, the 100-year flow rate and its uncertainty distribution is unknown. In this scenario, a hydrologic model must be included in the simulation process to compute a flow rate. This flow rate is then used in the hydraulic model, and the river stages from the hydraulic model are used in the floodplain delineation model (see Figure 3). This process is repeated a certain number of times. For each repetition, probability distributions of one or more parameters (such as the 100-year precipitation) are sampled to determine input values to the hydrologic and hydraulic models.

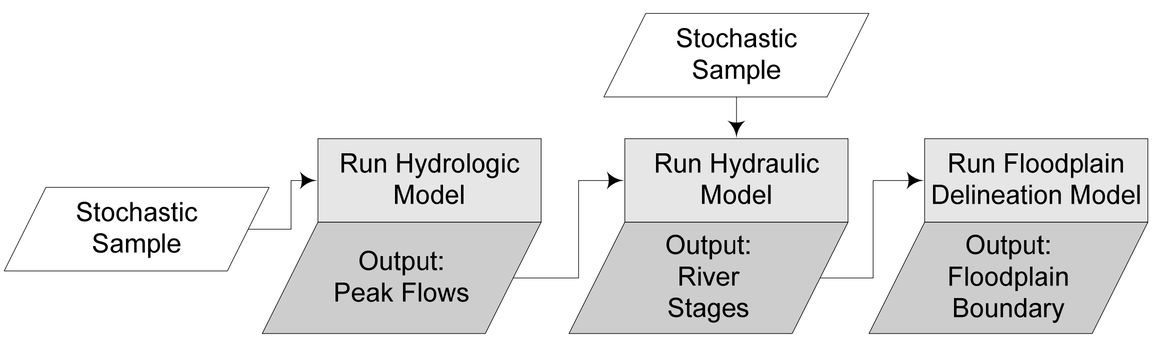

Figure 3: Creating a flood probability map when flow rates are unknown

The result from running either of these types of simulations is a flood probability map. This map shows contours that represent the probability of flooding at a certain recurrence interval (see Figure 4).

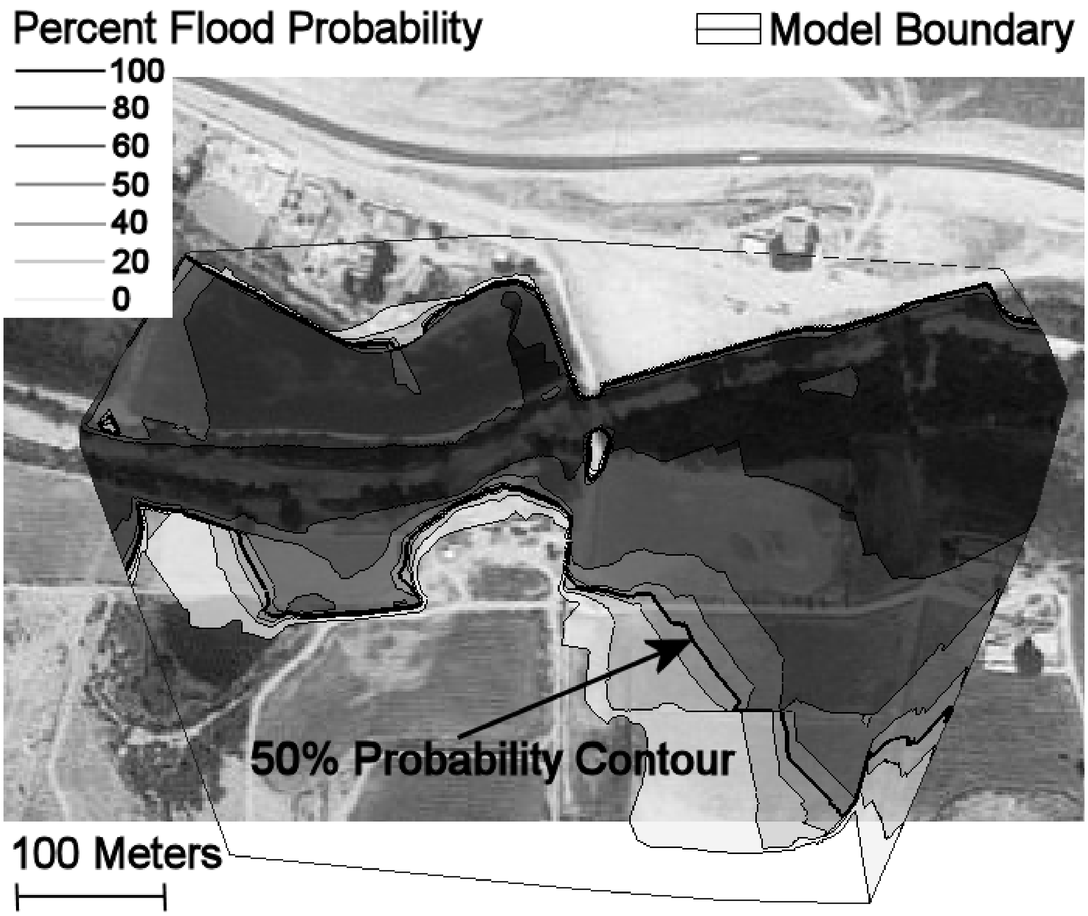

Figure 4: Flood probability map for a section of the Virgin River in southern Utah

Annual Exceedance Probability Mapping

Generating a flood probability map requires a probability distribution of precipitation or flow for a certain recurrence interval. What if, instead of wanting to determine what the probability of flooding is for a certain recurrence interval, you want to determine what the probability of flooding is during any single year? This probability is called the Annual Exceedance Probability (AEP), a value frequently used by the US Army Corps of Engineers (1996).

A map of annual exceedance probabilities is created in the same manner as a flood probability map; i.e. by defining a probability distribution and by running a floodplain delineation several times to generate a probability map. The key difference between the annual exceedance probability map and the flood probability map is how the probability distribution for discharge or precipitation is defined. Instead of defining the probability distribution for a single recurrence interval, the distribution for several recurrence intervals is defined (see Figure 5). The area of uncertainty associated with all the recurrence intervals is then sampled from this discharge or precipitation-probability curve (see Figure 6).

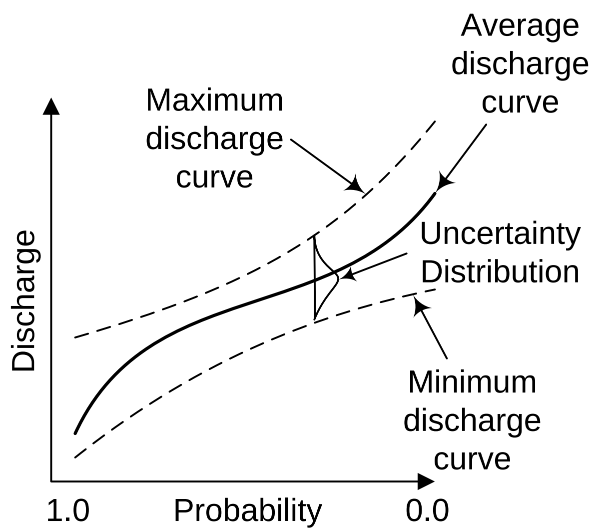

Figure 5: A discharge-probability curve and its uncertainty

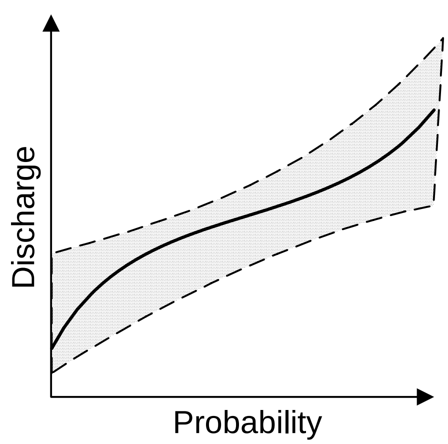

Figure 6: Canvassing the space of the discharge-probability curve

When generating an AEP map, each floodplain simulation is similar to generating a map for the peak discharge in a single year. That peak discharge could be the 2-year discharge or the 500-year discharge. For example, running 100 floodplain simulations is similar to simulating 100 years of peak discharges. This means that when the flood probability map is generated for an AEP simulation, the probabilities at any point on the map represent the probability of flooding at that point during any single year. In other words, if a point floods once during the hundred simulations, it has a flood probability of 0.01. This probability can be converted to a return period since Return Period = 1 / Probability. The return period for the 0.01 probability contour on the map represents the 100-year recurrence interval floodplain. The same result can be obtained by converting the 0.1 probability to the 10-year floodplain, the 0.02 probability to the 50-year floodplain, and so forth. This result is demonstrated by the AEP map for a section of Utah's Virgin River shown in Figure 7.

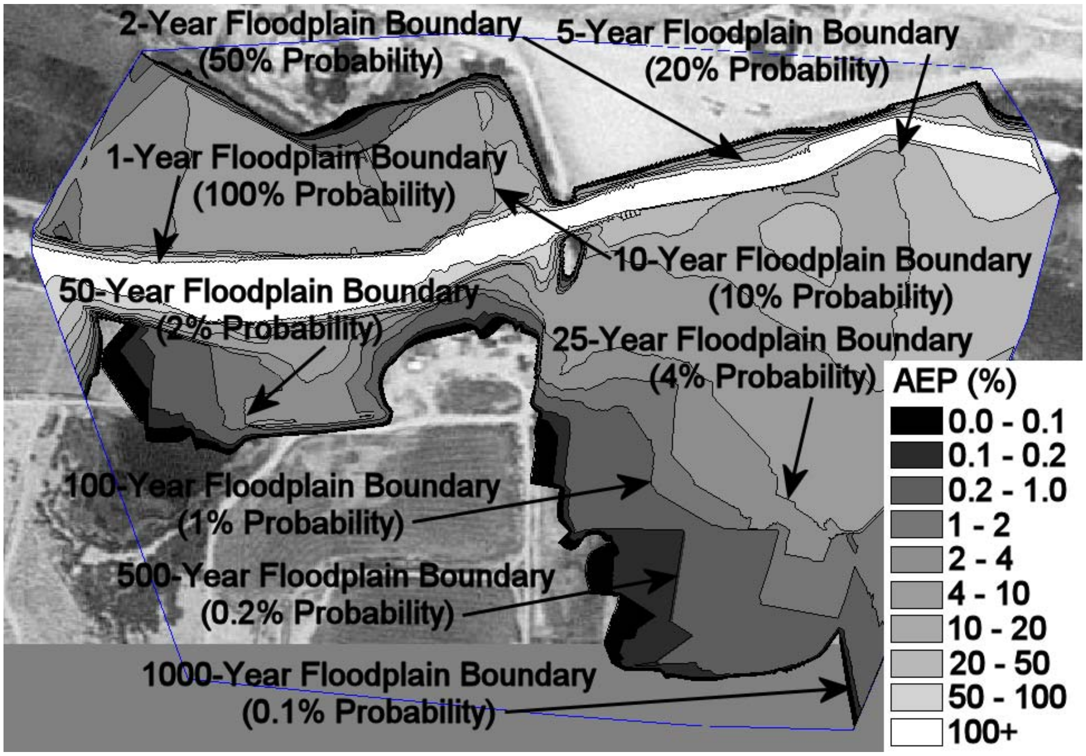

Figure 7: An AEP map showing the floodplain boundaries for different recurrence intervals

Future Development

Levee Failure Uncertainty Analysis

Currently, when one creates an AEP map using WMS, it is assumed that no levee failure occurs. However, levee failures can certainly occur, as evidenced by the recent levee failures during Hurricane Katrina in New Orleans. How can this levee failure be accounted for in a spatial manner? If the water surface is higher than the levee, there is a 100 percent chance that the levee has failed since the water depth is higher than the levee. This case is currently handled since the floodplain delineation process recognizes that the water surface overtops the levee. However, if the water surface is just beneath the top of the levee, there is still a probability that the levee will fail geotechnically at some location. As the water surface rises to the top of the levee, the probability of geotechnical levee failure increases. This phenomenon has been accounted for by the current USACE flood damage analysis procedures (1996).

The current USACE procedures define a curve depicting the probability of levee failure at each stage value. Two points are defined on this curve, a probable non-failure point (PNP) and a probable failure point (PFP), and the probability is linearly interpolated between these two points, as shown in Figure 8.

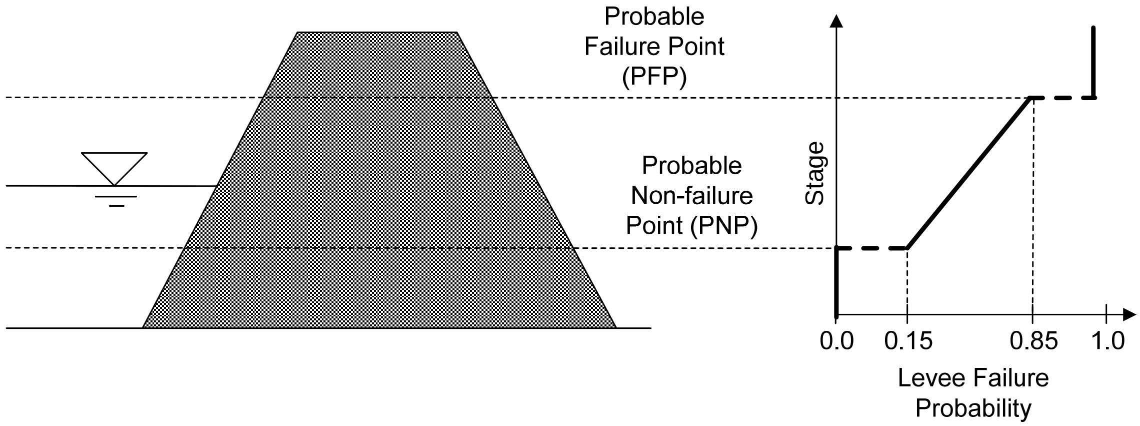

Figure 8: Process for determining geotechnical levee failure (adapted from USACE, 1996)

One limitation of the USACE approach is that it does not consider the spatial nature of levee failure. However, the spatial nature of failure could be incorporated into the process of creating an AEP map by defining the locations of levees as a line in a GIS-based model and breaking the levee into sections, with each section having the possibility of failing defined by the levee failure probability function shown in Figure 8.

Elevation Uncertainty

When generating an AEP map, one has the ability to define the probability distribution of single-variable inputs to hydrologic or hydraulic models such as loss coefficients and floodplain discharge. However, there is a degree of uncertainty in the elevation data used to generate the hydraulic model and to delineate the floodplain. Much research has been performed on evaluating the uncertainty of elevation data (see Jones et al, 1998 and Endreny and Wood, 2001).

Many times, since elevation data has been determined using remotely sensed data, areas of uncertainty can be divided into areas with different land covers. Areas with dense vegetation have a higher degree of uncertainty than areas with a small amount of vegetation. A hydraulic or a floodplain delineation model can be divided into different "uncertainty zones", with each zone having its own elevation uncertainty distribution. The elevations in each zone can be raised or lowered by a certain amount for each simulation. This would create a flood probability map that considers elevation uncertainty.

Two-Dimensional Finite Difference Hydrologic Modeling

GSSHA is a two dimensional, physically based, distributed-parameter hydrologic model intended to identify runoff mechanisms and simulate surface water flows in watersheds with both Hortonian and non-Hortonian runoff. It is capable of simulating stream flow generated from Hortonian runoff, saturated source areas, exfiltration, and groundwater discharge to streams. The model employs mass-conserving solutions of partial differential equations (PDEs) and closely links the hydrologic compartments to assure an overall mass balance and correct feedback (Downer et al., 2002).

The fact that GSSHA is a two-dimensional distributed model does not eliminate the fact that any set of input parameters, whether it is close to reality and actual measurements or not, will always produce output that can look professional. Yet, how reliable would these results be? And how would they compare to actual values and real-world measurements? It is probably worth mentioning here that, just like any other model, GSSHA input parameters exhibit some level of uncertainty that in most cases are hard to quantify without the help of some statistical technique, such as simulation or an arbitrary margin of safety (Walker, 2003).

To overcome these limitations, the GSSHA interface in WMS is incorporating a stochastic module so modelers can setup a "normal"; i.e. deterministic, GSSHA model for their respective watershed, using one unique set of arbitrary, suggested or calibrated input parameters. Modelers can expand on their deterministic models by defining a stochastic version of the model. In this stochastic model, some parameters of the choice of the modeler are picked from an assumed PDF to represent the parameter for one simulation. This picked value would vary in each simulation for a large number of simulations.

Water Quality Probabilistic Modeling

Water quality probabilistic modeling is a natural extension to hydraulic and hydrologic model uncertainty. Obviously, water quality modeling makes no exception to the above illustrated scheme. This is particularly true because of the overall lack of water quality data. And since one of the major benefits of simulation is that it overcomes the disadvantage of missing/lacking data (Salah et al, 2005), it is always advisable to simulate the values of input parameters in case any data are lacking.

Similar to the above stochastic flood plain delineation by linking HEC-1 and HEC-RAS, integrated water resources/water quality modeling is currently under development in WMS. Linkage of a two-dimensional finite difference hydrologic model (GSSHA) to a two dimensional laterally averaged hydrodynamic and water quality model (CE-QUAL-W2) is an addition to the integrated capabilities of WMS (Figure 9).

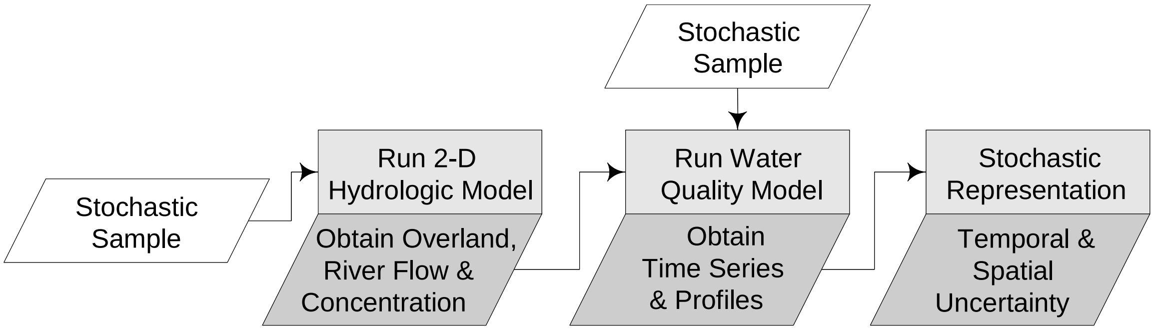

Figure 9: Temporal and Spatial Water Quality Modeling in WMS.

One of the typical questions answered by water quality models is: What is the time series of a certain pollutant at a specific point in time? Most water quality models can estimate a deterministic linear representation of concentration based on given set of input parameters. The other typical finding of a water quality model is a 2-d representation of the concentration of a water quality constituent along a horizontal plane (i.e. map) or longitudinal sections of the water body. These representations usually divide the water body area or section into regions of varying concentrations at a given time step. These regions are deterministically separated by a "single line boundary" (Figure 10).

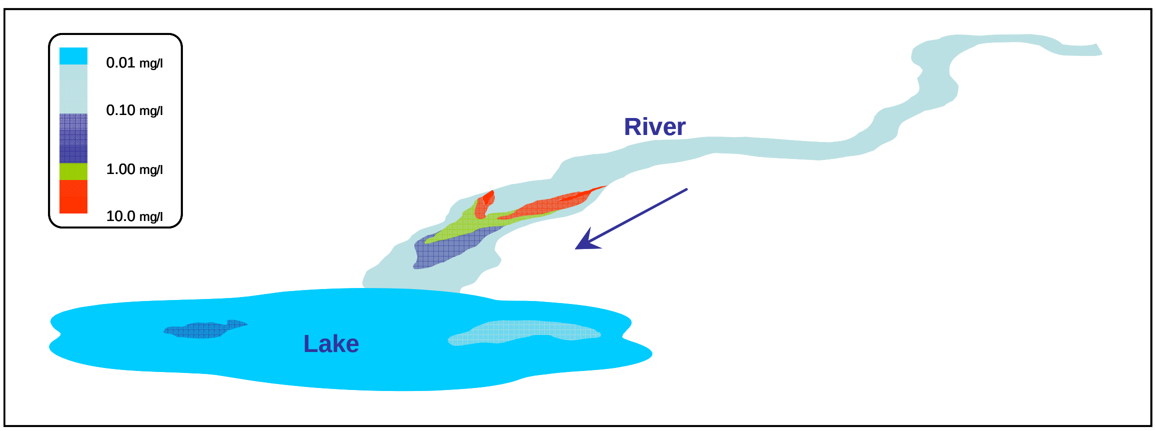

Figure 10: Hypothetical deterministic pollutant concentration map.

As a result of the stochastic integration of the GSSHA and CE-QUAL-W2, modelers would be able to view the output of CE-QUAL-W2 in a credible interval as opposed to the "single" line boundary or curve (for spatial and temporal uncertainty respectively, (Figure 11(a))). The credible interval (Figure 11(b)) would encapsulate all the distributional values that the accompanying probability entails. This means that a 95% credible interval of a time series would mean that there is a 95% probability that the concentration of a pollutant lies within this interval.

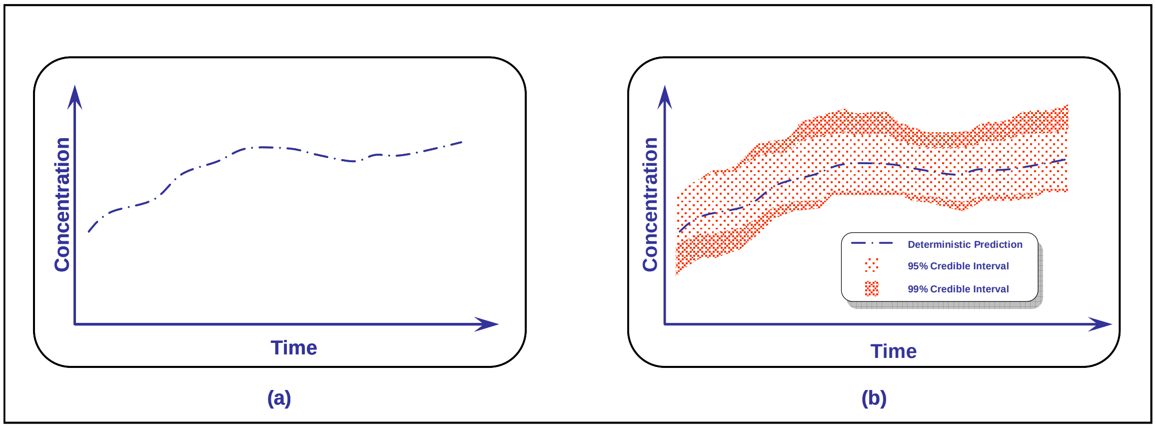

Figure 11: Stochastic vs. deterministic water quality models output.

Conclusion

Tools have been created allow people to quantify the uncertainty of input parameters to hydrologic, hydraulic, floodplain delineation, and water quality models. By quantifying the uncertainty of input parameters, the uncertainty of outputs can be evaluated using flood probability maps, AEP maps, and water quality probability maps. These maps are useful tools, but more research needs to be done in this area. Two areas of research include incorporating levee failure uncertainty and elevation uncertainty into the procedures presented in this paper.

These tools will enable engineers to provide more accurate spatial and temporal estimates of watershed discharge, flood extents, and pollutant concentrations. The stochastic integration of HEC-1, HEC-RAS, and the WMS floodplain delineation models provide for the generation of flood probability and AEP maps. Additionally, the integration of GSSHA and CE-QUAL-W2 will provide for the generation of water quality probability maps. These maps are useful tools for estimating possible flood damage areas or areas with undesirable pollutant concentrations.

References

Downer C.W., Ogden F.L., 2002. GSSHA User's Manual, Gridded Surface Subsurface Hydrologic Analysis. Version 1.43. ERDC Technical Report, Engineering Research and Development Center, Vicksburg, MS, USA.

Endreny, T. A., and Wood, E. F., 2001. Representing Elevation Uncertainty in Runoff Modelling and Flowpath Mapping. Hydrological Processes, 15:2223-2236.

Jones, J. L., Haluska, T. L., Williamson, A. K., and Erwin, M. L., 1998. Updating Flood Maps Efficiently: Building on Existing Hydraulic Information and Modern Elevation Data with a GIS. U.S. Geological Survey Open-File Report 98-200. *Available at* http://wa.water.usgs.gov/pubs/ofr/floodgis/. Accessed in November, 2005.

Salah A., Fields P., Miller A. W., 2005. Simulating Uncertainty in Mass Balance Modeling For Fresh Water Reservoirs; Case Study: Deer Creek Reservoir, Utah, USA. ***In:*** Winter Simulation Conference, Dec 4-7 2005, Orlando, FL, USA, Institute for Operations Research and the Management Sciences, On CD.

U.S. Army Corps of Engineers (USACE), 1996. Risk-Based Analysis for Flood Damage Reduction Studies: Engineering Manual 1110-2-1619, Washington, D.C., USA.

Veneziano, D., and Langousis, A., 2005. The areal reduction factor: A multifractal analysis. Water Resources Research, 41:W07008.

Walker, W. W. 2003. Consideration of Variability and Uncertainty in Phosphorus Total Maximum Daily Loads for Lakes. Journal of Water Resources Planning and Management, 129(4):337-344, American Society of Civil Engineers.