

A web-based environmental decision support system for legacy models

Saurav Kumar, Adil N. Godrej and Thomas J. Grizzard

ABSTRACT

An environmental decision support system (EDSS) was designed for the Occoquan system in Northern Virginia, USA. This EDSS is available through the internet using web-browsers, and enables stakeholders to interact with complexly-linked water resources models for the Occoquan system based on seven implementations of HSPF and two implementations of CE-QUAL-W2 software. Using the web-interface of the EDSS, users may delineate land use changes and simulate the water quality impact of such changes by remotely executing the water resources models. The EDSS utilizes a server cluster to share the computational load of simultaneously executing multiple instances of the linked Occoquan system models along with methods to limit 'similar' model executions. The server cluster was assembled from disparate machines with spare computing resource available on the local network, thereby eliminating the need for any additional hardware to execute an increased number of model simulations. It is expected that the enhanced accessibility to the water resources models through the EDSS may allow stakeholders to use the models as a planning and educational resource, without direct expert modeler's involvement. Further, this EDSS is comprised of modules that may be extended to other watersheds with similar legacy, calibrated modeling systems.

Key words | environmental decision support system, Occoquan, water resources model, web-based land use modification

Saurav Kumar (corresponding author)

Adil N. Godrej

Thomas J. Grizzard

Occoquan Watershed Monitoring Laboratory,
The Charles E. Via, Jr. Department of Civil and
Environmental Engineering,
9408 Prince William St,
Manassas, VA 20110,
USA
E-mail: saurav@vt.edu

INTRODUCTION

Computer-based water resources models designed to simulate natural interactions in a watershed-waterbody system are often applied by experts to manage such systems. Appropriate application of water resources models, with recognition of underlying assumptions and associated uncertainties of the processes simulated, enables expert modelers to test different solutions and analyze the impacts of changes in water resources indicators/constituents. Most water resources modeling software packages, such as Hydrological Simulation Program Fortran (HSPF), CE-QUAL-W2 (www.ce.pdx.edu/w2/), Soil and Water Assessment Tool (SWAT), and MIKE (www.mikepoweredbydhi.com/), with state-of-the-art science require significant expertise and data to calibrate for the watershed and/or waterbody of interest. Subsequent usage of calibrated

models as tools for planning and management of reservoirs and their watersheds also requires significant knowledge and familiarity with the modeling system. This complexity in usage of the models may exclude most non-expert actors, even those who may have commissioned the task, from many phases of the watershed planning process. Given that most decision making for the watershed often involves considerations apart from water resources, such as economic, and socio-political considerations (Kiker *et al.* 2005), exclusion of key stakeholders from the simulation tasks may, in-turn, reduce the weight associated to the output from water resources modeling in the decision making process for the watershed.

The new paradigm for partnership within a watershed seeks to involve not only decision makers but all citizens.

There is growing recognition, as evidenced by the US Environmental Protection Agency's handbook for developing watershed plans, that directly involving key stakeholders with the water resources model may be advantageous and is becoming a part of many watershed-based approaches of urban management (USEPA 2008). Various types of stakeholder partnership schemes have been tried and were shown to be beneficial for implementing and designing watershed based solutions. Some potential benefits of the stakeholder partnership, summarized from various studies in the literature (USEPA 2001, 2012; Korf-macher 2001; Leach *et al.* 2002; Irvin & Stansbury 2004; Murdock *et al.* 2005; Diduck *et al.* 2007; Reed 2008), include the following:

- gains from local knowledge and insight which experts might overlook;
- reduced likelihood of solution failures due to non-acceptance among stakeholders;
- increased fairness and equity in decision-making processes by involving all parties with a stake and addressing or taking into consideration their issues;
- better possibility for establishment of a common ground among stakeholders with different objective functions through negotiation;
- avoidance of failures due to over-reliance on expert opinion.

Water resources experts frequently derive optimal solutions for watershed management using models. Letting stakeholders interact with the models may help in highlighting the advantages of the optimal solution and improve the likelihood of its eventual acceptance. In addition, the process of stakeholder participation empowers the stakeholders to influence the formulation of the optimal solutions in many ways. For example, by choosing to address the uncertainty in simulations differently, from how the uncertainty was addressed by experts to derive the optimum solution, stakeholders may reach a different conclusion and solution. Often no clear optimum solution exists, due to the inherent uncertainty in the modeling process and the complexity of the natural processes modeled. Thus, involving stakeholders becomes even more appropriate as it lets people affected by the decisions (stakeholders) participate in making some assumptions necessary to derive the

decision (McIntosh *et al.* 2008). Furthermore, participatory exercises involving water resources models encourage negotiations among stakeholders with different objectives (USEPA 2001).

The literature contains examples of many methods to involve stakeholders in the water resources modeling. Some of these partnership schemes involve developing the water resources model with stakeholder involvement, such as, cooperative modeling (Cockerill *et al.* 2006). Other methods seek partnership at a later stage after development of models by experts. It has been reported that development of models with stakeholders encourages better stakeholder education, among other not readily measurable benefits (Cockerill *et al.* 2006). Variations of participatory modeling, such as Group Model Building, Mediated Modeling, and Companion Modeling, do not rely on robust traditional water resources models, but rather utilize modeling methods from the system dynamics domain, use Bayesian networks and agent based modeling approaches which are easier for non-expert stakeholders to use and understand (Voinov & Bousquet 2010). However, there are cases where calibrated and validated traditional water resources models for water-body and watershed systems already exist, and developing a new model may not be a viable option due to economic and technical considerations. In such cases, many benefits of involving stakeholders may also be achieved by extending these already calibrated models with sound science to the stakeholders. The problem, of course, is the fact that these models with state-of-the-art and robust science were not designed for the modern internet and cloud-computing centric world, hence the term 'legacy models'. With the available user interface and mechanisms used to calibrate and utilize most of these models, it is difficult to make them available to non-expert stakeholders. Further, the computation resources required in some cases may be prohibitory for using these models for decision-making and what-if type analysis.

The task of making the water resources models accessible to the stakeholders has become easier and more efficient than in the past because of developments in the internet and Geographical Information System (GIS) technologies in the last decade (Steiniger & Hunter 2012). There have been many attempts to develop web- and GIS-based systems in several water resources domains, such as a web-based

system designed to provide access to simulations of stream water quality in the USA (Booth *et al.* 2011), 'aquavoice' a system based on MIKE-SHE using JAVA technologies for ground water modeling and decision support (Jonoski 2002), a system for Idaho ground water management (Tuthill *et al.* 2000), a web-based decision support tool based on multi-criteria decision making (MCDM) (Khelifi *et al.* 2006), a web-based stakeholder collaboration tool for flood risk management (Almoradie 2014; Almoradie *et al.* 2015), and a real-time flood forecasting and analysis system for Shuangpai region in China (Li *et al.* 2006). Often simpler models or meta-models that work with information derived from robust calibrated modeling systems are used for web-based environmental decision support systems (EDSS). There is also some evidence that neural network and other empirical models perform adequately (Chen & Chau 2006; Muttil & Chau 2006; Wu *et al.* 2009; Taormina *et al.* 2012) and hence may be used with EDSS. In this experiment, however, our goal was to enable stakeholders to access robust legacy models in a user-friendly manner.

Leveraging recent technological advancements, a web-based EDSS was developed for the Occoquan Reservoir and its tributary watershed (Occoquan system) in Northern Virginia. The EDSS was designed to work with the existing water resources models for the Occoquan system in order to enhance stakeholder interaction with the watershed modeling software. Using the EDSS, stakeholders may work with complex legacy water resources models and evaluate their management solutions using a system similar to the 'two-level' EDSS suggested by Rizzoli & Young (1997). A two-level EDSS is a tool designed by experts for stakeholder involvement. The system described here is the first step in a broader experiment to make water resources tools and data more easily available for the stakeholders that was driven by local stakeholders, such as representatives of the water utilities, and local jurisdictions overseeing the monitoring and modeling programs. Requests for the ability to interact with water quality models to access the impact of land use changes is one such feature. Although designed for the Occoquan system, several EDSS components, such as the database and front-end Graphical User Interface (GUI), are independent of the models and can be extended to incorporate many other modeling systems for different sites. Lessons learnt from this experiment may also be very

useful for the integrated modeling community. It should be noted that a holistic decision support system for supporting land use changes will require more criteria, such as economic and other socio-political analysis than the single water resources criterion being addressed by this EDSS. It is envisioned that the next phase will extend the current system and include such modules.

THE STUDY AREA

The Occoquan Reservoir is located in the Northern Virginia, region of the Washington, DC metro area (Figure 1). The reservoir drains an area of about 1,476 km². Lake Manassas is the other major impoundment in the Occoquan watershed. The land use in the tributary watersheds for the two reservoirs consists of about 50% forest land, 30% agriculture land and 20% urban uses (Dougherty *et al.* 2007). The Occoquan Reservoir was one of the first large water supplies in the world where planned potable water reuse was implemented for improving drinking water yield. The reservoir, together with the Potomac River, continues to serve as an important part of the water supply source to more than 2 million people in Northern Virginia. The Occoquan Watershed Monitoring Laboratory (OWML) monitors water resources in the reservoir and its watershed, and maintains a database of the water resources in the reservoir and watershed system dating from 1973 to the present. The laboratory has also been involved in modeling watershed interactions for the complex watershed system through widely-accepted and open-source computer software (Xu 2005; Liu 2011). This modeling system was extended in this study to allow stakeholder interaction and simultaneous execution of the models on a local server grid. The Occoquan Reservoir is classified as eutrophic, and various efforts are underway to reduce the influx of nutrients from its contributing watershed to control further nutrient enrichment.

Occoquan Reservoir and watershed system model

The Occoquan system model was described as a complexly-linked model by Xu *et al.* (2007). This complexly-linked model was developed by linking various models to simulate the multiple reservoir watershed system. The model is

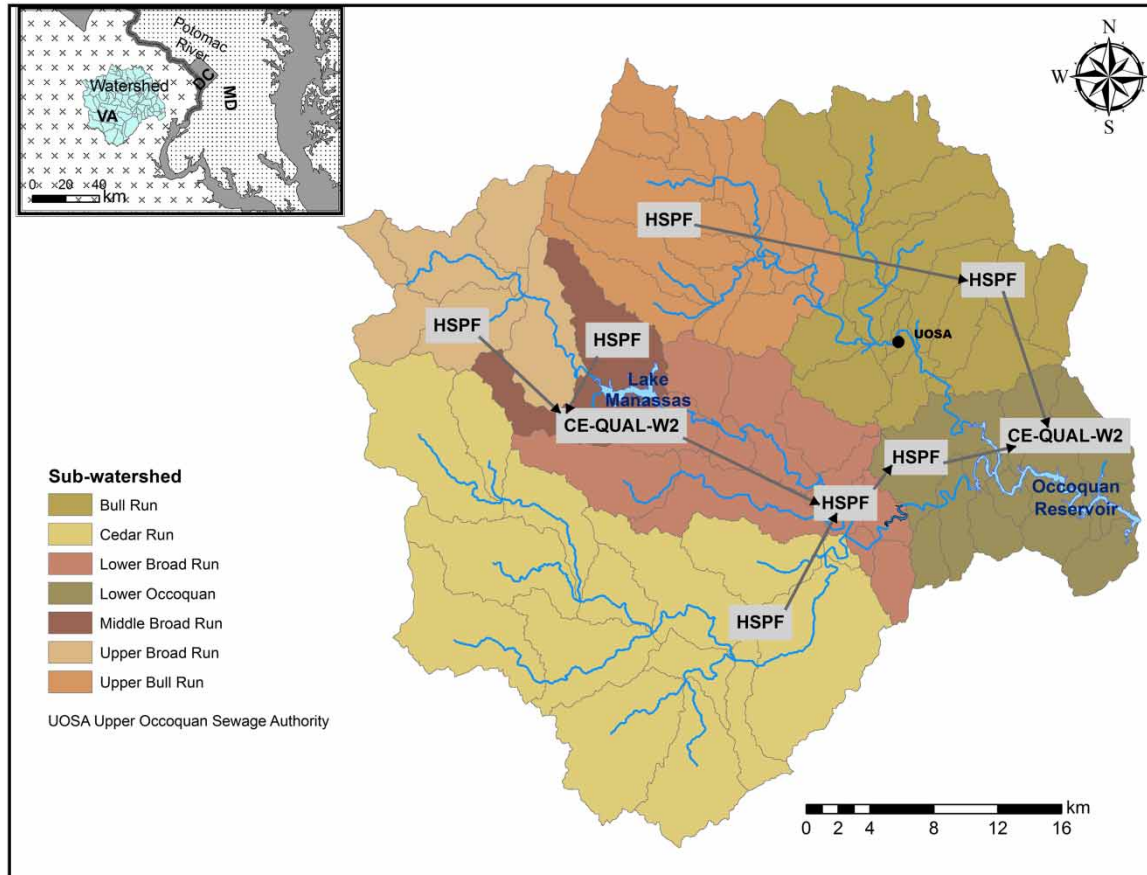


Figure 1 | Various models constituting the complexly-linked Occoquan system model with the model connections and their sub-basins.

considered to be complexly-linked because the downstream model components utilize the outputs from upstream models as input, instead of using observed data set. In addition, there are parallel branches in the model as opposed to most other traditional linearly linked system. This type of complex-linkage is expected to mimic the natural system better than simple-linkage (Xu *et al.* 2007). The newest version of the system model uses seven implementations of Hydrological Simulation Program–Fortran (HSPF) (Bicknell *et al.* 2001) and two implementations of CE-QUAL-W2 (Cole & Wells 2008) software. All HSPF models are executed at an hourly time step and CE-QUAL-W2 uses a variable time step using the largest one that is numerically stable at any point in the execution. Two prominent reservoirs in the system, Occoquan Reservoir and Lake Manassas, are simulated by CE-QUAL-W2, and their contributing watersheds are simulated by HSPF.

Seven HSPF implementations were used because HSPF has computational step limitations and the finer resolution (segmentation) required for the model exceeded the capacity of the software. Furthermore, the presence of Lake Manassas in the mid-watershed region required that the watershed model upstream of the lake feeds into the lake model, and then the lake model output feeds into a downstream HSPF model. HSPF and CE-QUAL-W2 software were chosen to simulate the Occoquan Reservoir and Watershed system, as both are open source software with a wide user base. Model interconnection standards, such as Open Modeling Interface (OpenMI) (Gregersen *et al.* 2007), were not used to connect the models. Instead, more typical file-based data sharing methods were used to share data among models. The primary reason for not using OpenMI in this system was the fact that the linked system was developed before such interconnection standards were

widely adopted and no usable OpenMI wrappers were available for HSPF and CE-QUAL-W2 software. For the Occoquan system model, upstream model outputs still drive downstream models, similar to the OpenMI system and different from a scenario where all models were acting independently. However, there is no linkage at the time step scale. For new versions of the linked system, an OpenMI based approach is being investigated using wrappers for HSPF and CE-QUAL-W2 that may also allow experimentation with different simulation software for which OpenMI wrappers already exist. It may be noted that several other modeling systems exist with their respective strength and weaknesses (Borah & Bera 2003, 2004), and the choice of modeling systems for the Occoquan system was based on continuity with the existing program and modeling efforts in the regional Chesapeake Bay Program.

HSPF is a lumped parameter, hydrological and water resources modeling software that may be used to simulate hydrological process, such as surface runoff, interflow, base flow, and snowmelt, and various water resources parameters, such as nutrient concentrations, dissolved oxygen concentrations, and sediment transport, associated with pervious land, impervious land, and well mixed waterbodies. HSPF is a flexible model, and it allows simulation of selected water resources parameters at various time scales using established empirical, theoretical, and other laboratory derived process algorithms. The smallest discrete land unit in HSPF is defined as a Segment, which may be composed of various lumped land use types that dictate the effective pervious and impervious area. CE-QUAL-W2 is a two-dimensional, fully hydrodynamic model, written in Fortran, which is widely used to simulate lakes, rivers, estuaries, and reservoirs (Cole & Wells 2008). It is a computationally robust model that ensures efficient model execution by adjusting the timestep to the maximum allowable value without violating the hydrodynamic stability requirements. It allows for simulation of the fate and transport of various physical, chemical and biological constituents of water, such as temperature, dissolved oxygen, nutrients, algae, zooplanktons, macrophytes, etc. However, CE-QUAL-W2 does not simulate overland runoff from the watershed. To simulate the conditions in the Occoquan system the models are executed in sequence as shown in Figure 1. Each sub-watershed in the figure corresponds to one HSPF implementation

and the two prominent waterbodies are represented by two CE-QUAL-W2 models. For calibration purposes, each HSPF implementation had at least one stream sampling station at or near the outlet of the sub-basin being simulated. The whole linked system, however, is also calibrated to the conditions observed in the Occoquan Reservoir. For more details on the models and the calibration procedure employed, refer to Xu (2005).

THE OCCOQUAN EDSS DESIGN

The Occoquan EDSS was designed with three key modules (Figure 2): the web-server (web module), the bridge module, and the Distributed Model Execution module (DME module). Table 1 lists the functions of each module along with key design software and key tasks performed by each module. As opposed to an integrated single system, this modular design was chosen to allow extension to a more generalized development platform. With a generalized development platform, others seeking to implement a similar system may easily change modules and adapt water resources models to generate a similar EDSS. Choosing a modular architecture may have increased the development time, but it is expected to be more usable in the long run. This EDSS has been designed for the Occoquan system, but it has the potential to act as a more general Environmental Integrated Modeling Framework (EIMF) for similar systems.

An EIMF is a set of components that may be reused to deliver an EDSS (Rizzoli *et al.* 2008). For this EDSS, the web module and the DME module are independent of the Occoquan system software and may be easily extended to any EDSS designed to analyze impacts of land use change on water resources. The bridge module, however, is specific to the Occoquan system. It uses information about the Occoquan system modeling software to receive input from the web module, process the land use change input, give direction to execute the models on the DME module, and finally, retrieve results, as required. The three modules were designed primarily using Oracle® JAVA. JAVA was chosen due to both the authors' familiarity with the programming language and the very-wide adoption of the language. Numerous other programming languages such as

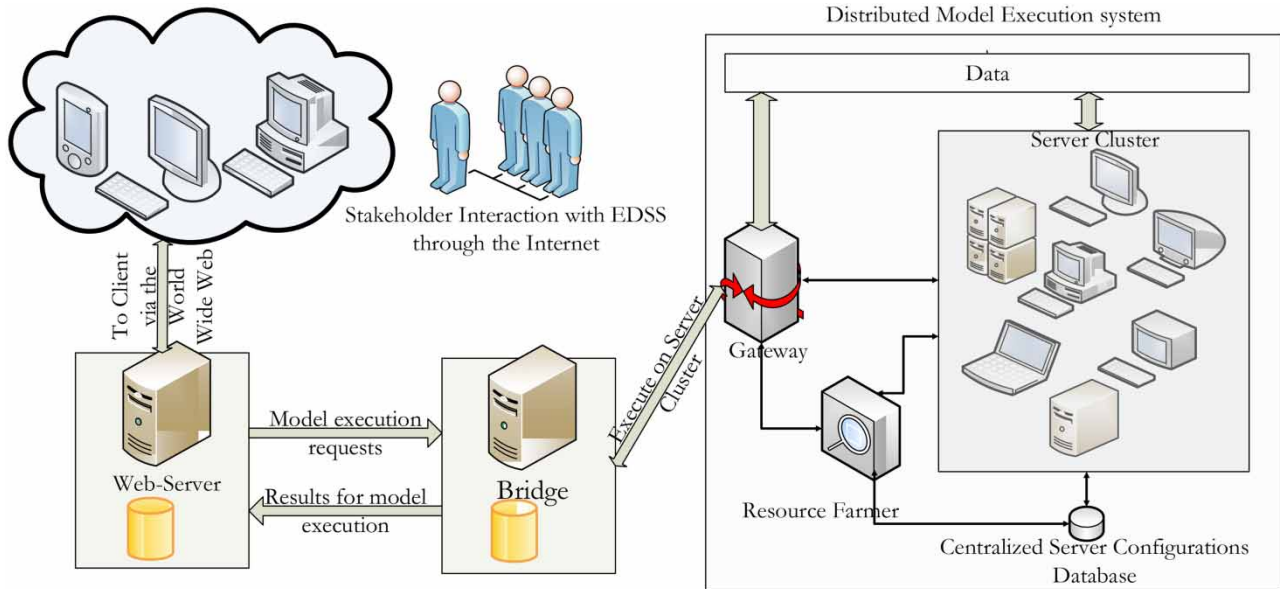


Figure 2 | The three modules that constitute the EDSS.

Table 1 | Key tasks and design software for the three main modules in the EDSS

Module	Key purposes/tasks	Key design software/packages
Web module	<ul style="list-style-type: none"> Organize and present information about Runs^a and Scenarios^b in the EDSS Generate reports with results of model executions and comparisons of Runs in different Scenarios Compare individual Run with the Base Run^c Allow stakeholder to create, copy, and edit Runs or Scenarios Manage stakeholder authentication 	Oracle Java EE 5 ArcGIS Server Java Glassfish v 2
Bridge module	<ul style="list-style-type: none"> Analyze land use changes made by stakeholders Decide if one or more models constituting the linked Occoquan system model have to be executed Execute models on DME-module Mine results after model execution and add it to the web module database 	Java SE (JDK update 19) 32 bit ArcEngine Java
Distributed Model Execution module	<ul style="list-style-type: none"> Manage local server cluster to execute models Execute models on local server cluster Return results of model execution 	Java SE (JDK update 19) Java RMI

^aRuns are individual cases where a stakeholder changes the land use.

^bA Scenario is a collection of Runs; typically, a collection of Runs in a Scenario has some common broad objective, such as Runs done to analyze possible locations of a new development may be grouped in one Scenario.

^cBase Run is the Run for 2002–2007 observed conditions.

C++, C#, and Python may have been used to develop this system. All model executions for the EDSS were carried out to simulate the impact of land use modification. Thus, all other variable inputs (except land use) were kept the

same as those observed for the 6-year period from January 2002 to December 2007, which is the calibration period for the Occoquan system model. It should be noted that the January 2002–December 2007 model set was the latest available

at the time, and the model set using the latest land use data (January 2008–December 2012) is close to being completed and will be used in the near future to better represent the current conditions.

Web module

The web module organizes all the individual land use modifications (*Runs*), available in the database, into *Scenarios*. *Scenarios* are a collection of individual *Runs*, created by users, to group one or more *Runs* for easier analysis. The EDSS allows all authenticated users (users) to view all *Runs* and *Scenarios* available in the database via a web-browser. Authentication is managed using an OpenID based system, where the OpenID URL for the user (obtained from OpenID providers such as Google, Yahoo, etc.) is added to the approved users database. Later, while logging in, to authenticate their identity users are redirected to the OpenID providers' website. Authenticated users (users) are also allowed to edit all *Runs* or *Scenarios* they have created. *Runs* or *Scenarios* created by other users may be copied and then edited. Users also have the option to create new *Scenarios* and place existing *Runs* from other *Scenarios* in them.

To modify land use, users may delineate polygons and other shapes over the current land use map or any other previously modified land use map. These land use modifications may be performed from the web browser in multiple sessions, saving as required after each session. Once all modifications are complete, users may then submit the *Run* to be analyzed. After analysis of the *Run*, results are available to view for all users of the EDSS. The web module also presents the results comparing various *Runs* in a *Scenario*. While comparing *Runs*, instead of displaying absolute magnitudes for the selected water resources parameter, relative difference from the current situation is displayed for all *Runs* compared. The results are available at various geographical scales from the whole watershed to the smallest subunit modeled. Results are grouped by years and seasons (winter, spring, summer, and fall) within the simulation period.

Selected water quality and quantity parameters are displayed as results on the web module. Nitrogen and phosphorus nutrient loads represent major concerns for the watershed, hence, species of nitrogen and phosphorus

were chosen as water resources indicators to analyze the impact from land use modification. Dissolved forms of total inorganic nitrogen (TIN) and orthophosphate-phosphorus (OP) were chosen to represent nitrogen and phosphorus, respectively, as the Occoquan system models simulate these constituents consistently better than other forms of nitrogen and phosphorus. Also, these two dissolved inorganic forms (TIN and OP) are readily available for algal uptake. Along with nitrogen and phosphorus fluvial loads, hourly peak streamflow (>90th percentile) and dissolved oxygen (DO) criteria violations in the reservoirs may be analyzed. The DO criterion is considered to be violated if in an unstratified reservoir or in the epilimnion of a stratified reservoir the simulated DO concentration is less than 4 mg/L. The epilimnion in a stratified reservoir is defined in the EDSS, as the section of reservoir above the thermocline, where the temperature gradient is more than 1°C/m. This methodology is similar to the assessment method used by the Commonwealth of Virginia to classify a reservoir as impaired for DO by the Virginia Department of Environmental Quality (VADEQ 2010). For the purpose of comparison, the violations are computed at the outlet of the reservoir based on the modeled daily average output and summed to get the annual DO criteria violations. Although only the above-mentioned water resources parameters were selected for analysis at this time, other parameters and water resources indicators (simple or complex) may be incorporated into the system. Using a knowledge management system similar to the one described by Chau (2007), expert knowledge may also be used to design indicators.

Bridge module

The bridge module analyzes the *Runs* submitted by users for execution. It uses the area and location of land use modifications to determine if any similar *Run* exists in the database. If a similar *Run* is found, then the results for that *Run* are used for analysis. Otherwise, the Occoquan system modeling software is executed, and new results are added to the database. Only the models simulating the portion of the watershed downstream of the area where changes were made are executed, as no water resources changes are expected in the upstream region. For the

Occoquan system models, a *Run* is deemed as similar to another *Run* if the percentage difference in the area for each land use types within all segments is less than 5%. For this study the minimum 5% land use change within each segment to ascertain similarity was based on the expert judgement for the Occoquan modeling system. Besides expert judgement, other more sophisticated methods for specifying this threshold may depend on sensitivity analyses of the type carried out by Liu (2011) or base it on type of change (from pasture to industrial, for example). In the linked scenario, sub-basin models where the changes are below the threshold may be deemed similar and stored results for the sub-basin may be used, while others generally require additional model execution. In situations where some (but not all) of the models are similar, all the models downstream to the non-similar sub-basin model are executed. If multiple similar *Runs* exist, then the first similar *Run* found is utilized. This method of identification of similar model takes advantage of HSPF's lumped parameter modeling property. However, by utilizing appropriate spatial relationships, this method may be extended to other watershed modeling software.

Distributed model execution module

The DME module executes water resources models as required by the bridge module on a server cluster. This module has three core components: the server cluster, the resource farmer, and the gateway (Figure 2). Other supporting components, such as configuration databases and bulk data transfer mechanism that may be required for some operations, are also incorporated in this module. The server cluster is constituted of computers with disparate hardware and software configurations available on the local network. These may include machines dedicated for executing water resources models and other computers that may be only available to execute models during off-time hours. All machines that participate in the server cluster have a server program running on them that enables communications with other components of the DME module. The gateway is a dedicated computer that acts as a fixed router for the bridge module. The bridge module requests all model executions and retrieves results from the DME module using the gateway. The resource farmer

typically is a separate thread running on the gateway machine; however, it may also be implemented on a separate machine. The resource farmer identifies the status, such as processor load, and participation times of each machine in the server cluster and also acts on the request of the gateway to identify the best model to execute. Figure 3 shows schematically the activity diagram for the process of execution on the gateway. 'Process Model' sub-activity from Figure 3 is detailed in Figure 4, illustrating the sequential steps taken by the gateway to execute any software. While processing the model, the gateway selects the server to execute the model on by querying the resource farmer. Figure 5 illustrates the communications between different components for the 'Setup Model on Server' sub-activity of Figure 4.

To execute a model on the server cluster a JAVA wrapping class was created for each of the nine possible models for the Occoquan system model (seven HSPF and two CE-QUAL-W2). These wrapping classes standardize

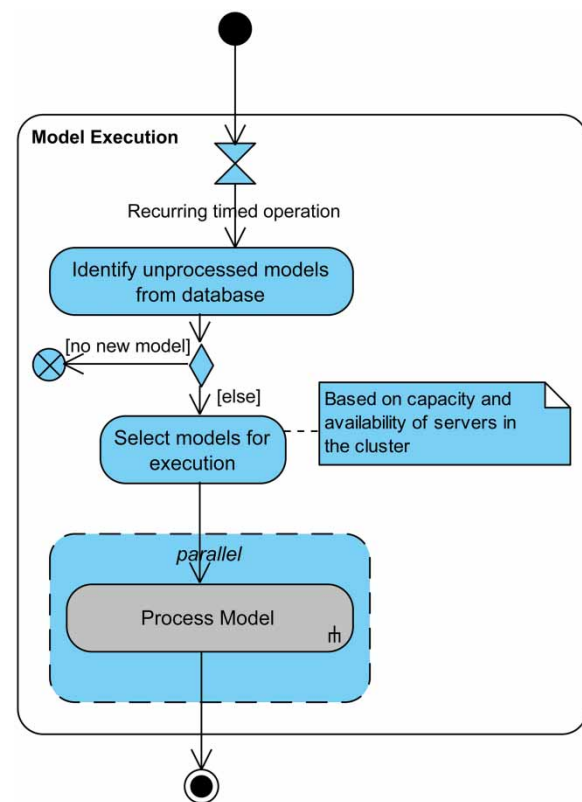


Figure 3 | Overall activity diagram for the gateway, executing models in parallel on the server cluster.

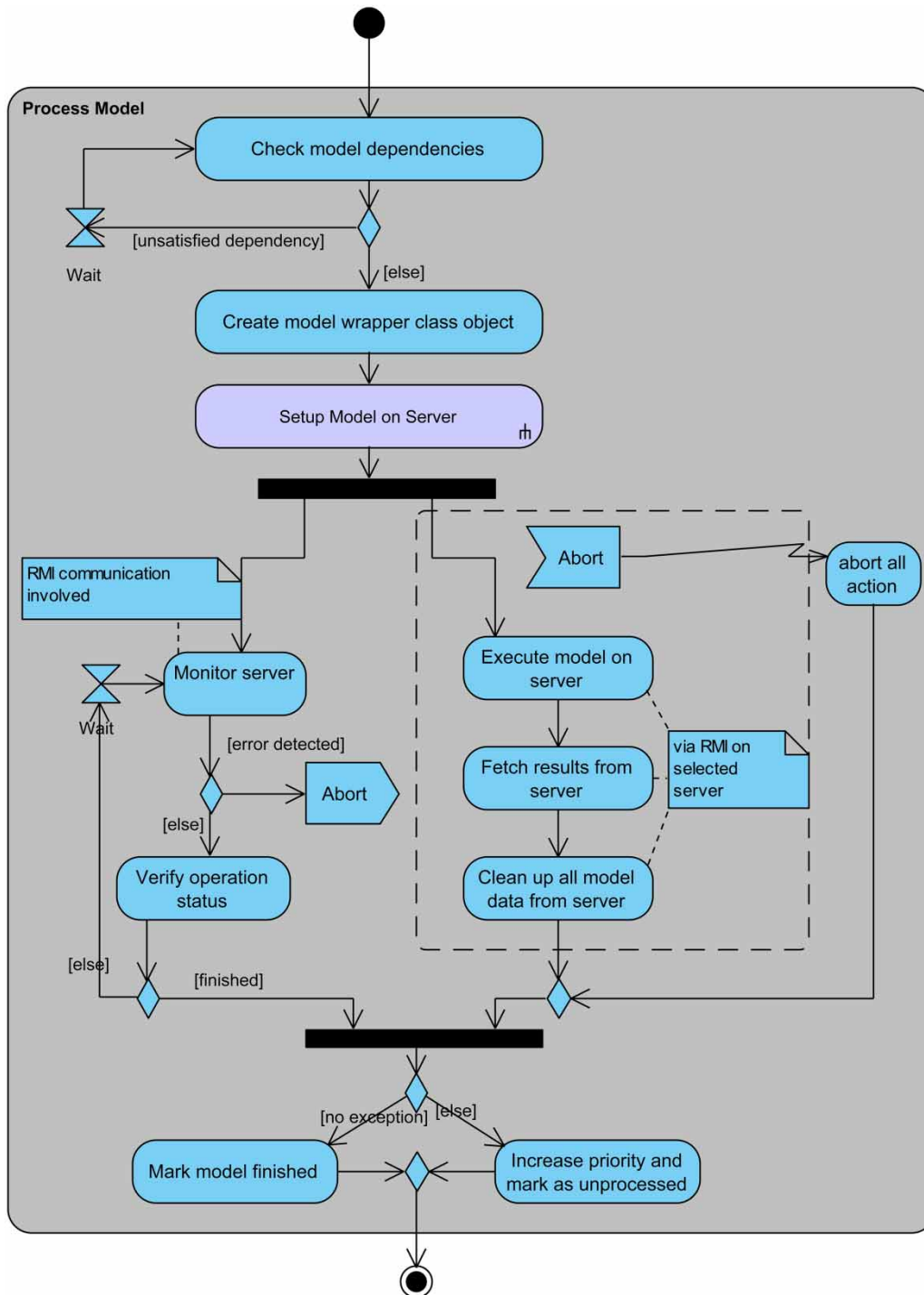


Figure 4 | Sequential steps taken by the gateway to execute any model.

the input, output, software dependencies (HSPF or CE-QUAL-W2) and execution procedures. To extend this system to other modeling software (for example SWMM),

a similar wrapping class may need to be developed. All these wrapping classes expose the following five basic methods that enable model execution on the cluster:

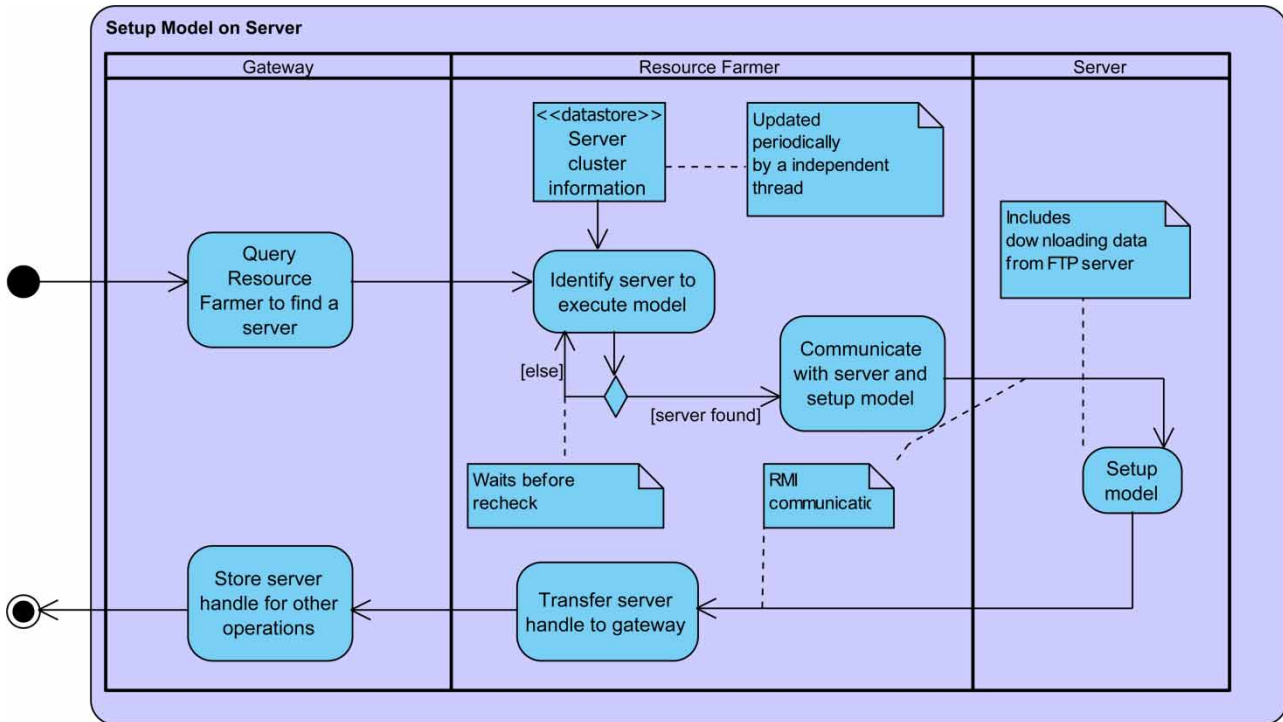


Figure 5 | Illustration of the 'Setup Model on Server' activity involving communication between the three components.

1. Setup method: downloads codebase including executable file if required, performs input conversions, and ensures all the required external libraries (such as HSPF software) are present.
2. Execute method: starts execution of the modeling software on the selected server by issuing operating-system specific commands.
3. Stop execution: stops execution and cleans up all the resources used. If required, this process may use the operating system's specific kill thread command to kill non-responding applications.
4. Return results: returns results to the gateway and communicates errors encountered, if any.
5. Find current execution status: identifies current model execution status based on processor and memory loads for the thread executing the model along with other model specific parameters such as output file size.

For a typical model execution on the DME module, once an unprocessed model has been identified (Figure 4), the gateway creates a model wrapping object for the model using the inputs provided by the bridge module. The gateway then

queries the resource farmer to identify a suitable server to execute the model. The resource farmer queries the server configuration database for suitable servers that may be used to execute the model. This list of suitable servers is developed based on software installed on the server (for example HSPF), comparison of maximum models allowed to be executed on the server at one time with current executions, and comparison of typical time required to execute the model with time left in the availability of the server to participate in the cluster. From the list of available possible servers, the server that took the least time to execute this model (based on past execution data) is used by the resource farmer as the appropriate server. The resource farmer also executes the setup method implemented by the wrapper class on the selected server using JAVA remote method invocation (RMI) protocol (Figure 5). A handle to the selected server is then returned to the gateway, which manages further model execution and result retrieval (Figure 4) using RMI protocol. A new thread is created on the gateway dedicated to managing execution of the model on server. More detail about this module including error handling and scheduling information is available from Kumar et al. (2014). There are other components in the

DME module which allow configuration of the servers and enable data transfer. The bulk of the data are transferred between the gateway and server cluster using a File Transfer Protocol (FTP) server and clients.

This configuration is close to the traditional three-tiered server cluster, where the client sends all requests and receives responses from a common gateway, different from systems where the gateway may just act as a switch that identifies the server in the cluster to process requests and all future request responses are handled directly by the server (Tanenbaum & Van Steen 2007). Because the client always requests and gets a response from the same fixed gateway, the server cluster may be hidden from the client as opposed to interacting with the assigned server directly, allowing for easier client application development and enhanced security for the server cluster. A drawback of the three-tiered architecture is that a failure of the gateway may render the whole cluster unusable, and redundant gateways may be required to overcome this drawback. Some other open-source and proprietary frameworks also provide similar grid execution functionalities such as HTCondor (Thain *et al.* 2005) and Apple® X-GRID; however, most similar frameworks are complex requiring significant configurations. For this test application, a simpler DME module, described above, was used. A typical execution of the Occoquan system model in series takes about 6 hours on a dedicated desktop machine. Since the Occoquan system model is made by connecting nine models, some time saving may be obtained by running the models in parallel (when possible) as in the case with the DME module. Further, the DME module may be particularly beneficial in obtaining faster results for cases when multiple model execution requests are received from the web. The ability of the cluster to dynamically add or drop machines further adds to the extensibility and scalability of the EDSS when occasionally model execution requests spike.

TYPICAL APPLICATION OF OCCOQUAN EDSS: NEW DEVELOPMENT UPSTREAM OF LAKE MANASSAS

Lake Manassas is the primary water supply source for the City of Manassas and other areas around the city. The Lake Manassas watershed, which lies entirely inside the

Occoquan Reservoir watershed, faces increasing development pressure due to its proximity to Washington DC (DC) and easy accessibility from DC through Interstate 66. In this example, four different spatial arrangements (*Runs* 1 to 4) for new developments were analyzed (Figure 6). All four arrangements increased the urban land uses by the same amount. The spatial locations of these four land use modifications, however, were different. At different locations, the difference in the nature of the base land use replaced and other factors related to flow routing are expected to change the amount of net nutrients flux exported to the lake. To estimate these differences in fluvial loads exported, fluvial loads for OP and TIN were analyzed at the point of entry into the lake and compared with the base (pre-development) scenario. The objective of the study was to use the simulation results to identify the best location, from the four considered, for the new development, along with the expected increase in the fluvial nutrient load due to these developments. Such information may be used to design appropriate on-site or off-site mitigation strategies to reduce net nutrient flux to pre-development levels.

To simulate the four test cases a new *Scenario* was created using the web interface. Subsequently, *Runs* were added to the *Scenario*. The new development proposed was delineated over the existing land use. For this *Scenario*, *Run* 1 was delineated using the tools within the EDSS, and appropriate urban land uses were assigned. *Runs* 2, 3, and 4 were created as copies of *Run* 1, with the location of the new developments moved to reflect the setting being analyzed. After completing the edits, *Runs* were marked as finished. Once marked as finished, the EDSS executed the required water resources modeling software. After completion of the model execution, results of the simulation were available for analysis.

Results analysis

Following model execution, data from the EDSS may be used to generate graphs as shown in Figure 7, which compares the annual load input to Lake Manassas. The graphs may be used to analyze the water resources changes due to the developments being considered. Year 2003 had the highest precipitation in the 2002–2007 period and, consequently,

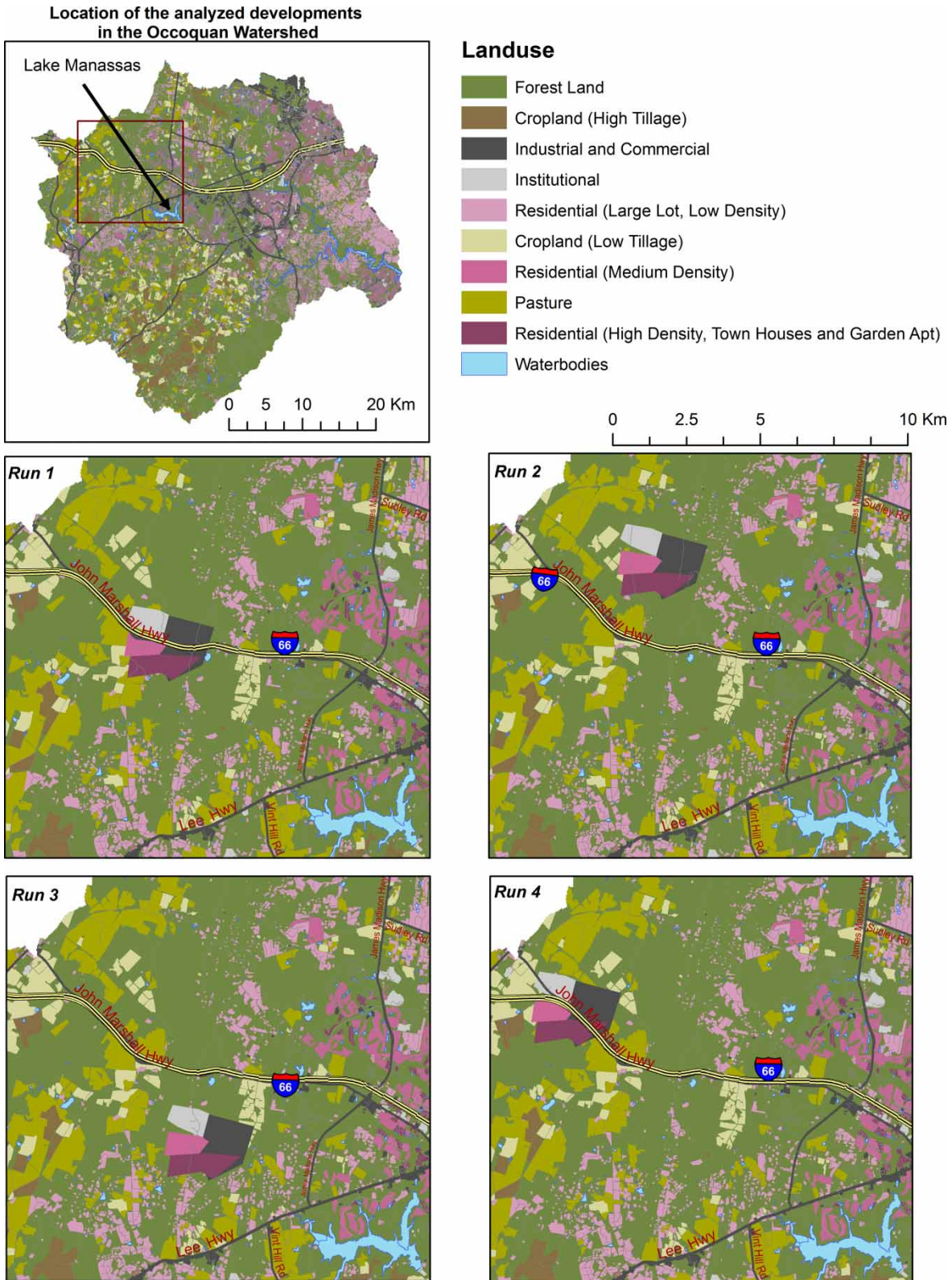


Figure 6 | Developments near Lake Manassas analyzed for water resources impact.

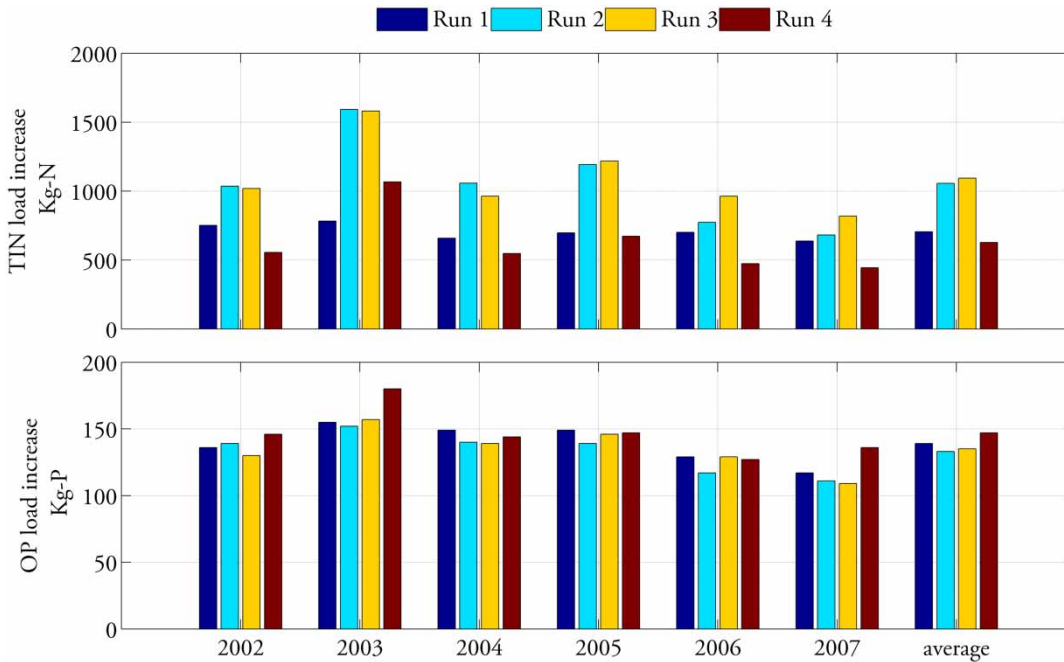


Figure 7 | The fluvial load increase observed from the development scenarios analyzed.

the fluvial load exported in that year was higher. The increase in fluvial loads from the base condition showed similar trends for both constituents examined, with the increase in both TIN and OP highest for 2003 (Figure 7). The OP fluvial load increase in all four *Runs* was similar on average and had limited value in discriminating between *Runs*. Note that OP is actively cycled in the natural environment, and similar OP loads do not imply that other species of phosphorus or total phosphorus will be similar.

TIN fluvial flux showed larger differences among the *Runs* analyzed. *Runs* 2 and 3 had higher average increases in TIN fluvial loads than *Runs* 1 and 4. Between *Run* 1 and 4 in 2003, the year with the maximum increase in fluvial nutrient loads, *Run* 1 showed a lower increase in TIN and OP fluvial load compared to *Run* 4. From these observations, it may be concluded that the land use modification represented by *Run* 1 is the better alternative from a water resources standpoint. Year 2003 had higher fluvial loads along with higher runoff volume. *Run* 1, in 2003, has the least excess nutrient export for TIN and close to the least excess export for OP among the *Runs* analyzed. This observation reinforces the choice of *Run* 1 as the preferred alternative. It may be noted that *Run* 1 replaced the least

amount of undisturbed forest area for developments among the alternatives, which might have contributed to the lower fluvial load transport in *Run* 1.

This analysis presented the simulation results for the four *Runs* from a water resources standpoint. In a real decision making scenario, stakeholders may use this information to make other changes to the proposed development and use those results along with other considerations to make an informed decision. Also, the water resources at the entrance to the lake, which was chosen to evaluate the schemes, is one of many possible ways to analyze results obtained from the EDSS. Other examiners may choose to analyze water resources at other critical locations and possibly arrive at different conclusions. It is vital to note that due to the inherent uncertainty in simulating natural processes, these results provide only an estimate of expected water resources impacts. However, in the absence of any other better estimate these results may be used for planning purposes. *Run* 1, which appears to be the best development scenario from the analysis presented, might not be viable for other reasons (e.g. financial, legal, regulatory, or other practical considerations) and another *Run* might prove to be preferable. Another important

output from the EDSS is the expected magnitude of increase in nutrient fluvial loads, which may be used to design best management practices as mitigation strategies.

DISCUSSION AND CONCLUSION

A decision support system (DSS) is a computer/technology solution to structured problems, which, in turn helps in the overall management of unstructured scenarios. The Occoquan system EDSS is a similar tool that may be used to assess the impacts on water resources due to development, and other land use changes in the watershed. This was demonstrated by the example which utilized the EDSS to assess four alternative development scenarios in the northern region of the Occoquan watershed (above Lake Manassas). As discussed earlier, the results obtained from the analysis could also be employed to design mitigation strategy that may be used to limit the extra fluvial nutrient loads expected from the chosen development. The Occoquan system EDSS provides stakeholders the ability to perform following tasks:

- modify the current land use and create *Runs*;
- group *Runs* into *Scenarios* for comparisons and further analysis;
- share *Scenarios* and *Runs* with other users;
- execute water resources models to assess the impacts of the land use modifications;
- present results in a manner that allows for comparison of various water resources indicators, for each *Scenario* and/or *Run*, at various geographical scales.

This EDSS was designed for analysis of water resources impacts due to land use modifications, and hence the features mentioned above were developed. However, the utility of the EDSS is not limited to analyzing the water resources impact of land use modifications. The system, with some additions, may be employed to analyze and educate the stakeholders about other issues, such as the impact of climate change that might be analyzed by modifying the precipitation patterns along with land use. One of the issues that have limited wide adoption of robust legacy modeling software by non-expert watershed planners has been the lack of user-friendliness (Smith Korfmacher 1998; Korfmacher 2001; Murdock et al.

2005; Koehler & Koontz 2008; Johnson 2009). Web-based EDSS may remove some of these usage barriers by making the model interface more user-friendly and intuitive. The EDSS system described here demonstrated an innovative modern user interface (UI) that may be developed for rejuvenation of legacy models and use them with modern decision support systems. The UI described in this paper borrows from contemporary popular web-designs that may be assembled using easily available software. The back end is also designed to take advantage of networked systems and reduce the time taken to execute the water resources modeling software.

Executing complex water resources models without the added functionality of the EDSS to analyze the proposed land use changes may be expensive in terms of time and monetary investment. Further, these simulations, if carried out without an EDSS, would rely heavily on water resources and modeling experts, which might subject the simulation outcomes to expert biases. Through an EDSS, the underlying tasks may be performed efficiently, and the eventual consumer of information may have more control over the operations, including designing scenarios to be analyzed, and analyzing the water resources result obtained. By providing active stakeholders with an opportunity to utilize these calibrated models directly, the EDSS may also enhance the implication of the water resources models on the decision-making process. In addition, a significant stakeholder education and outreach benefit may be expected from the EDSS regardless of the outcome and influence of the EDSS on the decision making process. In real decision-making scenarios for land use modification, this EDSS may only provide water resources impact, which is not the only criterion used by decision makers. Most such decisions require collaboration and negotiations amongst stakeholders and evaluation of several other criteria, such as economics. It may be noted that to make this system a holistic DSS, other components such as a negotiation module and MCDM support have to be added. Using an ontology-based knowledge management system, such as the one described by Chau (2007), coupled with MCDM along with negotiation and judgement modules (Jonoski 2002; Khelifi et al. 2006; Almoradie 2014), some of the present EDSS limitations may be addressed in the next phase of development.

Although advantageous for encouraging stakeholder partnership and education, the EDSS relies on the robustness of the underlying modeling software to simulate the impact on water resources from land use change. Well established HSPF and CE-QUAL-W2 software drive the Occoquan system model. These models have been widely applied for forecasting water resources as these models are based on fundamental governing equations and verified empirical relationship (Bicknell *et al.* 2009; Cole & Wells 2008). Further, the land use modifications performed through this EDSS are limited to the land use types already utilized in the Occoquan system model. No new, previously un-calibrated, parameter is thus introduced by changing land use in these simulations. Nevertheless, when applying the EDSS, it must be kept in mind that water resources results obtained through this system and any other modeling system are subject to various uncertainties. These uncertainties stem from raw data unreliability, knowledge deficiency in model software development, deficiencies in model development, validation or calibration, and various other human factors (Brugnach *et al.* 2008; Maier *et al.* 2008). Users of this system should be aware of these limitations of software models before gainfully utilizing the EDSS.

It is clear that web technologies and the internet backbone has changed from the time when most of the legacy water resources model were developed, and most of the legacy models were not designed/updated to leverage the power of the internet and modern network computing environments. It has also been demonstrated in this paper that by using modern techniques it is possible to make the robust legacy model available to stakeholders without compromising the state-of-the-art science. There is some evidence in the literature that moving models to the web and providing an easier to use interface may disguise and underplay the complexity of modeling natural systems. However, it has also been argued that the educational advantages of making such a system available to a wider audience, along with the potential to be used in multi-criteria decision-making, outweighs the risk associated with the perception of diminishing complexity. To guard against misuse of the system at this stage, only selected users have access to the system and we are developing methods with stakeholder feedback to highlight uncertainty of the modeling results.

Evaluation instruments for the EDSS are also being designed which takes these concerns into account. There are several other limitations that need to be addressed, such as derivation and communication of uncertainty to stakeholders. Most legacy models, such as HSPF and CE-QUAL-W2, do not provide confidence bounds on results, but some estimate of uncertainty and, as a corollary, confidence bounds, may be derived from sensitivity analysis. Another limitation that is inherent from the fact that this system uses web-based technologies instead of conventional software that run on a computer is the need for constant updates and security patches, both to the server backend and the front-end browser based UI. Large parts of the EDSS UI is based on JavaScript and sometimes with web-browser updates new bugs are introduced in previously working JavaScript which need to be fixed. Also, as evident from the software stack of the EDSS, several of the servers used are outdated and need updating. Although inconvenient, these updates also force the developers to refresh the system with often efficient and more contemporary design. Designing EDSS involves configuring water resources models and web-UI so that inputs can be passed to the models and results may be retrieved. One of the solutions to the problem of constantly upgrading the EDSS to keep up with underlying software upgrades might be an open source, collaborative software development model (Herbsleb & Moitra 2001), where communities that need such an EDSS may work collaboratively to develop and update the required system and also benefit from the inherent review system of collaborative development (Rigby *et al.* 2012).

Developing with open source GIS software has been problematic in the past due to availability, maturity, and ability of open source GIS software, along with expertise required to use and extend the software. However, the last few years have seen a tremendous growth in the availability and abilities of open-source GIS platforms and standardization, and, together, these have made it easier to develop using open source GIS. For example, ArcGIS server and engine, proprietary platforms, chosen for this system may be replaced with a combination of JavaScript application programming interface (API) terraformer (<http://terraformer.io/>), ESRI Geometry API (<https://github.com/Esri/geometry-api-java>) along with any of the open-source or free web-mapping platform such as Google Maps API (<https://developers.google.com/maps/>),

OpenLayers (<http://openlayers.org/>), and Leaflet (<http://leafletjs.com/>). Further, open model and data sharing standards developed by the environmental community under the aegis of the Open Geospatial Consortium (OGC) such as OpenMI, WaterML2, RiverML, and OGC Web Mapping Services (WMS), and rise of cloud computing that may allow water resources model to be executed as a web service, will tremendously benefit the next generation of web-based EDSS. This standardized water resources model independent UI shell will also allow other models to be coupled and effectively initiate collaborative software development for the EDSS UI shell.

ACKNOWLEDGEMENTS

The authors thank the diligent anonymous reviewers for their suggestions that have helped us improve this manuscript. This work is supported by the Occoquan watershed monitoring and modeling program, and we express our gratitude to the local jurisdictions and public utilities in the Occoquan watershed region for their continuing financial support.

REFERENCES

- Almoradie, A. D. S. 2014 *Networked environments for stakeholder participation in water resources and flood management*. PhD thesis. UNESCO-IHE, Institute for Water Education, Delft, The Netherlands.
- Almoradie, A., Cortes, V. J. & Jonoski, A. 2015 *Web-based stakeholder collaboration in flood risk management*. *J. Flood Risk Manage.* **8** (1), 19–38.
- Bicknell, B. R., Imhoff, J. C., Kittle Jr, J. L., Jobs, T. H., Donigan Jr, A. S. & Johanson, R. C. 2001 *Hydrological Simulation Program–Fortran: HSPF, Version 12 User's Manual*. Aqua Terra, Mountain View, CA.
- Booth, N. L., Everman, E. J., Kuo, I. L., Sprague, L. & Murphy, L. 2011 *A web-based decision support system for assessing regional water-quality conditions and management actions*. *JAWRA J. Am. Water Resour. Assoc.* **47** (5), 1136–1150.
- Borah, D. & Bera, M. 2003 *Watershed-scale hydrologic and nonpoint-source pollution models: review of mathematical bases*. *Trans. ASAE* **46** (6), 1553–1566.
- Borah, D. & Bera, M. 2004 *Watershed-scale hydrologic and nonpoint-source pollution models: review of applications*. *Trans. ASAE* **47** (3), 789–803.
- Brugnach, M., Pahl-Wostl, C., Lindenschmidt, K. E., Janssen, J. A. E. B., Filatova, T., Mouton, A., Holtz, G., van der Keur, P. & Gaber, N. 2008 *Complexity and uncertainty, rethinking the modelling activity*. In: *Developments in Integrated Environmental Assessment* (A. J. Jakeman, A. A. Voinov, A. E. Rizzoli & S. H. Chen, eds). Elsevier, Amsterdam, The Netherlands, pp. 49–68.
- Chau, K. W. 2007 *An ontology-based knowledge management system for flow and water quality modeling*. *Adv. Eng. Softw.* **38** (3), 172–181.
- Chen, W. & Chau, K. 2006 *Intelligent manipulation and calibration of parameters for hydrological models*. *Int. J. Environ. Pollut.* **28** (3), 432–447.
- Cockerill, K., Passell, H. & Tidwell, V. 2006 *Cooperative modeling: building bridges between science and the public*. *J. Am. Water Resour. Assoc.* **42** (2), 457–471.
- Cole, T. & Wells, S. 2008 *CE-QUAL-W2: A Two-Dimensional, Laterally Averaged, Hydrodynamic and Water Quality Model, Version 3.6*. US Army Corps of Engineers, Washington, DC.
- Diduck, A., Sinclair, J., Pratap, D. & Hostetler, G. 2007 *Achieving meaningful public participation in the environmental assessment of hydro development: case studies from Chamoli District, Uttarakhand, India*. *Impact Assess. Project Appraisal* **25** (3), 219–231.
- Dougherty, M., Dymond, R. L., Grizzard, J. T. J., Godrej, A. N., Zipper, C. E. & Randolph, J. 2007 *Quantifying long-term hydrologic response in an urbanizing basin*. *J. Hydrol. Eng.* **12** (1), 33–41.
- Gregersen, J., Gijsbers, P. & Westen, S. 2007 *OpenMI: open modelling interface*. *J. Hydroinf.* **9** (3), 175–191.
- Herbsleb, J. D. & Moitra, D. 2001 *Global software development*. *Software IEEE* **18** (2), 16–20.
- Irvin, R. A. & Stansbury, J. 2004 *Citizen participation in decision making: is it worth the effort?* *Public Admin. Rev.* **64** (1), 55–65.
- Johnson, M. S. 2009 *Public participation and perceptions of watershed modeling*. *Soc. Natur. Resour.* **22** (1), 79–87.
- Jonoski, A. 2002 *Hydroinformatics as Sociotechnology*. CRC Press, Rotterdam, The Netherlands.
- Khelifi, O., Lodolo, A., Vranes, S., Centi, G. & Miertus, S. 2006 *A web-based decision support tool for groundwater remediation technologies selection*. *J. Hydroinf.* **8**, 91–100.
- Kiker, G. A., Bridges, T. S., Varghese, A., Seager, T. P. & Linkov, I. 2005 *Application of multicriteria decision analysis in environmental decision making*. *Integr. Environ. Assess. Manage.* **1** (2), 95–108.
- Koehler, B. & Koontz, T. 2008 *Citizen participation in collaborative watershed partnerships*. *Environ. Manage.* **41** (2), 143–154.
- Korfmacher, K. S. 2001 *The politics of participation in watershed modeling*. *Environ. Manage.* **27** (2), 161–176.
- Kumar, S., Godrej, A. & Grizzard, T. 2014 *Using locally distributed computing to aid water quality modeling*. In: *Proceedings of the 7th International Congress on Environmental Modelling and Software* (D. P. Ames, N. W. T. Quinn & A. E. Rizzoli, eds). San Diego, CA, USA.
- Leach, W. D., Pelkey, N. W. & Sabatier, P. A. 2002 *Stakeholder partnerships as collaborative policymaking: evaluation*

- criteria applied to watershed management in California and Washington. *J. Policy Anal. Manage.* **21** (4), 645–670.
- Li, X.-Y., Chau, K. W., Cheng, C.-T. & Li, Y. S. 2006 A Web-based flood forecasting system for Shuangpai region. *Adv. Eng. Softw.* **37** (3), 146–158.
- Liu, Y. 2011 *Effective Modeling of Nutrient Losses and Nutrient Management Practices in an Agricultural and Urbanizing Watershed*. PhD Dissertation, Virginia Tech, Manassas, Virginia.
- Maier, H. R., Ascough II, J. C., Wattenbach, M., Renschler, C. S., Labiosa, W. B. & Ravalico, J. K. 2008 Uncertainty in environmental decision making: issues, challenges and future directions. In: *Developments in Integrated Environmental Assessment* (A. J. Jakeman, A. A. Voinov, A. E. Rizzoli & S. H. Chen, eds). Elsevier, Amsterdam, The Netherlands, pp. 69–85.
- McIntosh, B. S., Giupponi, C., Voinov, A. A., Smith, C., Matthews, K. B., Monticino, M., Kolkman, M. J., Crossman, N., van Ittersum, M., Haase, D., Haase, A., Mysiak, J., Groot, J. C. J., Sieber, S., Verweij, P., Quinn, N., Waeger, P., Gaber, N., Hepting, D., Scholten, H., Sulis, A., van Delden, H., Gaddis, E. & Assaf, H. 2008 Bridging the gaps between design and use: developing tools to support environmental management and policy. In: *Developments in Integrated Environmental Assessment* (A. J. Jakeman, A. A. Voinov, A. E. Rizzoli & S. H. Chen, eds). Elsevier, Amsterdam, The Netherlands, pp. 33–48.
- Murdock, B. S., Wiessner, C. & Sexton, K. 2005 Stakeholder participation in voluntary environmental agreements: analysis of 10 Project XL case studies. *Sci. Technol. Human Values* **30** (2), 223–250.
- Muttill, N. & Chau, K.-W. 2006 Neural network and genetic programming for modelling coastal algal blooms. *Int. J. Environ. Pollut.* **28** (3), 223–238.
- Reed, M. S. 2008 Stakeholder participation for environmental management: a literature review. *Biol. Conserv.* **141** (10), 2417–2431.
- Rigby, P. C., Cleary, B., Painchaud, F., Storey, M. & German, D. M. 2012 Contemporary peer review in action: lessons from open source development. *Softw. IEEE* **29** (6), 56–61.
- Rizzoli, A. E. & Young, W. J. 1997 Delivering environmental decision support systems: software tools and techniques. *Environ. Modell. Softw.* **12** (2–3), 237–249.
- Rizzoli, A. E., Leavesley, G., Ascough II, J. C., Argent, R. M., Athanasiadis, I. N., Brilhante, V., Claeys, F. H. A., David, O., Donatelli, M., Gijssbers, P., Havlik, D., Kassahun, A., Krause, P., Quinn, N. W. T., Scholten, H., Sojda, R. S. & Villa, F. 2008 Integrated modelling frameworks for environmental assessment and decision support. In: *Developments in Integrated Environmental Assessment* (A. J. Jakeman, A. A. Voinov, A. E. Rizzoli & S. H. Chen, eds). Elsevier, Amsterdam, The Netherlands, pp. 101–118.
- Smith Korfmacher, K. 1998 Water quality modeling for environmental management: lessons from the policy sciences. *Policy Sci.* **31** (1), 35–54.
- Steiniger, S. & Hunter, A. J. S. 2012 The 2012 free and open source GIS software map – A guide to facilitate research, development, and adoption. *Comput. Environ. Urban* **39**, 136–150.
- Tanenbaum, A. S. & Van Steen, M. 2007 *Distributed Systems: Principles and Paradigms*. Pearson Prentice Hall, Upper Saddle River, New Jersey.
- Taormina, R., Chau, K.-W. & Sethi, R. 2012 Artificial neural network simulation of hourly groundwater levels in a coastal aquifer system of the Venice lagoon. *Eng. Appl. Artif. Intell.* **25** (8), 1670–1676.
- Thain, D., Tannenbaum, T. & Livny, M. 2005 Distributed computing in practice: the Condor experience. *Concurr. Comput.* **17** (2–4), 323–356.
- Tuthill, D., Ciscell, M., Petrich, C., Kissinger, B. & Oakleaf, J. 2000 Migration from tabular to spatial data analysis techniques for water management in Idaho. *J. Hydroinf.* **2**, 183–195.
- USEPA 2001 *Stakeholder Involvement & Public Participation at the U.S. EPA: Lessons Learned, Barriers, & Innovative Approaches (EPA-100-R-00-040) Rep.* Office of Policy, Economics, and Innovation, US Environmental Protection Agency, Washington, DC.
- USEPA 2008 *Handbook for Developing Watershed Plans to Restore and Protect Our Waters (EPA-841-B-08-002), report Rep.* Office of Water, Nonpoint Source Control Branch, US Environmental Protection Agency, Washington, DC.
- USEPA 2012 Superfund Community Involvement, edited, US Environmental Protection Agency. www.epa.gov/superfund/community/.
- VADEQ 2010 *Water Quality Assessment Guidance Manual (Memo No. 09-2006) Rep.* Water Division, Department of Environmental Quality (DEQ), Commonwealth of Virginia, Richmond, Virginia.
- Voinov, A. & Bousquet, F. 2010 Modelling with stakeholders. *Environ. Modell. Softw.* **25** (11), 1268–1281.
- Wu, C. L., Chau, K. W. & Li, Y. S. 2009 Predicting monthly streamflow using data-driven models coupled with data-preprocessing techniques. *Water Resour. Res.* **45** (8), W08432.
- Xu, Z. 2005 *A complex, linked watershed-reservoir hydrology and water quality model application for the Occoquan watershed, Virginia*. PhD Dissertation, Virginia Tech, Falls Church, Virginia.
- Xu, Z., Godrej, A. N. & Grizzard, T. J. 2007 The hydrological calibration and validation of a complexly-linked watershed-reservoir model for the Occoquan watershed, Virginia. *J. Hydrol.* **345**, 167–183.

First received 29 December 2014; accepted in revised form 11 June 2015. Available online 16 July 2015