An Overview of Non-Point Source Pollution Modeling: Current Status and Future Prospect

Hossein Ahmadi

Department of Biological Systems Engineering, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA

ABSTRACT
Non-point source (NPS) pollution control is an increasingly important topic in aquatic environment protection due to its significant adverse effect on water quality. Characterizing NPS pollution loads is crucial for preparing BMPs to improve water quality and protect the aquatic ecosystem. Modeling tools are commonly used for simulating NPS pollution in watersheds; however, the modeling process contains uncertainties and complications, which makes the prediction of NPS pollution loads challenging and complicated. To deal with this issue, significant progress has been made to address NPS pollution modeling problems in the past few decades. The current study reviews different approaches being used for NPS pollution modeling. In this context, the main methods in NPS modeling are described and classified into three categories. 1. Empirical models, 2. Physically based models, and 3. Simulation based optimization models. The present study contributes to fulfilling the gaps in the classification of NPS pollution modeling approaches and highlights the advantages and drawbacks of each approach in order to set a standard for choosing a proper simulation tool for estimating NPS pollution loads based on the limitations and requirements of the understudy circumstance.

Keywords: Non-Point Source Pollution, Modeling, Empirical Models, Physically Based Models, Simulation Based Optimization Models

Introduction
Non-point source (NPS) pollution refers to diffuse contamination, which is not discharged from a few localized points or single indefinable sources. Sources of NPS pollution include agricultural, mining, forest logging, urban runoff and stream bank erosion [1]. NPS pollution resulting from land runoff, precipitation, atmospheric deposition, drainage, seepage or hydrologic modification is commonly intermittent and generally correlated with meteorological events, including precipitation and runoff, and land characteristics such as soil properties, and topography [2-4].

NPS pollution, in contrast to point source pollution, is characterized by unpredictable occurrence, complicated mechanisms and processes, variable spatial and temporal pollution, and challenges in monitoring, simulation, and control. These properties and characteristics of NPS pollution makes modeling of it complex and challenging [5,6]. Nonpoint source (NPS) pollution is the major cause of impairment of US surface waters It harms the aquatic ecosystem and greatly reduces water quality which leads to a decrease the capacity of natural water resources for drinking water and recreation purposes. On the other hand, controlling this type of pollution is difficult, and NPS pollution is usually controlled through prevention rather than treatment [7,8]. With the rapid development of agricultural technology and urbanization in the previous decades, NPS pollution has turned into an important side effect of agricultural production and urbanization [9-12]. The amount of NPS pollution and its grave consequences on the environment and human health increase yearly around the globe due to the development of agricultural technologies, putting more land under cultivation, the rising use of chemical fertilizers and pesticides, and urbanization. Thus, NPS pollution has become a primary threat to surface water quality and evolve into the primary contributor to water-related problems such as water contamination, aquatic ecology deterioration and eutrophication [13-17]. Proper management of agricultural and urban runoff is a large concern for the U.S. Environmental Protection Agency (USEPA) and the U.S. Department of Agriculture (USDA) [18]. In recent decades, agricultural and rural NPS pollution has become the leading contributor to water quality degradation across the world, resulting in the importance of controlling agricultural and urban NPS pollution loads in protecting the aquatic environment pollution has remarkably increased [19-21]. Identifying NPS pollution characteristics, tracking NPS pollutants pathways, and estimating the NPS pollution loads in a watershed greatly aid researchers to get an acute insight into the entire processes of NPS pollutants and determine the complete impact of NPS pollution in order to control water pollution and create and implement BMPs. However, tracking NPS pollutants is largely difficult from production to the final fate [22-24]. Modeling is a common tool for estimating NPS pollution loads, and in order to control NPS pollution having an accurate model to predict NPS pollution is essential. NPS pollution models simulate the spatial and temporal variation of NPS pollution by considering the entire basin system and the all complicated pollution-generating process. These simulations also
evaluate the effects of various BMPs on controlling NPS pollution, thereby providing a basis for environmental management plans [25]. Various approaches have been widely used throughout the last decades to precisely estimate NPS pollution in watersheds. These approaches include empirical modeling, physically based modeling, and simulation-based optimization modeling. There is a large body of literature associated with NPS pollution modeling; nevertheless, little direction is in place selecting and applying the appropriate approach to simulate NPS pollution under different circumstances. Therefore, this work attempts to fill the mentioned gaps by classifying the NPS pollution modeling approaches to help researchers choose the suitable method and model to estimate NPS pollution loads. The present paper reviewed and categorized models which have been generally used for NPS pollution modeling.

**Classification**

The hydrological models commonly used for NPS pollution modeling could be clustered differently based on various criteria. In this paper, the approaches to simulate NPS pollution have been classified as follows:

1. **Empirical models**
2. **Physically based models**
3. **Simulation based optimization models**

A schematic of the categorizing different models for the simulation of NPS pollution is shown in Figure 1.

**Empirical Models**

Empirical modeling is based on observations rather than on a mathematical equation to describe the behaviour of the system; therefore, empirical models benefit data and observations of the system to drive a specific pattern to characterize the hydrological parameters [26]. These models, which are also known as a black box, data-driven models, neglect some processes and parameters to simplify NPS pollution modeling processes [17]. The principal benefits of these models are they need a lower demand for input data and consist of a rather simpler calculation process. However, these models cannot adequately reflect the contamination migration process and cannot be applied to regional scale problems[27]. Empirical modeling is not an option perfect for the prediction of NPS pollution loads in basins where complex land cover and geomorphic units exist [28].

Empirical modeling techniques include three different categories statistical methods, export coefficient model, and hydrograph separation methods.

**Statistical Methods**

Statistical methods were developed based on simultaneous monitoring data for water quality and quantity in the runoff. This method’s fundamental assumption is to ignore the actual pollutant migration process on the surface and calculate the pollutant concentration based on the quality of receiving waters [29]. As there are a wide variety of variables associated with affecting water quality, statistical analysis has become a powerful tool to study water quality. A large number of investigations have been carried out to study NPS pollution in recent years with statistical methods [30,31]. As the statistical method requires a large deal of data, it is appropriate for a watershed with adequate data in order to estimate the regional pollution load sufficiently accurately. Thus, the utilization of this method is limited to some specific cases since this method is data-intensive and expensive.

**Export Coefficient Mode (ECM)**

The export coefficient model (ECM) is based on the concept that the nutrient load exported from a basin is the sum of the produced by catchments with different land-use types [32,33]. This simple and relatively efficient method is widely used for simulating NPS pollution loads according to large time steps (monthly or annual) on a watershed scale. In recent decades, a great number of ECM based studies have been carried out for simulating NPS pollution [34-37]. The ECM model has a simple and straightforward layout, less number of parameters and easier operation [38,39]. However, this method suffers from not considering some influence factors of NPS pollutants, such as climate, topography, soil type, land cover type, and other human activities [40]. To address the drawbacks of ECM a method known as the improved export coefficient modeling method (IECM) was created in order to improve the ECM by accounting for the effects of the temporal-spatial heterogeneity of precipitation and terrain on NPS pollution, and applying other characteristics of watersheds such as climate, sediment, nutrient decay, soil erosion, or bioactive ingredients of fertilizer and pesticide. Various scholars and researchers have used IECM for predicting NPS pollution loads [20,41-48]. The ECM and IECM approaches have been widely accepted as methods to estimate NPS pollution loads in watersheds with sufficient accuracy, benefiting from advantages such as limited input data requirements, fewer parameters, and easy operation. These methods can be scaled up to a regional scale, in addition, these methods are particularly appropriate for large watersheds, in which observed data are inadequate. Their export coefficients, however, are fundamentally quite varied and reflect unique site circumstances for each case and cannot extend to other cases.

**Hydrograph Separation Method**

This method uses runoff hydrographs to calculate point source (PS) and non-point source (NPS) pollution. In this method runoff hydrograph is divided into the base flow and storm flow to estimate the point source (base flow) and NPS loads (storm flow) [49]. In order to separate baseflow from total storm flow several methods have been presented [50,51]. The methods of hydrograph separation can be differentiated into two main groups: graphical approaches and filtering approaches [52]. Graphical approach is based on stream flow data [53,54], and emphasizes on determining the points in which base flow intersects the rising and falling limbs of the quick flow response. The recursive digital filtering approach is another method to quantify base flow contributions in which data processing of the entire stream hydrograph derives a base flow hydrograph [55,56]. In recent years the hydrograph-separation approach has been developed and utilized in various for predicting agricultural and urban NPS pollution loads [57,58].
Overall, this technique is quite straightforward and broadly used in the hydrological field; however, the storm flow and base flow separation does not distinguish between NPS (storm flow) and point source (base flow). In addition, this approach is often restricted to small basins with a short amount of data.

**Physically Based Models**
The physically based models are based on some mathematical equations characterizing the physics of the hydrological processes. These equations are spatially and temporally discretized to describe the hydrological phenomena over time and space. Deterministic models based on spatial complexity aspect are classified as lumped, distributed, and semi-distributed models [26]. In the last years, deterministic physically-based modeling (more often called simply physically-based modeling) has received much attention for NPS pollution modeling [59-62]. Physically based models incorporate hydrological model, soil erosion model and pollutant transport model into a generally a complete model system providing quantities and qualitative description of NPS pollution. These models also referred to as “white-box” models, take into account the fundamental mechanism of the pollution process and consider distinct spatial and temporal feature distributions. These models can be applied to large-scale studies; however, it needs a large body of data to calibrate the models [29]. Physically based models can be divided into single event simulation and continuous simulation models.

**Single Event Simulation**
The single event modeling uses simple equations to simulate hydrological processes. This type of modeling is easy to run; nevertheless, it cannot consider the variability of parameters and several variables such as soil moisture conditions are assumed; in addition, single event models are not generally applicable for long term simulation [63].

Table 1 shows the single event models which have been widely used for modeling NPS pollution and present their properties.

<table>
<thead>
<tr>
<th>Model</th>
<th>Temporal scale</th>
<th>Watershed delineation</th>
<th>Runoff</th>
<th>Erosion</th>
<th>Sediment</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGNPS</td>
<td>Storm event</td>
<td>Hydrological unit</td>
<td>Curve number</td>
<td>USLE</td>
<td>Sediment routing</td>
</tr>
<tr>
<td>ANSWERS</td>
<td>Storm event</td>
<td>Hydrological unit</td>
<td>Manning</td>
<td>ANSWERS</td>
<td>Storm event</td>
</tr>
<tr>
<td>CASC2D</td>
<td>Long term</td>
<td>2D grid</td>
<td>Diffusive wave</td>
<td>USLE</td>
<td>Not simulated</td>
</tr>
</tbody>
</table>

Among the models shown in Table 1, the AGNPS model and ANSWERS model are considerably used in recent studies.

**AGNPS**
AGNPS (Agricultural Non-Point Source Pollution Modeling System) is developed by USDA for the NPS pollution modeling in rural areas. This event-based model simulates runoff, sediment, and nutrient transport from agricultural watersheds. This model deploys cells to cover the computational domain. These cells, which are uniformly square areas represent the watershed and enable considering and defining features as a point within a watershed [64]. The AnnAGNPS (Annualized Agricultural Non-Point Source Pollutant Loading Model) is the improved version of the AGNPS model which is based on continuous simulation [65]. AGNPS and AnnAGNPS models have considerably been applied to different watersheds to simulate hydrological processes and NPS pollution loads [66-68].

The drawbacks of the AGNPS model are requiring intensive input data, incapable of simulating pollutant transformations, and not considering baseflow. [63]

**ANSWERS**
ANSWERS (Areal Nonpoint Source Watershed Environment Response Simulation) is a distributed, and event-based model developed for estimating the impacts of land use on NPS pollution loads. This model utilizes a distributed parameter concept to model spatially variable runoff, seepage, underground drainage and erosion. ANSWERS-2000 is the enhanced and continuous version of the ANSWERS model developed by at Virginia Tech. ANSWERS-2000 continuously simulates nutrient load within a watershed. In addition, this model can consider different BMPs (agricultural and urban) for decreasing sediment and nutrient delivery to streams and leaching of nitrogen. The incapability of simulating chemical processes, requiring intensive computation calculations, and sensitivity to input data are the main downsides of the ANSWERS model.

**Continuous Simulation**
Continuous hydrologic modeling is an approach to simulate the entire hydrological cycle by considering different parameters such as soil type, moisture, and storage. This model increases the simulation accuracy by taking into account historic hydrological events. Continuation models are generally used to model hydrological processes over longer periods of time such as months and even years, to consider all the precipitation-runoff events during the period [63].

The important continuous simulation models for NPS pollution are shown in Table 2.

<table>
<thead>
<tr>
<th>Model</th>
<th>Temporal scale</th>
<th>Watershed delineation</th>
<th>Runoff</th>
<th>Erosion</th>
<th>Sediment</th>
</tr>
</thead>
<tbody>
<tr>
<td>SWAT</td>
<td>Long term</td>
<td>Basin and subbasins</td>
<td>Curve number</td>
<td>MUSLE</td>
<td>Bagnold’s stream power</td>
</tr>
<tr>
<td>HSPF</td>
<td>Long term</td>
<td>Basin and segment</td>
<td>Empirical equation</td>
<td>Splash detachment and wash off</td>
<td>Toffaleti or Colby methods</td>
</tr>
<tr>
<td>MIKE SHE</td>
<td>Long term</td>
<td>2D grid</td>
<td>Diffusive wave</td>
<td>NI</td>
<td>NI</td>
</tr>
</tbody>
</table>

The SWAT model and HSPF models are considerably used for the simulation of water quantity and quality, and for investigating the effect of different BMPs on water quality in a watershed [69,70].

**SWAT**
SWAT (Soil and Water Assessment Tool) is an advanced, physically based, distributed, basin scale, hydrological model developed by USDA-ARS [71]. The hydrological model includes various processes such as surface runoff, peak flows, groundwater, evapotranspiration, etc. based on water balance equation and simulates the transport process of many substances including nutrients, sediment, heavy metals, etc [71]. Recent studies have shown that the SWAT model can predict sufficiently accurate runoff and NPS loads in a watershed. On the other hand, this model needs a great deal of data about under study watershed, which brings about uncertainties. In addition, there is an unsuitable
mechanism for run off calculation and description of the interaction of groundwater and surface water, which requires further research and improvement.

**HSPF**

HSPF (Hydrological Simulation Program—Fortran) is a distributed model for watershed scale model developed by the United States Environmental Protection Agency (EPA) [72]. HSPF uses three main modules to simulate hydrological processes in a watershed. These modules include PERLND, IMPLAND, and RCHRES. Each of these can be further divided into several compartments simulating different processes. Pervious Land-segments (PERLND) module contains 12 sections to simulate water quality and quantity parameters in pervious land segment. The Impervious Land-segments (IMPLND) module is divided into 6 sections estimating runoff and water quality parameters in impervious land. The RCHRES module is divided into eleven sections to simulates quantity and quality of the water in rivers of watershed. These modules are linked together to provide a comprehensive hydrological simulation for different segments of a watershed.

The HSPF is a model that employs empirical equations for hydrologic simulation and estimates hydrologic parameters and pollution loads. The HSPF has been considerably used in NPS pollution investigations in watersheds, providing a solid basis for the establishment and implementation of watershed management plans. As HSPF contains many empirical equations it includes high uncertainties for modeling and demands extensive data for the calibration of the model.

**Simulation Based Optimization Models**

A considerable part of NPS pollution modeling, including the spatial and temporal variations of hydrological parameters, can be represented using deterministic models. However, hydrologic data and equations have been subjected to uncertainties leading to bias and error in deterministic modeling. As a wide variety of parameters are used in the NPS pollution simulation, physically based models are sensitivities to weather, soil types and land use data and a slight change in input data lead to significant changes in the predicted results. Thus, the calibration process of physically based models is so critical in the simulation of NPS pollution. In addition, some simplifications and assumptions lead to an inevitable uncertainty of predicted results [73]. In addition, the placement and optimization of BMPs for controlling NPS pollution in watersheds is a complex and challenging problem. These problems almost contain a large number of variables which create computational efforts. This problem can be transformed into an optimization problem with spatial and temporal features [74,75]. In order to deal with the mentioned issues, several models based on an optimization-simulation approach have been developed to meet the demands. These models are composed of a deterministic core within a surrogate modeling frame for optimization. Simulation-based optimization models combine simulation modeling and optimization techniques known as simulation optimization (SO) models. Depending on the number of objective functions, optimization problems are categorized as single-objective or multi-objective.

**Single-Objective Optimization**

When there is just one objective function to optimize, the process is called single-objective optimization. In the single objective function all different objectives are lumped into one function representing the goal of the problem and the main target of single objective optimization is to get the best solution, corresponding to the either minimum or maximum value of the objective function [75].

Over the past years, the single-objective optimization has been widely used for NPS pollution modeling and selection of suitable BMPs [76-78]. The single objective optimization is an advantageous and applicable tool providing decision makers with a great understanding of the nature of the problem, but usually cannot provide a set of alternative solutions that trade different objectives against each other.

**Multi-Objective Optimization**

Optimization problems that need to address more than one objective are called multi objective optimization problems and may present several optimal solutions. In the multi-objective optimization problems, the optimal values are found out of many objective functions and there is no single optimal solution. The interaction among different objectives induces to a set of impaired solutions, largely known as the trade-off, or Pareto-optimal solutions [75]. Multi-objective optimization models as a tool for targeting BMP implementations to control NPS pollution has been received great attention lately [79-81]. In the SO simulation, SWAT, HSPF, and ANNAGNPS are widely used as hydrological models and the optimization algorithms are commonly NSGA-II or Ant Colony. Taking advantage of the multi-objective optimization model along with the simulation hydrological model greatly enhances the quality of the simulation process and provides a more realistic modeling. The main disadvantage of the multi-objective optimization simulation is that it is time intensive, and it needs a large computation time to run dynamically linked hydrological models.

**Future Prospects of NPS modeling**

NPS pollution is prevalent source of water pollution, and it has complex mechanisms and processes. The NPS pollution modeling is a complex procedure which requires a large data associated with climate, geological, land cover, and hydrological conditions. In this study, three main categories have been introduced for NPS pollution modeling by considering the spatial and temporal variations. This classification includes empirical models, physically based models, and simulation-based optimization models. The empirical models offer a simple equation to characterize the relationship between hydrological components and NPS pollution loads, particularly in the case of lacking monitoring data. Empirical models fall into three classes, including statistical methods, export coefficient model, and hydrograph separation methods. The empirical model commonly does not need a great deal of input data and contains a simple calculation technique; however, it lacks simulating of spatial and temporal variation of NPS pollution. Compared to empirical models, physical-based models contain good physical and chemical mechanisms and could map the spatiotemporal distribution of NPS pollution at larger scales, and they can represent a system's behavior by partial differential equations based on the physics of hydrological processes. Physically based models include two main groups of single-event simulation and continuous simulation models. The physically based model demands a great deal of data on climate, hydrology, land uses, etc. This model generally suffers from data scarcity issues, which reduce the accuracy and efficiency of this simulation process. Simulation-based optimization models combine optimization techniques into simulation analysis to address complicated problems. According to the number of objective functions, these models are divided into single and multi-objective functions. Simulation-based optimization models provide an efficient way to find BMP or a combination of BMPs based on various constraints; although, these models are typically computationally expensive.
Over the past decades, NPS pollution modeling has continuously evolved to provide an accurate simulation of NPS pollution; nevertheless, there are still some limitations which should be addressed. The accuracy of physically based modeling largely depends on the calibration of the input data. Due to the large number of parameters commonly used by physically based NPS models, both calibration and validation of these models commonly deal with some issues and challenges. Thus, conventional methods for either calibration or validation are generally cumbersome. Therefore, there is a need for developing an efficient framework for the calibration of NPS models. NPS pollution simulation normally only focuses on the processes of hydrology, erosion, and pollutant transportation; however, this process is quite complex including multiple processes of hydrology, environments, chemical, and ecological process. Consequently, the integration of different models in order to consider various parameters should be taken into account as a future trend in NPS pollution modeling. Multiscale modeling paves the way for integrating standalone models to provide a comprehensive simulation and efficiently pass information across temporal and spatial scales. Most of NPS pollution models just deal with surface water pollution and neglect the groundwater impacts; however, pollution of surface water can result in degradation of ground water quality and conversely contamination of groundwater can decrease surface water quality. Thus, considering the complicated interaction of surface water and groundwater is necessary for a comprehensive simulation of NPS pollution.

Developing a decision support system for selecting the optimum combination of BMPs to control NPS pollution within a watershed greatly helps decision-makers and policy-makers to increase the efficiency of the regulations and policies. Consequently, providing a framework to investigate the effect of different BMPs on controlling NPS pollution based on multi-model comparison could be crucial for future studies.

**Conclusions**

Determining the spatiotemporal variation in NPS pollution is a prerequisite for enhancing water quality and protecting the environment. Providing an accurate simulation of NPS pollution in a watershed is essential for water resource protection plans. However, considerable investigations have been carried out on NPS pollution modeling and defining and implanting BMPs to control NPS pollution, but the terminology and categorizing different approaches for NPS pollution simulation have not yet been well explained in the literature. The classification of NPS pollution modeling is vital to help researchers choose a proper technique for NPS pollution according to the limitations and purposes of modeling. A review of approaches to simulate NPS pollution is presented in this paper. The most dominant methods have been reviewed and categorized in the current study based on their properties, which could suggest a direction for researchers to choose the optimal approach to simulate NPS pollution in a watershed.

The approaches for modeling NPS pollution are classified into the following categories.

1. **Empirical models**
2. **Physically based models**
3. **Simulation based optimization models**

Empirical models offer simplified solutions for estimating NPS pollution loads in watersheds based on observed and monitored data. Empirical models are simple, easy to operate and have low-demand data; however, it lacks some accuracy to perform specifically on a watershed scale.

The physically based models provide comprehensive modeling by considering all hydrological processes in a watershed. As physically based models require a wide variety of input data, they could be operated when various data, including hydrology, geology, and so on, are available.

Simulation-based optimization models integrate a hydrological model with an optimization algorithm which results in high computational cost and effort in modeling. Simulation-based optimization models are commonly utilized for complex problems, which provide optimization for BMPs placement. Future NPS pollution modeling should attempt to decrease the limitations of the current simulation models to provide an accurate and realistic simulation. Thus, more attention would be inevitably paid to the calibration of NPS pollution models, multiscale modeling, considering the groundwater-surface water interaction, and developing a decision support tool in order to have a comprehensive NPS pollution modeling.

**References**


57. Song JH, Her Y, Guo T (2022) Quantifying the contribution of direct runoff and baseflow to nitrogen loading in the Western Lake Erie Basins. Scientific Reports 12: 9216.
77. MJE Skardi, A Afshar, SS Solis (2013) Simulation-

