Developmental Strategies for a Water Quality Assessment Model with Limited Datasets – A Case Study from River Bhavani, India

A Ravinashree¹, C Sivapragasam²*, M Vasudevan³*

¹Post-Graduate Student, Department of Civil Engineering, Kalasalingam Academy of Research and Education, Krishnankoil- 626126, Tamil Nadu, India.

²Senior Professor, Department of Civil Engineering, Kalasalingam Academy of Research and Education, Krishnankoil- 626126, Tamil Nadu, India.

³Assistant Professor, Department of Civil Engineering, Bannari Amman Institute of Technology, Sathyamangalam – 638401, Tamil Nadu, India.

*Corresponding authors Email: vasudevan@bitsathy.ac.in; sivapragasam@klu.ac.in

Abstract. Understanding the significance of rational attributes for a water quality model depends largely on the availability and reliability of virgin datasets pertaining to their spatial-temporal variations and interactions. Present study aims to identify certain crucial attributes for deriving a water quality model for a selected reach of River Bhavani in Tamil Nadu, India using the results of the QUAL2K model based on a limited dataset. The spatial variations in the concentrations of inorganic suspended solids (ISS) and carbonaceous biochemical oxygen demand (cBOD) were studied in detail corresponding to the possible variations in other model parameters (pertaining to the hydraulic, environmental and climatic conditions). The results indicate that the cBOD has a significant dependency on temperature, dissolved oxygen and pH, while the ISS has shown more dependency on alkalinity, conductivity and bottom algae. It is further inferred that variations in pH and dissolved oxygen can impose triggering effects when external pollution loads (point and nonpoint sources) are introduced while generating the corresponding profiles of ISS and cBOD. Based on the results, the conceptual strategies for developing a data-limited model is presented, which can be further developed to simulate water quality profiles for similar data-limited scenario.

Key Words: Biochemical oxygen demand; Data-limited Modelling; Inorganic suspended solid; QUAL2K; Water quality model

1. Introduction

The global scenario of accessibility for clean and safe drinking water is certainly challenging as driven by the exploitative captivation of surface and ground water resources. The key challenges presently existing and are likely to emerge in the near future pertaining to the water sector can be broadly understood under the classification of accessibility, availability, quality, ecosystem impacts, crisis preparedness and shared governance. As evidenced from many earlier studies, river pollution is one of the most serious challenges faced by the water sector of the present age. Rivers have become invariably
the immediate recipients and conduits of different types of wastewaters as imposed by the growing population adding with higher industrial production requirements. Unfortunately, most of the perennial rivers are on the verge of imminent bereavement and eventual transition to mere effluent drains as they are practically supplemented by the incessant sewage flows. In order to address some of the existing conflicts in water resources management activities, usage of surface water quality models can play a major role by simulating the spatial-temporal variations in the pollutant concentrations thereby depicting appropriate strategies for effective waste management and ecosystem restoration.

As the source and nature of pollutants entering into a water body vary from place to place, it is quite difficult to ascertain their exact distribution within a given water resource profile without having sufficient data collection and monitoring network. Various attempts were made earlier to devise reliable systems for collection, monitoring and assessment of real-time water quality data from the rivers for the purpose of developing water quality models [1-4]. However, many of such models were intrinsically limited in their applications in terms of site-specific characterization of model parameters, requirement of huge dataset and incomplete derivation platform for developing futuristic decisions for pollution abatement/restoration of resources. Hence selection of suitable water quality assessment tools is very crucial to understand the present situation, to simulate futuristic scenarios and to derive a meaningful decision-support tool.

Though many sophisticated simulation tools are being developed for surface water quality assessment by various research groups, the open source tool QUAL2K (and its previous version QUAL2E) has been quite extensively used in simulating the water quality profile for medium and large-scale river basins [1]. Cho and Ha [5] had performed parametric optimization for QUAL2K model using calibration with influence coefficient algorithm. They reported that the majority of the errors were found for dissolved oxygen and chlorophyll-A concentrations in a multi-reach river simulation. Zhang et al. [6] developed a water quality improvement model using QUAL2K for Taihu Lake Basin, China to select an optimal management program. Fan et al. [7] reported a combined modeling approach using QUAL2K and HEC-RAS to assess the impact of tidal effect on river water quality. They have concluded that such calibration essentially requires enormous corrections on reaction rate constants especially when the location is suffering from data-insufficient condition. Chowdhury et al. [8] attempted scenario-based modeling for Sungai Selangor Basin using QUAL2K for remediation management and planning. They have used the water quality index based on a minimum number of parameters (DO, BOD, COD, SS, NH3-N and pH) and validated QUAL2K results with the observed readings.

In general, QUAL2K is mainly employed for simulating various waste loading scenarios in polluted streams as part of basin-based environmental management plan [6-7]. In order to integrate various water quality models based on a conceptual interpretation, Bui et al. [9] successfully attempted to utilize the inputs of the SWAT model (in terms of pollutant concentrations and loading rates) as a feed for the QUAL2K model. Vasudevan et al. [10] studied the waste loading scenario in River Yamuna for a certain reach near Delhi using QUAL2K modeling with detailed analysis. However, one of the prominent research topics related to the water quality assessment is the DO-BOD interactions which were modeled extensively using QUAL2K [7, 11-12].

Since most of the water quality models are influenced by the parametric sensitivity and scalability of the datasets, the data-driven models using soft computing techniques are emerging as reasonable solutions to predict the trend of common water quality features. One specific advantage of these models is that the emphasis on obvious input parameters can be isolated and a rational relationship can be brought out which is purely acceptable based on the statistical evidence. There are many studies related to the prediction of water quality (in terms of dissolved oxygen, biochemical oxygen demand, electrical conductivity etc.) using soft computing methods such as genetic programming and artificial neural networks [13-15]. Such models are useful to predict the discharge variations and eutrophication (in terms of algal blooming) in water bodies [16-17]. Such models provide avenues to address the constraints such as restricted parameters with limited dataset. In some cases, a non-parametric approach is considered to be better that can eventually address the issues pertaining to limited data sets and constrained repeatability.
Though the QUAL2K model is quite user-friendly in developing the simulation scenario and capable of providing free-formatted results, a realistic representation of the field requires a huge number of input data owing to account for most of the possible complex interactions happening in a selected river section. This is particularly difficult to apply under situations of limited measured data in order to justify the prevailing assumptions and to verify the results. In an attempt to address this issue, the present study investigates the identification of crucially influencing parameters and their parametric sensitivity for two categorically representative pollutants (inorganic suspended solids – ISS and carbonaceous biochemical oxygen demand – cBOD). The simulation is particularly carried out for a selected stretch of River Bhavani located in Sathyamangalam, Tamil Nadu, India. Considering the limitations on acquiring all the required dataset for the QUAL2K model, the present study aims to provide a sensible characterization of various developmental strategies for simulating futuristic water quality under limited datasets.

2. Materials and Methods

2.1. Study Area
The study area consists of a selected stretch of River Bhavani at Sathyamangalam comprising of 5km longitudinal distance (from 11°30’18.30” N & 77°13’53.50” E to 11°28’50” N & 77°15’37.54” E) covering the key aspects of the localities as mentioned in Figure 1. The Bhavani River originates from the Western Ghats’ Nilgiri Hills, runs through Kerala’s Silent Valley National Park, and returns to Tamil Nadu. It is a perennial river having a length of 217 km, draining from a watershed area of 0.62 million hectares which is scattered throughout the states of Tamil Nadu (87%), Kerala (9%) and Karnataka (7%). The Bhavani sub-basin is primarily characterized as agriculture-based with supportive forest and plantation area in the upper regions. It is provided with an intensely irrigated agricultural system with remarkable groundwater depletion. In order to ascertain the intricate aspects of water pollution in River Bhavani, a few point and nonpoint sources were identified for sampling as shown in Figure 1. The selected points are representative sources of unconnected drainage outlets of municipal effluents which are directly discharged to the river without any treatment.
2.2. Data Collection

The closed-type layout of the entire river sub-basin has the advantage of effective utilization of surface water if sufficiently harvested. Therefore, the biggest disadvantage of this layout is the persistence of pollutants emerging from the neighboring industrial belt. The primary data collection includes the monitoring and analysis of basic water quality parameters such as pH, alkalinity, electrical conductivity, inorganic suspended solids (ISS), dissolved oxygen (DO) and biochemical oxygen demand (cBOD) from the selected points of reference within a reach of 5km as shown in the Figure 1. The spatial sampling method was consistently followed to study the variations in the distribution of pollutants over the selected stretch of the river. Moreover, the selected QUAL2K model requires steady state conditions which has been safely considered to be a typical representation of the extreme conditions of persisting pollution loading to the river. Hence, temporal variations in the pollutant concentrations are omitted in the present study although a seasonal variation can be of significance to study the futuristic impacts of the modeling strategies. Apart from the water quality data, other required information about the hydraulics and weather conditions of the study area were gathered from the literature about the study area with suitable assumptions which are explained in the upcoming sections.

2.3. Model Description

The selected QUAL2K model is a one-dimensional, steady state, non-uniform flow and transport model where the channel is assumed to be well-mixed both vertically and laterally. It is a very simple spreadsheet based model which can be downloaded (https://www.qual2k.com/) and installed on any medium-sized processor-system. In addition to the flow simulation, the heat budget and temperature profiles can also be modeled as a function of meteorological parameters on a diurnal time scale. Another
salient feature of this model is the inclusion of extensive modeling parameters pertaining to the kinetics of water quality which enables one to simulate the selected features on a daily/hourly time scale. As the focus of the present study is to identify the significant factors for simulating the water quality profile under limited dataset conditions, a reference condition is initially defined with all available data in order to serve as a means for the comparison of the results. The selected parameters and their values are described in Table 1.

Table 1: Sample database prepared for calibrating the QUAL2K model for the given scenario

| Parameter type      | Parameter description                              | Value | Unit       | Remarks/ Significance                                                                 |
|---------------------|----------------------------------------------------|-------|------------|---------------------------------------------------------------------------------------|
| Hydraulic           | Headwater flow rate                                | 1.5   | m³/s       | Normal range (1-10); selected for calibration                                        |
|                     | Flow measuring method                              | -     |            | To compare with a weir, rating curve or Manning's formula                             |
|                     | Temperature                                        | 27    | C          | Can vary diurnally                                                                    |
|                     | Electrical Conductivity                            | 800   | umhos      | Chance of many dissolved ions                                                        |
|                     | Inorganic Solids                                  | 30    | mg D/L     | Presence of sediment, mineral soils; highly variant                                  |
|                     | Dissolved Oxygen                                  | 5     | mg/L       | Vary between 4-8 depending on pollution                                              |
|                     | Carbonaceous Biochemical Oxygen Demand (fast degrading type) | 5     | mg O₂/L    | Readily degradable organic matter                                                    |
| Water quality       | Organic Nitrogen                                  | 1800  | ug N/L     | Mainly from nonpoint source pollution                                                |
|                     | NH₄-Nitrogen                                      | 500   | ug N/L     | Depends on undigested organic matter from point source                                |
|                     | NO₃-Nitrogen                                      | 300   | ug N/L     | Depends on degradation of organic matter from point source                           |
|                     | Organic Phosphorus                                | 10    | ug P/L     | Mainly from nonpoint source pollution                                                |
|                     | Pathogen                                          | 30    | cfu/100 mL | Based on total coliforms                                                             |
|                     | Alkalinity                                         | 150   | Mg CaCO₃/L | Total value                                                                           |
|                     | pH                                                | 8     | -          | Depends on water flow rate and mixing rate                                           |

2.4. Model Calibration
The QUAL2K model is fed with the necessary initial dataset to run for the hydraulic, geometric and climatic conditions for River Bhavani. Prior to further trails, the model is first calibrated by adjusting the hydraulic parameters based on the observed discharge values. For this, three conditions were considered as possible means of accessing realistic discharge values – Manning’s formula, Weir formula
and Rating curve. For each case, the model is run to fine-tune the corresponding flow model parameters (velocity coefficient, depth coefficient and roughness coefficient) while other parameters were kept unaltered during each run in order to check the satisfactory match with the reported results.

2.5. **Sensitivity Analysis**

The model is further analyzed for the sensitivity of various model parameters such as discretization limit, headwater quality parameters, lateral loading of pollutants, diurnal variations in temperature and dissolved oxygen as well as the variability in pollutant loading rates. A series of case studies were framed for simulation experiments by varying the limiting values under low, medium and high categories. This approach will help not only to identify the sensitivity towards individual parameters, but also to estimate the extent to which the results (i.e. downstream water quality profile) can vary. The details of the case study conditions are represented in Table 2.

Table 2: Description of the case studies used in the QUAL2K model

| S. No. | Parameter      | Value                      | Remarks                                                                 |
|-------|----------------|----------------------------|-------------------------------------------------------------------------|
| 1.    | River length   | 5.0 km                     | Representative of the study area with identified pollution sources      |
| 2.    | No. of reaches | 5                          | Selected after trails on the optimum number of flow elements            |
| 3.    | Headwater flow rate | 1.5 m³/s                | Representative of the average flow for the last 15 years              |
| 4.    | ISS            | 10, 100 and 500 (mg/L)     | Typical range limits selected based on extremely cases of low flow and flood flow |
| 5.    | DO             | 4, 7 and 10 (mg/L)         | Typical range limits selected based on literature values for varying oxygenation potential |
| 6.    | cBOD           | 5, 10 and 20 (mg/L)        | Typical range limits selected based on observed and reported values representing the background concentration |
| 7.    | ISS            | 10, 100 and 500 (mg/L)     |                                           |
| 8.    | DO             | 4, 7 and 10 (mg/L)         | Similar range selected based on the diurnal peak factor and flow variability due to the influence of change in effluents |
| 9.    | cBOD           | 5, 10 and 20 (mg/L)        |                                           |
| 10.   | pH             | 4, 7, 10 (-)               |                                           |

3. **Results and Discussion**

3.1. **Calibration Results**

The simulation results of the QUAL2K model for the selected reach of River Bhavani has been analyzed on a pragmatic conceptual platform in order to draw some plausible conclusions for developing a simplified data-limited model. Based on the calibration tests with flow parameters, the model is found to be quite stable and consistent with the hydraulic components of simulation. The simulated values of discharge rate using the three conditions were found to be consistent for the selected reach of the river (Figure 2).
Figure 2: Calibration results for flow simulation using QUAL2K

3.2. Simulated Contamination Profile for Inorganic Pollutants

The spatial variations of organic and inorganic pollutants in the river are simulated for the selected reach of the river. It is observed that the inorganic pollutants have reached high concentrations towards the outer end when considered as a constant point source at the selected entry point. However, the concentration values are found to be lower after assuming a settling rate of 0.5 m/day for the ISS (Figure 3). This is in accordance with the assumption that the pollutants carried by the influent can be quite turbid in nature and may get deposited to the river bed as they travel with the given low discharge value (1.5m$^3$/s). The same trend can be extended to analyze the expected profile of suspended solids within the vicinity of any point source of pollution. However, it is to be understood that the resuspension of sediments due to the turbulence can be a limiting criteria within the mixing reach of such loading.

Figure 3: Comparison of removal efficiency of inorganic suspended solids (ISS) within the reach for various input and loading conditions
3.3. Simulated Contamination Profile for Organic Pollutants

The extent of organic pollution in the river is studied in terms of the variations in DO and cBOD using various simulation case studies as mentioned in Table 2. Based on the close analysis, the concentration of DO is found to be directly dependent on the temperature profile in the river in consideration with its diurnal as well as seasonal variations. It is observed that the concentration profile of cBOD increased drastically within the selected regime despite the fact that DO is sufficiently supplied through reaeration throughout the simulation. When the initial cBOD is low (5mg/L), the depletion in DO was higher when compared to the increased input of cBOD (20mg/L). In addition, the increase in cBOD profile towards the outer end of the selected reach shows the significant impact of waste loading situations. This indicates the limited influence of dilution in absence of turbulence in bringing back the DO levels. However, it is observed that there is no significant increase in the ISS concentrations for the corresponding increase in cBOD indicating the predominant organic nature of the point sources emanating from the Sathyamangalam Township. The higher values of cBOD are of concern in order to provide sufficient degradation period as they are assumed to be fast degrading species.

![Figure 4: Comparison of simulated concentrations of dissolved oxygen (DO), inorganic suspended solids (ISS) and carbonaceous biochemical oxygen demand (cBOD) for various initial conditions of cBOD within the reach](image)

Based on the above mentioned results, the influence of organic pollutants on the overall environmental condition of the river can be understood in terms of the three simulated parameters, viz. DO, ISS and cBOD. However, in order to justify the selection of these parameters, the influence of other most prominent parameters was also analyzed. It is found that there is a drastic decrease in the electrical conductivity (EC) values under low DO conditions (4 mg/L) whereas the decrease in EC seems to be marginal for higher DO values (7 and 10 mg/L) (Figure 5). Since the river is observed to receive large amounts of organic-dominated effluent, the apparent change in the concentration of inorganic species seems to be limited mainly towards gravitational settling. Based on the results, the overall biodegradation potential of the river can be expected to be still viable under the selected ranges of the organic pollutants in order to abate the pollution loads within a certain reach limit. However, this is not the case when the model is to be used for an extreme contamination scenario.
3.4. Strategies for Modelling River Water Quality

The success of management operations in the water resources depends largely on the effective monitoring and protection of water quality based on the available data [18-19]. As observed with the above results, the organic pollutants can be understood to have an unsurpassable role on the overall environmental condition of the river despite the incremental changes in the distribution of inorganic species. Quite apart from performing a rigorous model developmental approach, it is important to adopt and articulate a comprehensive strategy for identifying the important factors and necessary steps to derive a data-limited model for futuristic prediction. As highlighted in the introduction as well as based on the experience from the study, an outline for a data-driven modeling framework is proposed here (Table 3) with a view to extend the modeling exercises towards successful prediction of water quality under practical constraints on data availability.

Table 3: Proposed strategies for selecting modeling constraints for water quality assessment under data-limited conditions

| Strategy type                      | Suggested steps                  | Suggested factors                        | Remarks from experience                                      |
|------------------------------------|-----------------------------------|------------------------------------------|--------------------------------------------------------------|
| Selection of inputs with limited records | Calibration and sensitivity analysis | Statistical error limits                  | Dependency on relevant literature is unavoidable. Model performance was satisfactory based on three-level sensitivity studies. |
| Correlation of grouped parameters  | Correlation and regression factors |                                          |                                                               |
| Adaptation of design of experiments procedure | Dimensionality based on principal component analysis or clustering |                                          |                                                               |
| Data split for training           | Optimization-based supervised methods | Statistical error limits                  | Not required in the selected model                           |

Figure 5: Comparison of simulated concentrations of dissolved oxygen (DO), inorganic suspended solids (ISS) and electrical conductivity (EC) for various initial conditions of DO within the reach.
Selection of model structure and architecture

| Selection between linear and non-linear input-output functions | Computational speed and efficiency | Over-dependency on large number of model coefficients can be minimized by adopting simplified rule-based selections |
| Selectivity of data-based models for hybrid approach | | |
| Trials on extended domains of same study area (replicability) | Statistical error limits | |
| Application to similar scenario (or futuristic) (predictability) | Selected model is quite extendable for higher reaches. Scenario-based simulations resulted in good predictability |
| Reduction of uncertainties (structural validity) | | |

4. Conclusions
The present study focuses on the water quality assessment using the QUAL2K model for a perennial stream under data limited conditions. The QUAL2K model is calibrated using the available hydraulic and environmental data for a limited stretch of River Bhavani in Tamil Nadu and further analyzed for the sensitivity of various model parameters. A number of case studies were developed to study the response of the model for the selected range of parameters. The model results suggest that ISS, DO and cBOD can frame the basic structure of the model if no other details of pollutants are available. The influence of DO on both organic and inorganic pollution was found to be critical in this study. The biodegradation potential of the river can be expected to be consistent under the selected ranges of the organic pollutants. However, this is not the case when the model is to be used for an extreme contamination scenario. The scope of developing a data-limited model can be evaluated based on the strategies for selection of input parameters, data splitting, selection of model structure and validation plan.

References
[1] Chapra SC 2008 Surface water-quality modeling (Waveland Press, Inc.) p 235-257
[2] Wu W, Dandy GC and Maier HR 2014 Environ Model Soft 54 108-127
[3] Moriasi DN, Gitau MW, Pai N and Daggupati P 2015 Trans ASABE 58 1763-85
[4] Ji ZG 2017 Hydrodynamics and water quality: modeling rivers, lakes, and estuaries (John Wiley & Sons) 473-503
[5] Cho JH and Ha SR 2010 Sci Total Environ 408 1985-91
[6] Zhang R, Qian X, Yuan X, Ye R, Xia B and Wang Y 2012 Int J Environ Res Public Health 9 4504-21
[7] Fan C, Ko CH and Wang WS 2009 J Environ Manage 90 1824-32
[8] Chaudhary S, Dhanya CT and Kumar A 2018 Hydrol Res 49 1042-55
[9] Bui HH, Ha NH, Nguyen TN, Nguyen AT, Pham TT, Kandasamy J and Nguyen TV 2019 Ecolhydrol Hydrobiol 19 210-23
[10] Mangottiri V, Nambi IM and Suresh Kumar G 2011 Int J Adv Engg Technol 2 336-344
[11] H Ismail A and A Abed G 2013 J Kerbala Univ 9 257-73
[12] Kalbargi PB, Shareefa RN and Deshannavar UB 2015 Int J Engg Manufa 5 15-25
[13] Vanitha S, Sivapragasam C and Namoothiri NV 2016 Int Conf Theor Comput Sci Discr Math (Springer, Cham) p 445-452
[14] Mahalakshmi R, Sivapragasam C and Vanitha S 2019 Inform Commu Technol Intelligent Sys (Springer, Singapore) p 401-408
[15] Sivapragasam C, Jegatheesan V, Arun VM and Vanitha S 2010 Int J Eng Sci Technol 2 3128-36
[16] Sivapragasam C, Muttil N, Muthukumar S and Arun VM 2010 Marine Pollu Bullet 60 1849-55
[17] Sivapragasam C and Muttil N 2005 Water Resour Manage 19 505-20
[18] Singh RP, Mangottiri V and Pandiyan B 2020 IOP Conference Series: Materials Science and Engineering Ed Vasudevan M, Soundara B, Karthiga N and Preetha V (IOP Publishing) p 012085
[19] Sivakkumar SN, Mangottiri V, Narayanan N and Saragur SK 2020 IOP Conference Series: Materials Science and Engineering Ed Vasudevan M, Soundara B, Karthiga N and Preetha V (IOP Publishing) p 012103