OPTIMAL WATER QUALITY MANAGEMENT IN

SURFACE WATER SYSTEMS AND

ENERGY RECOVERY IN WATER DISTRIBUTION NETWORKS

A Thesis
Presented to
The Academic Faculty

by

Ilker Tonguc Telci

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Civil and Environmental Engineering

Georgia Institute of Technology
December 2012

OPTIMAL WATER QUALITY MANAGEMENT IN

SURFACE WATER SYSTEMS AND

ENERGY RECOVERY IN WATER DISTRIBUTION NETWORKS

Approved by:

Dr. Mustafa M. Aral, Advisor
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Marc Stieglitz
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Jiabao Guan
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Turgay Uzer
School of Physics
Georgia Institute of Technology

Dr. Seong-Hee Kim
School of Industrial and Systems Engineering
Georgia Institute of Technology

Date Approved: October 5 2012

To my family

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor Prof. Mustafa M. Aral for the opportunity, guidance and inspiration. I will always be grateful for his patience, understanding and the fact that his door is always open to our long discussions. What I have learned from him is way beyond the limits of this thesis.

I would like to thank Dr. Jiabao Guan for his valuable comments and guidance as a member of Multimedia Environmental Simulations Laboratory (MESL) group and a member of the Ph.D. committee. I place on record, my sincere gratitude to the Ph.D. committee members Dr. Seong–Hee Kim, Dr. Marc Stieglitz and Dr. Turgay Uzer for their comments and enlightening suggestions.

I gratefully acknowledge Hydro Research Foundation (HRF) for giving me the opportunity to be a Hydro Fellow and for their financial support. I would also like to thank to HRF for our round tables where I learned and enjoyed a lot.

I would like to express my sincere appreciation to the MESL crew Dr. Wonyong Jang, Dr. Elcin Kentel, Dr. Jinjun Wang, Dr. Sinem Gokgoz Kilic, Dr. Kijin Nam, Dr. Scott Rogers, Dr. Recep Kaya Goktas, Mr. Andi Zhang, Mr. Biao Chang, and Mr. William

Morgan for their support, friendship and valuable comments and suggestions during MESL seminars.

I owe my deepest gratitude to my mother Gulay and my father Fikret Telci, my grandmother Turkan Arin, my sister Cansu Telci-Kahramanogullari and her spouse Kerem Kahramanogullari for their unconditional love and support felt from the other side of the world. Our scientific discussions with my father and mother have important contributions in my way of thinking and helped me to become the engineer I am today.

Finally, I cannot find words to express my gratitude to my wife Ayten Memmedova-Telci and my son Deniz Anar Telci for their love, support and patience in this long journey. We, as a family, overcome many difficulties which strengthen our love. This thesis would have remained a dream and my life would be so empty without them.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Water quality monitoring and search for environment friendly energy sources is becoming two of the most popular engineering research topics as we better understand the limits of our planet. In this thesis, first an optimal design methodology for water quality monitoring networks in river systems is developed. Next, a data interpretation approach is proposed to identify pollution source locations utilizing the water quality measurements supplied by the monitoring network. As the third topic, the thesis introduces an optimal design technique for energy recovery systems in water distribution networks.

In the first part of this thesis, an optimization algorithm is developed for the water quality monitoring system. In this process, the best monitoring locations are determined by utilizing the outcomes of a simulation model. The results of the simulation model is an essential component of this approach since they incorporate the unsteady and stochastic nature of hydrodynamics and the contaminant fate and transport processes in rivers into the optimization model. In this approach, the ideal monitoring locations are determined through a multi-objective optimization technique. One of the objectives of the monitoring system is specified as the early detection of the contaminants and the other as the reliability of the monitoring network. The methodology developed was first applied to a simple hypothetical river system to demonstrate the importance of the unsteady hydrological properties of the watershed on the optimal locations of the monitoring

stations. Then, it is tested on a realistic river system. The results show that the design technique developed can be effectively used for the optimal design of monitoring networks in river systems.

In the second part of the study, a methodology for rapid identification of contaminant source locations is introduced. Since this is an ill posed problem which has non-unique solutions, a classification routine which correlates candidate spill locations with the measurements at the water quality monitoring stations is developed. For this purpose, the breakthrough curve of a contaminant measured at monitoring site is parameterized using its statistical moments. Then, a large number of spill scenarios are simulated for the training of the monitoring system. After the training process, the method is ready for sequential elimination of the candidate locations which leads to the identification of spill location for a breakthrough curve observed at the monitoring station. The model developed is applied to real river system and the results show that this technique can be a reliable starting point for the contaminant source investigation projects.

The third part of the thesis is devoted to renewable energy production from water distribution systems. The main idea behind this study is to harvest as much available excess energy as possible by utilizing micro turbines. The energy production at these turbines is constrained by the minimum pressure limit set by the management. Moreover, the unsteady nature of the flow in the network results in variations in the available excess energy. These aspects of the water distribution systems necessitate operation schedules

for the micro turbines. In this study, a simulation-optimization method is developed which maximizes the energy recovered at the micro turbine(s). This simulation-optimization model is based on Genetic Algorithms (GA). A smart seeding of the GA is introduced to lower the computational burden. The algorithm tests several energy recovery system configurations which has different turbine locations and turbine types. Then the best configuration which has the highest energy production is selected. The methodology is first applied to a real pump driven network. Then, this network is converted into a hypothetical gravity driven system and the optimization model is tested on this new system. The results show that the energy recovery systems in water distribution networks can provide significant economic and environmental benefits and the methodology introduced is not only an optimal design tool but also an effective means of assessing the renewable energy potential in water distribution systems.

CHAPTER 1

INTRODUCTION

Two of the most important environmental challenges in the 21st century are to protect fresh water resources and to utilize renewable energy sources to lower greenhouse gas emissions. This study contributes to the solution of the first challenge by considering river systems which constitute a major component of fresh water supplies. Since the most important means of protecting river water quality are water quality monitoring networks, this study focuses on design of monitoring networks and interpretation of data supplied by the monitoring network to identify pollution source locations. The solution of the second challenge depends on discovering new renewable energy sources and finding new means to harvest this energy. Recent literature demonstrates that water distribution systems host important amounts of clean excess energy and this study presents a methodology to recover this energy. This chapter discusses the motivations of this study and introduces the concepts utilized to contribute to the solutions of these challenges.

1.1 Real-Time Water Quality Monitoring Networks for River Systems

Rivers are one of the most important natural resources, which have been the primary fresh water supply for humans throughout history. Also rivers are habitation places for a large diversity of species. This vital water resource has crucial influence on the

economies of the countries as well. United States Environmental Protection Agency (USEPA) states that some of the designated uses of these water bodies in the United States are public water supply, aquatic life harvesting, agriculture, recreation, fish, shellfish and wildlife protection and propagation (USEPA 2009). Therefore, the quality of river waters has important impacts on public health, wildlife and economy. To restore and maintain the chemical, physical and biological integrity of this natural resource, in 1972, The U.S. Congress passed The Clean Water Act (CWA) which requires states and other jurisdictions to assess the health of their waters and the extent to which their waters support water quality standards (USEPA 2009). Similar actions are taken on the other side of the Atlantic as well; in 2000, European Parliament passed the European Union Water Framework Directive (WFD) (Directive 2000/60/EC). In all these attempts for the protection of rivers, water quality monitoring is the key step to understand the current condition and to promote informed decisions on the use and management of these water resources.

The primary purpose of a water quality monitoring system in a river network is to provide a system that would generate sufficient and timely information to enable the managers to make informed management decisions regarding the quality of life of the populations that are utilizing this resource. The secondary purpose of water quality monitoring may include issues such as monitoring the quality of the environment for the essential needs of the habitat, identification of pollution sources thus potential polluters, immediate initiation of clean-up operations after an accidental or deliberate spill, concerns on potential terrorism events and the use of the data collected to identify

stringent rules and regulations to avert the adverse effects that may be caused by the consequential environmental degradation, that is the Total Maximum Daily Load management (TMDL). A TMDL is the total amount of pollutant that can enter a water body without causing it to violate a water quality standard for that pollutant. When the primary purposes identified above are considered as the criteria of the real-time monitoring network design, the design objectives can be easily identified. For the purposes cited above, the design objectives are the early detection time and the reliability of the monitoring system designed. These two criteria are essential to protect humans from adverse effects of exposure to harmful contaminants. Focusing on these two objectives, this study proposes a methodology that is based on the transient behavior of random contamination event or events in a river network. The proposed model is based on the hydrodynamics and the contaminant fate and transport characteristics of the river system under study. The information gathered from this analysis is used in an optimization model to identify the best monitoring locations in the river network in real time which would satisfy the two objectives identified above.

1.2 Identification of Contaminant Source Locations in River Systems Using Water Quality Monitoring Networks

Once the real-time monitoring stations are optimally located in the river system, they provide continuous information about the quality of the river water. This information can be used for long term management of the river system such as in the decision of the

TMDLs and in the inspection of these rules. In addition to these current applications, improving real time sensor technologies assign new duties to monitoring systems such as rapid identification of contaminant source locations. This information on the source location can be used for remediation purposes and for environmental forensic studies. In addition, source identification provides important information to be used in the health risk analysis of potential water pollution events in river systems. Identification of contaminant source location is essential for reducing the risk of exposure by preventing recurring pollution events as well as providing timely response to deliberate or accidental pollution incidents in rivers.

When a pollutant enters to the stream, its concentration can be recorded at a monitoring station as a breakthrough curve. The question is whether this data can be used to reveal important information about the location of the contaminant spill. This source location identification problem is an ill posed problem because of the irreversible nature of the contaminant transformation and transport processes. This problem has been studied in different media such as groundwater, water distribution systems and indoor air and many different solution techniques are proposed such as simulation-optimization (Guan, Aral et al. 2006), solution of transport problem backwards in time (Bagtzoglou and Baun 2005) and pattern recognition (Datta and Peralta 1986) and artificial neural networks (Singh, Datta et al. 2004). One of the objectives of this study is to develop a methodology which solves this problem in complex river systems with minimum computational effort. Thus, a technique which can be classified into pattern recognition group is utilized to gather

information on the location of the pollution spill from its breakthrough curve observed at the monitoring stations.

1.3 Water Distribution Systems as a Source of Renewable Energy

Within the past few decades, it is understood that global warming has become one of the most important environmental problems affecting the future of human civilization. The increase in average temperature on the surface of the earth cause drastic problems in many environmental aspects such as global sea level rise, severe floods and droughts all around the globe. The future results of global warming may change the surface area of the land causing large migrations, reduce the amount of available clean water, cause problems in agriculture, and eventually bring severe economical and social crisis all around the world. Studies show that global warming is a result of increasing concentrations of greenhouse gasses in the atmosphere (IPCC 2007). The primary greenhouse gasses in the atmosphere are water vapor (H_2O), carbon dioxide (CO_2), nitrous oxide (N_2O), methane (CH_4) and ozone (O_3). The sources of these gasses can be both natural and anthropogenic. However, rapidly increasing human activities such as burning of fossil fuels to produce energy dramatically increased the concentrations of greenhouse gasses within the past few centuries. As a result, reducing or stopping global warming became one of the most important challenges in the history of mankind. One of the solutions to this problem is utilizing renewable energy sources which do not emit greenhouse gasses.

Renewable energy is widely understood as energy recovered from the natural processes. Main renewable energy sources are sunlight, wind, rain, tides and geothermal heat. However, recently new renewable energy sources which can be considered as anthropogenic are being added to this list. This energy is produced inevitably during human activities but dissipated if it is not recovered. The best example to this type of renewable energy can be found in water distribution networks. Water distribution networks are designed to satisfy the consumer demands at the outlet nodes. To achieve this goal, adequate pressures need to be maintained throughout the network. However, high pressures cause damage to the pipeline and increase the amount of leakage causing serious economic loss. Therefore pressures in the water distribution networks have upper and lower constraints. While the pressures lower than a minimum cause unsatisfied demands and pressure requirements for emergency needs such as fire protection, the excess pressures cause pipe damage and leakage problems. As the complexity of a water distribution network increase, maintaining target pressures becomes more difficult causing excess pressures in the network. The conventional solution to this problem is to install pressure reducing valves which adjust the local head loss to lower the downstream pressure to a set value. However, this process causes dissipation of significant amount of energy that can be recovered and used by the community without emission of additional greenhouse gases. This energy recovery is possible by utilizing micro hydroelectric power plants as an alternative means of pressure reduction to pressure reducing valves. A micro hydroelectric power plant has a capacity less than 100 kW (Monition 1984). When the goal is set as energy recovery from excess pressures in a water distribution system, then the next step is to design this energy recovery system. The design of this system is

an optimization problem for the decision of the number, capacities and the locations of hydroelectric power plants installed in the water distribution network. The objective of this problem can be set as the maximization of the energy produced by the energy recovery system and the constraints are the minimum and maximum pressures along with the requirement of satisfied consumer demands. In this study a methodology is developed for the optimal design of energy recovery systems for water distribution networks

1.4 Scope of the Study

This thesis is organized in six main chapters. Chapter 2 of the study provides background and literature on the three main problems of the study in three sections. In the first section, real-time water quality monitoring systems, their design criteria and corresponding design methodologies in river systems are reviewed. The second section reviews the techniques for the identification of contaminant source locations in river systems. In the third section of Chapter 2, past studies on the solution of excess pressure problem and methods proposed for energy recovery in water distribution networks are provided. Chapter 3 presents the solution methodology and applications of optimal design of real-time water quality monitoring networks. The applications are arranged in two main parts. In the first part, the methodology is tested in a hypothetical river system taken from the literature for comparison purposes and for the demonstration of the effects of hydrodynamics and hydraulic characteristics of the river system on the optimal solution. The second part of Chapter 3 presents the results of the proposed method for a natural

river system which is chosen as the Altamaha River in the State of Georgia, USA. Chapter 4 describes the methodology for the identification of contaminant source location by using the data provided by a real-time water quality monitoring system and demonstrates the application of the proposed technique in Altamaha River system. In Chapter 5, the optimization approach for the design of energy recovery systems for water distribution networks is described and the methodology is applied on the Dover Township water distribution system in the State of New Jersey, USA. Finally, Chapter 6 presents a summary of findings of the thesis and concludes the study.

CHAPTER 2

BACKGROUND AND LITERATURE REVIEW

2.1 Sensor Technologies for Water Quality Monitoring

The primary constituent of a river monitoring system is the method of chemical or biological analysis which provides quantitative information about the quality of the water analyzed. Increasing human population and rapidly growing industry cause an increase in the amount and diversity of the pollutants disposed in water bodies which necessitates fast and cost effective analytical methods in environmental monitoring programs. For screening basic water quality parameters such as temperature, conductivity and pH, specific sensors for continuous data acquisition are available (Harmon, Ambrose et al. 2007). However, detection of emerging contaminants such as complex organic compounds requires more sophisticated methods. Traditional analytical methods such as chromatographic methods require conventional sample collection, transport to the laboratory and instrumental analysis. Therefore, although these methods provide precise information on the concentration of the contaminants, they are not efficient in the case of an accidental or deliberate spill of contaminants which necessitates a rapid response. Current trend in water quality monitoring is to utilize emerging technologies such as immunoassays and biosensors (Farre, Brix et al. 2005; Allan, Mills et al. 2006). An immunoassay is a biochemical test to measure the concentration of a substance observing the reaction of an antibody to its antigen. Immunoassays can be used for real-time

monitoring of several compounds such as polychlorinated biphenyls (PCBs), pentachlorophenol, atrazine, polynuclear aromatic hydrocarbons (PAHs) (Sadik and Van Emon 1996) and benzene, toluene, xylene (BTX) (Gerlach, White et al. 1997). A biosensor, also called an immunosensor, is a device for the detection of an analyte that combines a sensitive biological component with a physicochemical transducer which senses the interaction of the analyte with biological element and produces a corresponding signal providing real-time and cost effective results (Rodriguez-Mozaz, de Alda et al. 2005). Tschmelak, J., G. Proll, et al. (2004) showed the capability of an immunosensor for the detection of several contaminants such as bisphenol A, estrone, atrazine, isoproturon, endocrine disrupting chemicals, pesticides and several other organic compounds.

To summarize, sensor technologies for environmental screening is a rapidly growing field and improvements in this field result in precise and cost effective methods for real-time water quality monitoring. Once these technologies are available the next question is at which locations in a river system we should utilize these sensors in order to maximize our monitoring performance.

2.2 Water Quality Monitoring Network Design

The first step in monitoring network design is to identify the design considerations such as why to monitor and expected information from the system. In other words, the

objectives of the monitoring network and the information expected for each objective should be clearly defined prior to actual technical design of the monitoring system (Steele 1987; Harmancioglu and Alpaslan 1994). Kwiatkowski (1991) identified these objectives for large scale monitoring networks as detection of violations of quality standards, monitoring trends in the environment, providing information to develop and validate predictive models and monitoring ecosystem health. However, as emerging sensor technologies develop and new environmental threats become possible, additional objectives such as detection of accidental or deliberate spills become necessary (Allan, Mills et al. 2006). Once these conceptual design objectives are determined, other aspects of the problem such as number and locations of monitoring sites, sampling frequency, parameter selection and data transfer can be studied (Strobl and Robillard 2008). An extensive review about the studies on the design of these aspects of monitoring systems can be found in (Dixon and Chiswell 1996). Among these activities, the determination of data collection points constitutes the most important step since the success or failure of the other steps relies on the performance achieved in this first step.

The first attempt to determine optimum sampling locations came from Sharp (1971). He used a topological method which describes the river network in terms of stream order numbers and divides the watershed into sub-catchments. The division points were assigned as the optimum sampling locations. One year later, Beckers, Chamberlain et al. (1972) developed a methodology to determine priorities of segments along a river in terms of probability of violations to water quality standards. Harmancioglu and Alpaslan (1992) utilized statistical entropy analysis to assess the spatial frequency of monitoring

stations using monthly observed values of 40 water quality variables. Applicability of genetic algorithms (Cieniawski, Eheart et al. 1995) and integer programming (Hudak, Loaiciga et al. 1995) were also demonstrated. Kriging method incorporated with a one dimensional steady-state river water quality model was proposed as another option to select the optimal locations of monitoring stations (Lo, Kuo et al. 1996). Dixon, Smyth et al. (1999) improved Sharp's (1971) method to determine optimum sampling locations by using geographical information system (GIS), graph theory and simulated annealing algorithm. As new environmental simulation algorithms develop and computing capabilities improve, simulation-optimization methods became feasible. Ning and Chang (2004) presented a strategy to expand water quality monitoring stations in a river system by integrating simulation and fuzzy optimization approach. They also proposed a method for optimal relocation of monitoring stations to meet the long term monitoring objectives inspired from Kwiatkowski (1991) utilizing the same simulation algorithm with compromise programming (Ning and Chang 2005). Concurrently, Icaga (2005) applied a procedure using genetic algorithms to select optimum station combination in an existing monitoring network without simulation but utilizing real observations. Studies on the relocation and expansion of existing monitoring stations continued with Park, Choi et al (2006). They used observations from an existing monitoring network as their database and applied genetic algorithms in association with GIS to obtain the optimum design. Ouyang, Yu et al. (2008) proposed an optimization method for sampling locations considering only geometry of a river system. Their approach was to minimize the required cost to locate the pollutant source using genetic algorithms. Karamouz, Nokhandan et al. (2009) utilized entropy theory to determine redundant monitoring sites

in an existing monitoring network. They used a river water quality simulation model to generate a water quality database for their statistical entropy analysis.

As a summary, several sophisticated methods have been developed to find optimal locations for river water quality monitoring stations since 1970's. When data requirements of these methods are compared, some of these methods consider only the geometry of the river network, some utilize existing water quality observations and others are based on water quality models. Most of these water quality models are designated to obtain steady-state solutions simply because of the complexity of the watersheds and the rivers that are considered in those studies. However, as the computational capabilities of the hydrodynamic and transport models improve, it becomes possible to simulate unsteady conditions in natural and man-made systems.

Today, there are examples of monitoring system designs which take into account the transient nature of environmental systems. Chadalavada and Datta (2008) developed a methodology for optimal design of a dynamic groundwater monitoring network using a simulation model which considers the transient flow and transport process in the aquifer. Guan, Aral et al. (2006) determined optimal monitoring locations for a water distribution system emphasizing dynamic behavior of fate and transport processes. Similar problem was studied by Rogers (2009) by introducing nodal importance concept to reduce the computational efforts. Nam (2008) have studied optimization of both stationary and mobile water quality monitoring systems in lakes. Recently, importance of transient

simulation models in river networks was demonstrated by Gevaert, Verdonck et al. (2009). (Rogers 2009) In this study, a methodology which considers the unsteady behavior of hydrodynamics, fate and transport of contaminants in river networks is proposed. With this consideration and advances in sensor technologies summarized in previous section, the performance criteria of new generation water quality monitoring systems are determined as early detection and reliability of the system.

2.3 Identification of Contaminant Source Location

The second law of thermodynamics which explains the principle of increase in entropy (or disorder) in natural systems and processes dictates the irreversibility of the contaminant transformation and transport processes. Due to this irreversible nature, contaminant source identification problems are always ill posed problems and they require special approaches such as response matrix method, simulation optimization and pattern recognition techniques.

Gorelick, Evans et al. (1983) formulated this problem in groundwater applications as an optimization model using concentration response matrix technique. In the response matrix approach, they developed a two dimensional response matrix which describes the simulated concentrations at the measurement sites due to unit leaks at the candidate locations. The concentration at any measurement location is a liner combination of these unit leaks. Hence, their objective was to minimize the difference between the superposed

simulated concentrations and the measurements also called residuals and they utilized linear programming and multiple linear regression as the optimization methods. Aral and Guan (1996) utilized genetic algorithms to search the groundwater pollution sources in a 2-D homogeneous aquifer using a similar concentration response matrix. They developed an improved genetic algorithm to overcome the limitations of genetic algorithms in constrained optimization problems. Sun, Painter et al. (2006) embedded the response matrix technique into an iterative optimization model to identify contaminant source location and release history in groundwater environment. In their model they used a robust least square estimator which takes into account the uncertainty and they used branch and bound method as the optimization tool. Aral, Guan et al. (2001) formulated the contaminant source identification problem in groundwater as a non linear optimization model and they used a progressive genetic algorithm which combines an iterative process of simulation and genetic algorithms producing a computationally efficient method. Singh and Datta (2006) utilized a linked simulation-optimization approach using genetic algorithms as the optimization tool. In this simulation-optimization technique, the objective was to minimize the difference between observed concentrations and simulated concentrations at the measurement locations. Guan, Aral et al. (2006) proposed a similar simulation-optimization method to identify the contaminant source locations in water distribution systems. They used reduced gradient method as the optimization tool in this application.

Besides optimization methods, researchers have used different approaches such as the solution of transport problem backwards in time. Bagtzoglou and Baun (2005) used

Marching Jury Backward Beam equation in conjunction with discrete Fourier transform technique to improve computational efficiency. However, they used an optimization algorithm due to ill posed nature of the inverse problem.

Bayesian inference approach was another method used to overcome the difficulties of this ill conditioned inverse problem (Keats, Yee et al. 2007). In this method, Bayesian statistics was utilized to calculate the conditional probability density function of the source location and strength given a set of concentration measurements and a Markov chain Monte Carlo approach was used to generate these posterior distributions of the source parameters. Backward probabilistic models which estimate location and travel time probabilities using backward modeling were also used to determine prior location of pollutants in groundwater (Neupauer and Wilson 1999; Neupauer and Wilson 2001; Neupauer and Wilson 2004; Neupauer and Lin 2006). Similar probabilistic inverse models were applied to determine source locations in indoor air environments (Liu and Zhai 2008; Liu and Zhai 2009).

Other interesting approaches to solve source identification problems were pattern recognition and artificial neural network applications. Pattern recognition is the classification of an observation as one of the finite number of patterns based on its features. Datta and Peralta (1986) introduced an expert pattern recognition method to solve this problem in groundwater using an optimal sequential pattern classification algorithm. They used the expert system to decide to continue or terminate the pattern

recognition algorithm which utilizes dynamic programming. They associated this termination criterion with the risk of false classification. Later, Singh, Datta et al. (2004) applied artificial neural network approach to solve the problem in groundwater. They trained the artificial neural network with a set of data produced by a groundwater flow and contaminant transport model. Then, they used this trained network to identify the pollution sources for a given concentration observation data at measurement locations. Afterwards, Singh and Datta (2007) further improved this artificial neural network approach for partially missing concentration observation data.

As a summary, the contaminant source location identification problem has received significant attention from the scientific community providing a diversity of solution approaches. Another interesting observation is that this problem is not addressed significantly in river systems. One example of an inverse solution in river systems came from Boano, Revelli et al. (2005). They assumed that source locations are known and applied geostatistical methods to estimate release histories of pollutants in a river reach.

In this study, a methodology for identification of a contaminant source location in complex river systems is proposed. This approach can be classified in the set of pattern recognition algorithms. The methodology proposed is based on an adaptive sequential feature selection algorithm developed by Jiang (2008) for neural signal decoding

2.4 Energy Recovery from Water Distribution Systems

The environmental impacts of utilization of fossil fuels as a major source of energy necessitate the research on renewable energy resources. The main renewable energy resources are hydropower, biomass, solar, wind, geothermal and ocean energy. Among these resources, hydropower has an estimated total theoretical potential of 150 Exajoules annually (Johansson, McCormick et al. 2004). The hydroelectricity generation is a mature technology which is not expected to be further improved. However, small-scale hydropower still has potential of growth (Johansson, McCormick et al. 2004).

The term “small-scale” is most widely used for the hydroelectric power plants which has a capacity less that 10 MW (Paish 2002). Smaller hydropower plants can be classified into *micro* when the power is less than 100 kW and *mini* when it is between 100 and 5000 kW (Monition 1984). Some of the advantages of micro hydropower plants cited by Paish (2002) are predictable available energy, no fuel and limited maintenance requirements, long lasting technology and no environmental impact.

The traditional application of small-scale hydropower is to construct a small hydraulic structure such as a weir in a river to produce a head difference for the turbine (Yuksek and Kaygusuz 2006) and this traditional application still has significant potential all around the world (Nunes and Genta 1996; Dudhani, Sinha et al. 2006; Pokharel, Chhetri et al. 2008; Kanase-Patil, Saini et al. 2010). Recent studies have shown that mini and

micro hydropower plants are feasible and economically viable for applications in municipal water distribution networks (Ramos, Covas et al. 2005; Bieri, Boillat et al. 2010; Soffia, Miotto et al. 2010). The idea behind energy production from water distribution systems comes from the fact that energy dissipation devices such as pressure relief valves have already been used to reduce excess pressures which occur due to operational necessities. Therefore, most of the studies aim to recover this dissipated energy by a mini or micro hydropower plant.

The first attempt to integrate hydropower plants in water distribution systems came from Afshar, Benjemaa et al. (1990). They proposed a methodology to find the optimum number of hydropower plants and their locations, the optimum capacities of the turbines to be used and optimum pipe diameters at each section of the main transmission line in the water supply system. They utilized dynamic programming to find the optimum combination of these design parameters by maximizing the annual net benefit. The constraints of their optimization model were upper and lower bounds for the diameters considered, heads for the hydropower plants and pressures allowed. Another constraint was satisfaction of the demands.

Ramos, Covas et al. (2005), showed experimentally that pressure reducing valves and micro-turbines have similar behaviors for steady state flows and different activities were observed under transient conditions. They observed that although in some cases micro hydropower plants perform better in pressure regulation in other cases a mixed solution

of micro turbines and pressure reducing valves is recommended. They also proposed an optimization method to minimize pressure and the number of pressure reducing valves in the water distribution network. They utilized genetic algorithms to solve this optimization problem.

In another recent study, first, pressure reducing valves were introduced in the water supply system at optimal locations and then these valves were substituted by micro hydropower plants (Giugni, Fontana et al. 2009). This study emphasized the leakage problem due to excess pressures in the distribution systems. Their simulations showed that similar leakage reduction was achieved after the pressure reducing valves are replaced by micro hydro power plants. Genetic algorithms were used as the optimization tool.

Another interesting study which demonstrates the future potential of renewable energy applications in water supply systems came from Vieira and Ramos (2009). In their study, they introduced an optimization model for the operation of a hybrid water supply system equipped with a pump, hydraulic turbine and a wind turbine. The optimization model to achieve maximum energy efficiency described the pumping operation schedule considering the economic benefits coming from the profit of wind energy to supply water pumping. The optimization algorithm developed was based on linear programming formulation.

As a summary, micro hydropower plants have significant prospective applications in water distribution systems. These clean energy recovery systems can be used efficiently for pressure regulation in water distribution systems as a replacement of pressure reducing valves already in use. In this study, a methodology to determine optimal number, location and capacity of micro hydropower plants to achieve maximum energy recovery is proposed. The main constraints of the problem are defined as satisfaction of the consumer demands and maximum and minimum permissible pressures.

CHAPTER 3

OPTIMAL DESIGN OF WATER QUALITY MONITORING NETWORKS

3.1 Introduction

The components of a river monitoring network design study would include the selection of the water quality variables, identification of the locations of sampling stations and determination of the sampling frequencies. These are primary design considerations which may require a variety of objectives, constraints and solution methods. This study focuses on the optimal river water quality monitoring network design aspect of the overall monitoring program and proposes a novel methodology for the analysis of this problem. In the proposed model, the locations of sampling sites are determined such that the contaminant detection time is minimized for the river network while achieving maximum reliability for the monitoring system performance. For comparison purposes, the proposed method is tested on a simple network that has been studied in the literature. The comparative analysis of the design generated in this study and the outcome presented in the literature is discussed for various contamination scenarios. The results indicate that steady state solutions or solutions based on the geometry of the river network may not provide a reliable solution for the network design problem and a dynamic analysis may be necessary to solve this important problem. Altamaha river system in the State of Georgia, USA is chosen to demonstrate an application of the proposed methodology on a

natural river system. The results show that the proposed model can be effectively used for the optimal design of real-time monitoring networks in river systems.

3.2 Methodology

The proposed methodology is based on two important steps: (i) determination of the dynamic behavior of a contamination event in a river network. In this step the data that will be utilized in the second step is generated and stored; and (ii) determination of the monitoring locations based on an optimization model. These two steps are described in more detail below.

3.2.1 Hydrodynamic and Contaminant Fate and Transport Simulation

For the hydrodynamics and contaminant fate and transport analysis, the EPA Storm Water Management Model (SWMM) is used. As stated in the user manual of the model (Rossman 2007), SWMM is a dynamic rainfall-runoff model which can be used for single event or long-term (continuous) simulation of runoff quantity and quality from primarily urban areas with possible extension to watershed analysis. Among many capabilities of SWMM, one of the most important objectives in this study is the handling of river networks of unlimited size. Also, SWMM can be used to simulate a river network system with a wide variety of standard closed and open conduit shapes as well as natural

channels. In the SWMM model, user defined external flows and water quality inputs can be used for a variety of watershed conditions or simple channel networks can be analyzed. Although the model is designed for urban areas, SWMM can be used for the preliminary analysis of large watersheds as discussed in this study.

SWMM uses the Manning's equation to calculate the depth of flow in all conduits and to calculate the flow within a conduit link SWMM employs the conservation of mass and momentum equations (i.e. the Saint Venant equations). Contaminant transport within a conduit link is performed with the assumption that each sub-reach of the conduit behaves as a continuously stirred tank reactor (CSTR). Since a river reach can be divided into several sub-reaches to improve numerical accuracy in an application, the outcome is again suitable for the purpose of this study. The model also performs contaminant fate and transport analysis based on the results of the hydrodynamic analysis for various contamination conditions. The software has a flexible graphics user interface and can be imbedded in optimization models of efficient handling of the output.

3.2.2 Optimization Model

In this study, the river monitoring system is designed based on two performance measures: (i) minimizing the average detection time of the contamination events; and, (ii) maximizing the reliability of the monitoring system. Accordingly, the performance of the monitoring system will increase as the time between the start of a contamination event

and its detection time decreases and the number of potential scenarios detected increases. Any contamination event that leaves the river system without being detected will decrease the reliability of the monitoring system. These cases are penalized in the evaluation of the objective function to emphasize this negative impact.

The river network is assumed to contain *N* monitoring stations. These are the candidate monitoring stations. If the required (desired) number of monitoring stations is specified as *M* where $(M \le N)$, we may use the vector *X* to represent the solution vector of the monitoring stations denoted as, $X = [x_1, x_2, ..., x_i, ..., x_M]^T$, where x_i is the index of a monitoring station with in the set of candidate monitoring locations. For a given solution *X*, the detection time of a monitoring station *n* in a monitoring station set for the contamination scenario *s*, $d_s^n(X)$, is defined as the time elapsed between a contamination event and occurrence of a concentration exceeding a predefined threshold at a monitoring station. The detection time of the monitoring system for the scenario *s*, $t_s(X)$ is defined as the shortest time among the detection times of the monitoring stations for a contamination event and a penalty value of $(t_{sim} - t_{s,inj})$ is assigned as the detection time for all non-detected scenarios as described in Equation (3.1), where t_{sim} is the total simulation time and $t_{s,inj}$ is the starting time of the contamination event for scenario *s*.

$$t_s(X)=\begin{cases} min\left\{d_s^1(X),d_s^2(X),...,d_s^i(X),...,d_s^M(X)\right\} \text{ if scenario } s \text{ is detected,} \\ \left(t_{sim}-t_{s,inj}\right) \text{ if scenario } s \text{ is not detected.} \end{cases} \quad (3.1)$$

Based on this definition, the average detection time of the monitoring network *X*, $\bar{t}(X)$, which represents one of the objectives, can be calculated by taking the average of $t_s(X)$ over all possible scenarios:

$$\bar{t}(X)=\frac{1}{S}\sum_{s=1}^{S}t_s(X) \quad (3.2)$$

where *S* is the number of all scenarios considered in the analysis.

The reliability of the design *X*, $R(X)$, is defined as the ratio of detected contamination scenarios to the total scenarios tested. Therefore, the reliability of the monitoring system, which is the second objective function of the optimization problem and can be calculated as:

$$R(X)=\frac{1}{S}\sum_{s=1}^{S}\left(\frac{1}{N_e}\sum_{k=1}^{N_e}\delta_k(X)\right)_s \quad (3.3)$$

where $\delta_k(X)$ is an indicator variable taking the values 0 or 1 for non-detection and detection events respectively and N_e is the number of contamination events in a scenario. The reliability equation given above is selected to consider the occurrence of multiple contamination events within the simulation period and the possibility of detecting only

some of these events. If S is very large, it may become impossible to simulate all possible scenarios and consider them in the optimization problem. Then $\bar{t}(X)$ and $R(X)$ can be defined as the expected detection time and reliability of a randomly selected subset of scenarios in a statistical sense. These expected values can be estimated as:

$$\begin{aligned} \bar{t}(X) &= E\left[t_s(X)\right] \\ R(X) &= E\left[R_s(X)\right] \end{aligned} \tag{3.4}$$

where $E[.]$ indicates the expected value computed as:

$$E[Y] = \sum_{y:p(y)>0} yp(y) \tag{3.5}$$

where Y is a discrete random variable which takes on the values y associated with a probability function $p(y)$.

In this study the purpose of the monitoring system is to detect the contamination events as quickly as possible with the smallest failure rate. To achieve this goal, average detection time should be as small as possible and the reliability of the monitoring system should be as high as possible. Therefore, the design of a monitoring network can be formulated as an optimization problem that can be mathematically stated as:

$$
\begin{aligned}
f_1 &= \underset{X}{Minimize}\left\{\bar{t}\left(X\right)\right\} \\
f_1 &= \underset{X}{Maximize}\left\{R\left(X\right)\right\} \\
s.t.\ M &= M_o
\end{aligned}
\tag{3.6}
$$

where M_o is the predefined total number of monitoring sites.

The multi-objective optimization problem described above can be solved using a multitude of methods. In this study we will use the genetic algorithm (GA) for large systems and the enumeration search approach when the example problems discussed can be handled with this approach within a reasonable computation time. For more complicated applications where multiple optimal solutions are possible we will use the Pareto optimal analysis. The computational procedures for the GA and Pareto analysis are standard procedures that can be found in the literature and will not be repeated here (Holland 1975; Goldberg 1989; Guan and Aral 1999). Applications of this methodology to the design monitoring systems in water distribution systems has been tested with considerable success (Guan, Aral et al. 2006; Guan, Aral et al. 2006; Nam and Aral 2007; Nam and Aral 2007; Rogers, Guan et al. 2007).

3.3 Applications

In this section the application of the proposed methodology for two different river systems are demonstrated in separate sections. In the first section the proposed methodology is tested for several cases on a simple hypothetical river network. The primary purpose of this work is to: (i) develop the basic methodology used in this analysis; (ii) compare the optimization solution with other studies from the literature to verify the outcome; and , (iii) understand the effects of the hydraulic parameters and watershed characteristics on the optimal solution. In the second section, the proposed methodology is applied to the Altamaha River system which is a hydrologically and hydrodynamically much more complex river system and the effect of the more complex scenarios on the optimal outcome is further investigated and discussed.

3.3.1 Hypothetical River Application

In order to test and compare the proposed methodology, a hypothetical river network is chosen from the literature (Ouyang, Yu et al. 2008). As shown in Figure 3.1, this river network is composed of 11 river reaches (capital letters) and 12 Junctions (numbers). The hydraulic characteristics of the river network are given in the following sections separately since some of the parameters are changed for illustration purposes.

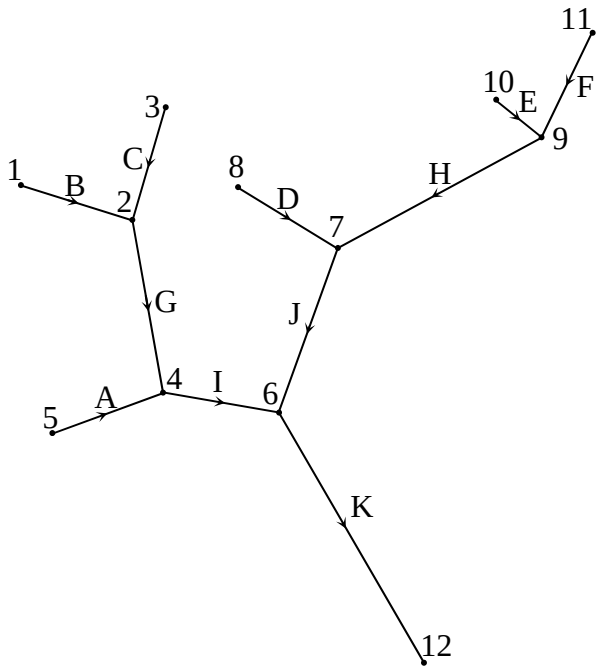


Figure 3.1 Hypothetical River Network.

3.3.1.1 Evaluation and Comparison of the Methodology

The methodology proposed is applied to several cases based on the contamination pattern in the river network. In Case 1, a single spill may occur randomly at any junction of the river system. In Case 2 and Case 3 the possibility of the occurrence of two random contamination spills at any two distinct junctions of the river network is considered. In Case 2, the occurrence of the two spills is simultaneous, whereas in Case 3 there is a 15 minute time lag between the two spill scenarios. In all cases, the best monitoring location is searched for a river monitoring system that is composed of three sensors. That is the optimization model given in Equation (3.6) is solved with constraint $M_o=3$. All three cases are investigated first on a simple hypothetical river system studied by Ouyang, Yu

et al. (2008). This river network is composed of 12 junctions and 11 reaches as shown in Figure 3.1. Therefore, the optimization algorithm will select the best placement scheme among $P = C_3^{12} = 220$ possible placement scenarios.

To be able to compare the results of this study with the results given in Ouyang, Yu et al. (2008), it is assumed that all river reaches have the same rectangular cross-section with a top width of 10 ft and lengths of channels selected have the same proportions as reported in the above mentioned study. Also, channel bottom slopes and Manning's roughness coefficients of all channels are the same throughout the river system. Furthermore, flow rates at the most upstream reaches are set to be equal to 10 ft^3/s. The resulting hydraulic characteristics of the river network are given in Table 3.1. In order to decrease the effect of CSTR assumption of SWMM, each river reach is divided into 100 ft long sub-channels.

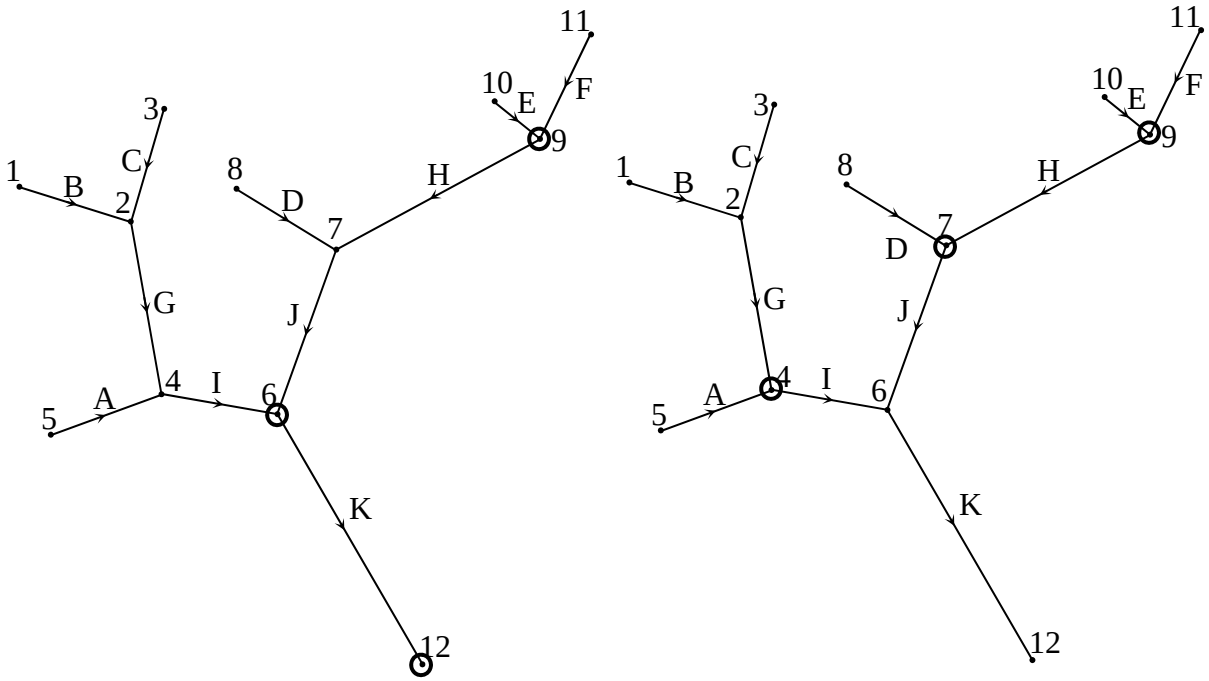
The reason for these assumptions is to render the river network system to behave similar to the system analyzed by Ouyang, Yu et al. (2008). The analysis used in that study was based on a geometric analysis of the system as opposed to the dynamic analysis used here.

Table 3.1 Hydraulic characteristics of the River Network

Reach	Length (ft)	Flow rate (ft^3/s)	Depth (ft)	Width (ft)	Channel slope	Manning's n
A	2000	10	1.31	10	0.0001	0.02
B	2000	10	1.31	10	0.0001	0.02
C	2000	10	1.31	10	0.0001	0.02
D	2000	10	1.31	10	0.0001	0.02
E	1000	10	1.31	10	0.0001	0.02
F	2000	10	1.31	10	0.0001	0.02
G	3000	20	2.08	10	0.0001	0.02
H	4000	20	2.08	10	0.0001	0.02
I	2000	30	2.75	10	0.0001	0.02
J	3000	30	2.75	10	0.0001	0.02
K	5000	60	4.53	10	0.0001	0.02

In Case1, contamination events are assumed to be single instantaneous spills occurring at the junctions of the river system randomly. Therefore, the number of possible spill scenarios to be simulated is *S=12*. After simulation, best solution satisfying Equation

(3.6) is searched using GA. Since the search space and number of scenarios are not large in this study, the global optimum solution is also determined by exhaustive search method comparing all possible outcomes and the GA solution is verified using this method. Optimization algorithm locates one of the monitoring stations to be at the outlet (Junction 12) in order to maintain 100% reliability (Figure 3.2(a)). However, for reliability values less than 100%, other junctions can be used such that detection time is minimized. Optimum monitoring locations for 100% and 83% reliabilities are given in Figure 3.2(a) and (b) respectively.



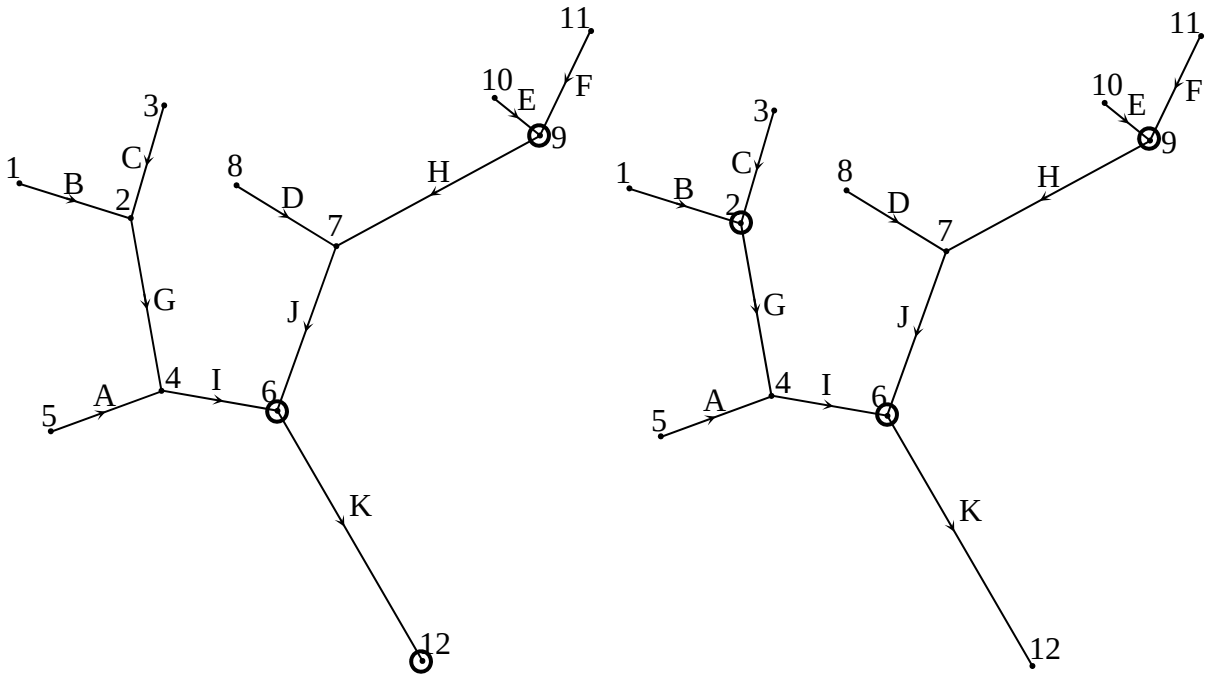
(a) 100% Reliability Solution (b) 83% Reliability Solution

Figure 3.2 Optimal monitoring stations for Case 1.

In Case 2, two instantaneous contamination spills are assumed to occur at distinct junctions spontaneously. Therefore, total number of possible spill scenarios to be

simulated is $S = C_2^{12} = 66$. For this case, both GA and exhaustive search method reached the same best placement scheme. As previously mentioned, the only difference between the contamination scenarios in Cases 2 and 3 is the 15 minutes time lag between the spills in Case 3. This difference doubles the total number of possible spill scenarios to be simulated in Case3 as $S = 2C_2^{12} = 132$. Again, exactly the same results were obtained from GA and exhaustive search method.

For Cases 2 and 3, although the optimized solutions do not have a monitoring station at the outlet (Junction 12), detection likelihood values are 100%. This result comes from the fact that even if one of the spills occurs at Junction 12 and is not detected, the monitoring station at junction 6 guarantees the capture of the other spill event. Therefore, in order to minimize detection time optimization algorithm moves the monitoring station at the outlet in Case1 (Figure 3.3(a)) to another junction for Cases 2 and 3 (Figure 3.3(b)).



(a) Solution for Case 1 (b) Solution for Cases 2 and 3

Figure 3.3 Optimal Monitoring Stations for Case 1, 2, 3.

Best monitoring locations for all three cases and corresponding detection times and their reliability values are given in Table 3.2. Since any non-detected scenario is represented by the simulation time, the detection time increases significantly as reliability decreases. Also, number of spills in the assumed scenarios has an important effect on detection time. Table 3.2 indicates that, for 100% reliability, single spill scenarios result in longer minimum detection time than two-spill scenarios. Furthermore, the time lag between two spills results in an increase in minimum detection time.

Finally, to demonstrate the effect of dynamic fate and transport processes, results of Case 1 are compared with the results of the study in which optimization was performed based only on geometrical characteristics of the river system (Ouyang, Yu et al. 2008). As seen from Figure 3.4 optimization algorithm of our study moves the monitoring station at Junction 7 to Junction 9. This increase in the importance of Junction 9 comes from the fact that Reach E is the shortest channel in the river system thus any spill at Junction 10 is captured at Junction 9 in a much shorter detection time. Also, the minimum detection time for the configuration proposed by Ouyang, Yu et al. (2008) is calculated as 72.5 minutes which is larger than that of this study (Table 3.2).

Table 3.2 Summary of results for hypothetical river system.

Case Number	Optimization Method	Optimum Sensor Locations	Minimum Detection Time (min)	Reliability (%)
1	Genetic Algorithm	6 – 9 – 12	63.75	100
		4 – 7 – 9	1000	83
	Enumeration Search	6 – 9 – 12	63.75	100
		4 – 7 – 9	1000	83
2	Genetic Algorithm	2 – 6 – 9	32.05	92
		6 – 9 – 12	37.73	100
	Enumeration Search	2 – 6 – 9	32.05	92
		6 – 9 – 12	37.73	100
3	Genetic Algorithm	2 – 6 – 9	38.09	92
		6 – 9 – 12	44.55	100
	Enumeration Search	2 – 6 – 9	38.09	92
		6 – 9 – 12	44.55	100

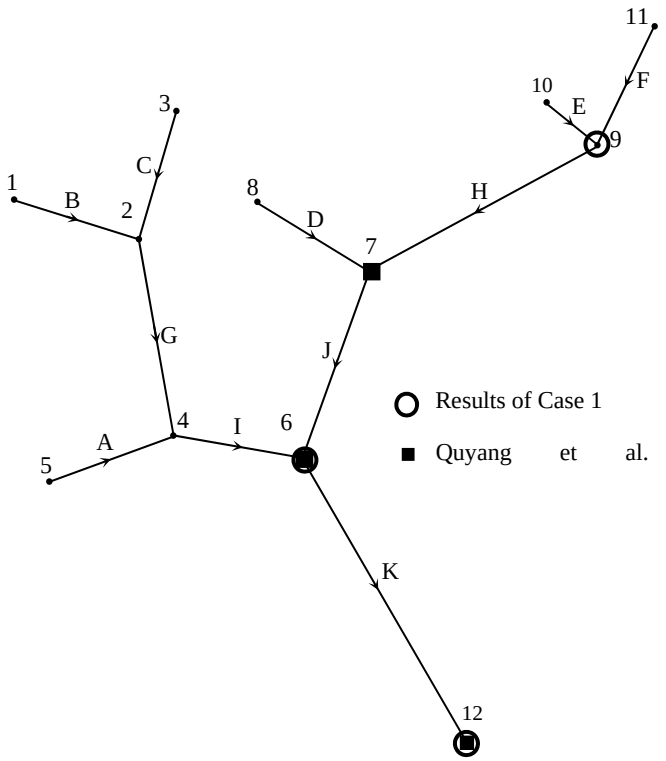


Figure 3.4 Comparison of the results of this study and Quyang et al. (2008)

3.3.1.2 Emphasis on Hydraulic Parameters

Up to this point, all hydraulic parameters such as Channel bottom slopes, Manning's roughness coefficients and width of the channels are assumed to be uniform trough out the river network (original configuration) as shown in Table 3.1. In order to demonstrate effect of these parameters on the optimal solution, only the Manning's roughness coefficient of reach G is increased from 0.02 to 0.05 and all other parameters are kept the same in the modified configuration. This increase in the roughness of channel G will decrease the velocity and increase the depth from 2.08 ft/s to 3.96 ft/s slowing down the transport of contaminant in that channel. Consequently, the optimization algorithm which tries to minimize the average detection time will prefer a location upstream of channel G.

Figure 3.5 shows that for 100% reliability solution in Case 1, the monitoring station downstream of channel G at Junction 6 in the original configuration is moved to an upstream position at Junction 2. This action further affects the position of the monitoring station at Junction 9 in the original configuration moving it to Junction 7 in modified configuration. The optimum solutions for all three cases are summarized in Table 3.3 which demonstrates that slow transport at channel G increases the importance of Junction 2 and Junction 9 is still assigned a monitoring station for two spill scenarios. Table 3.3 also indicates that average detection times for all cases are increased as expected due to slow motion of contaminant through channel G.

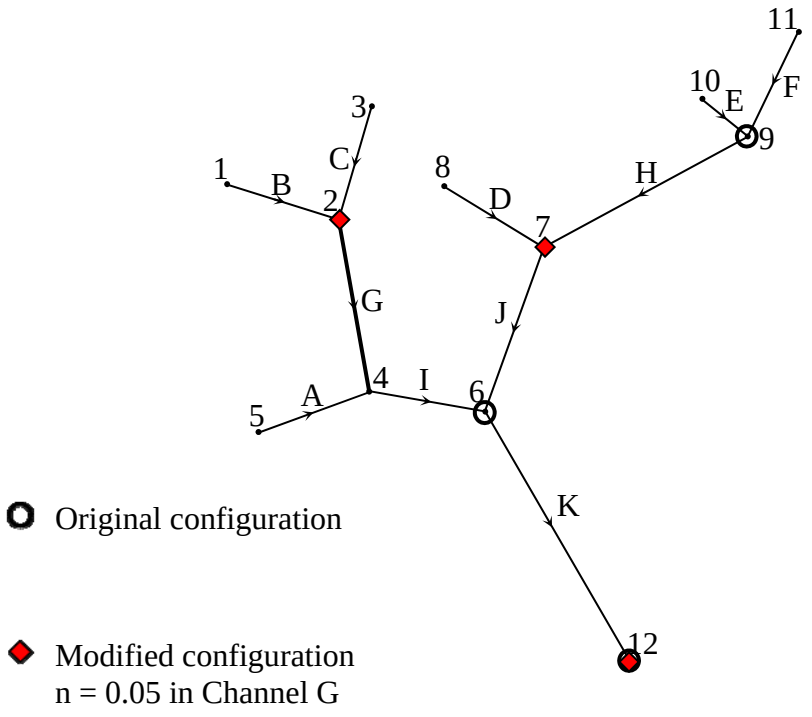


Figure 3.5 Comparison of optimum solutions for the river network with original and modified hydraulic parameters for Case 1 with 100% reliability.

Table 3.3 Effect of hydraulic parameters on optimum solutions.

Case	Hydraulic Configuration	Optimum Sensor Locations	Average Detection Time (min)	Reliability (%)
1	Original	6 – 9 – 12	63.75	100
	Modified	2 – 7 – 12	70.00	100
2	Original	6 – 9 – 12	37.73	100
	Modified	2 – 9 – 12	39.32	100
3	Original	6 – 9 – 12	44.55	100
	Modified	2 – 9 – 12	46.14	100

3.3.1.3 Emphasis on Watershed Characteristics

The aim of this part of the work is to illustrate the effect of watershed characteristics such as areas of subcatchment discharging the rain water into the river network and the intensity of this rain on the optimum placement of the monitoring stations. From another point of view, influence of the unsteady behavior of the river flow as a consequence of rain events on the optimum solution is investigated in this section of the study. To reach this purpose, hypothetical subcatchments are assigned for each junction of the river system. Area of each subcatchment is assumed to be directly proportional to the total length of the upstream river reaches discharging into the corresponding junction with the

ratio 1/2000 ac/ft. For the most upstream subcatchments, it is assumed that there is an imaginary river reach which has the same length as the real river reach just downstream of the corresponding junction. Possible rainfall intensities are assumed to be 1, 3 and 5 in/hour with a rain fall duration of 1 hour. All other hydraulic parameters are kept the same as listed in Table 3.1. One contamination event is assumed to occur with one rain event simultaneously at independent junctions as a possible scenario. Thus, there exist 12 possible rain locations with 3 different rain intensities and 12 possible contamination spill points resulting in a total number of possible scenarios to be simulated as $S = 3 \cdot 12 \cdot 12 = 432$.

The assumption that area of each subcatchment is proportional to the upstream river reaches results in subcatchments of different areas connected to junctions of the river network. For example, using the proportionality constant of 1/2000 ac/ft, while the area of the subcatchment assigned to Junction 1 in Figure 3.6 can be calculated as 1 ac, the area connected to Junction 7 is 3 ac. These differences in the areas of the subcatchments result in different rainfall runoff behaviors such as total volume of rain water, duration of rainfall runoff and peak discharge. Consequently, each rain event produces a different unsteady flow rate scheme in the river system. This leads to a faster transport of contaminant at some reaches than others and affects the optimum solution in a more complex but similar manner as explained in previous section. Figure 3.6 shows the best sensor placement for 100% reliability with an average detection time of 55.65 min. This decrease in the average detection time from 63.74 min for Case1 in Table 3.2 to 55.65 min for the scenarios with the rain events comes from the fact that rainfall runoff makes

the contaminant transport faster causing the contaminant to reach the sensor locations earlier.

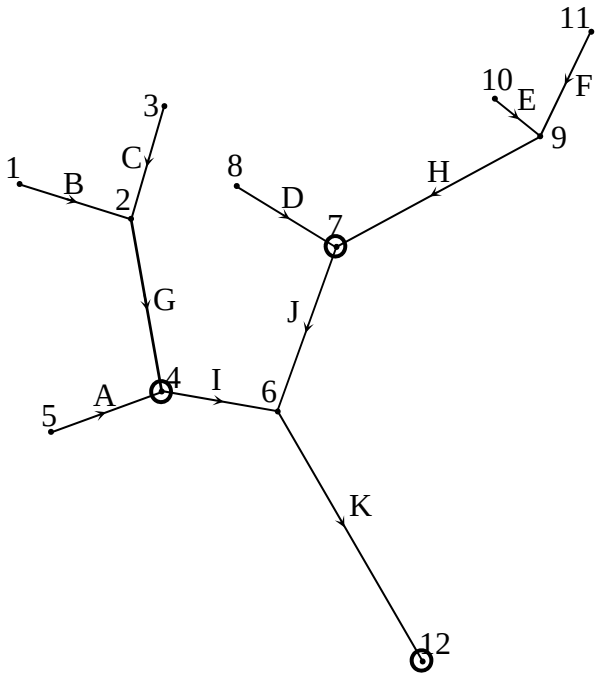


Figure 3.6 Optimum locations of monitoring stations with 100% reliability and average detection time of 55.65 min considering rainfall events.

3.3.2 Altamaha River Application

Altamaha river basin is the largest watershed in the State of Georgia, USA draining the water collected from about 25% of the state area to the Atlantic Ocean in a south-east direction. It is also the third largest basin discharging to the Atlantic Ocean to the east of Mississippi river. As shown in Figure 3.7, Altamaha River system is formed by the confluence of Ocmulgee, Oconee and Ohoopee rivers. This river network is composed of 60 river reaches 32 of which are upstream channels discharging into 28 internal streams.

All these channels form 30 confluences and 32 upstream points resulting in a total of 62 river junctions.

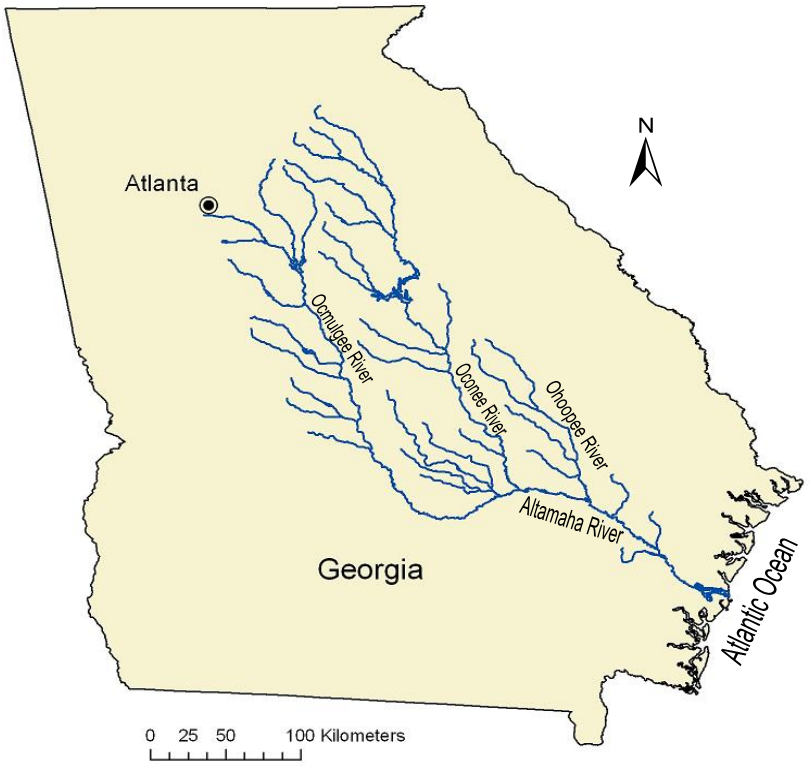


Figure 3.7 Altamaha River network in the State of Georgia USA.

Complexity of the Altamaha River system does not only come from the number of river reaches and the number of confluences they form. In addition to these, irregular characteristics of the natural topography, natural river cross-sections and complexity of lateral inflow rates are some of the difficulties that have to be overcome during the analysis of the hydrodynamics and the contaminant fate and transport characteristics of the system.

During the simulation step, one of the most important parameters that need to be estimated is the slope of the river beds. This information is obtained from the digital elevation data using ArcGIS. The digital elevation data maps are supplied by U.S. Geological Survey (USGS) in the National Elevation Dataset (USGS 2008). In this study, the digital elevation dataset which has the highest precision available is used with a resolution of 1 arc second (30 m). From this dataset, elevations of all junctions are determined and the slope of each river reach is assumed to be constant and calculated from the elevation difference between corresponding upstream and downstream junctions.

An important hydraulic dataset for the hydrodynamic analysis is the cross-sections of the river channels in the river system. In this study cross-section of a river channel is approximated as a trapezoid as shown in Figure 3.8. There are limited cross-section data available for the river system. This dataset can be obtained from bridge crossings and existing USGS monitoring site locations. These data are more extensive for the Lower Altamaha region which is downstream of the confluence of Ocmulgee and Oconee rivers (Figure 3.7). The data for cross-sections are obtained from the data available for this region (Gunduz and Aral 2003; Gunduz and Aral 2003) and meticulously extended to upstream regions of the river system considering the site conditions and data that exist for these regions. For this purpose, an upstreamness factor which indicates relative position of a river reach in the system is defined as described in Figure 3.9. Then, the geometric parameters for the river channels are determined assuming linear relation with respect to the upstreamness factor as shown in Table 3.4.

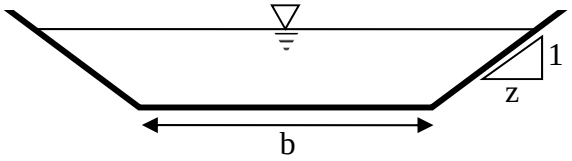


Figure 3.8 Trapezoidal approximation for the river cross-sections

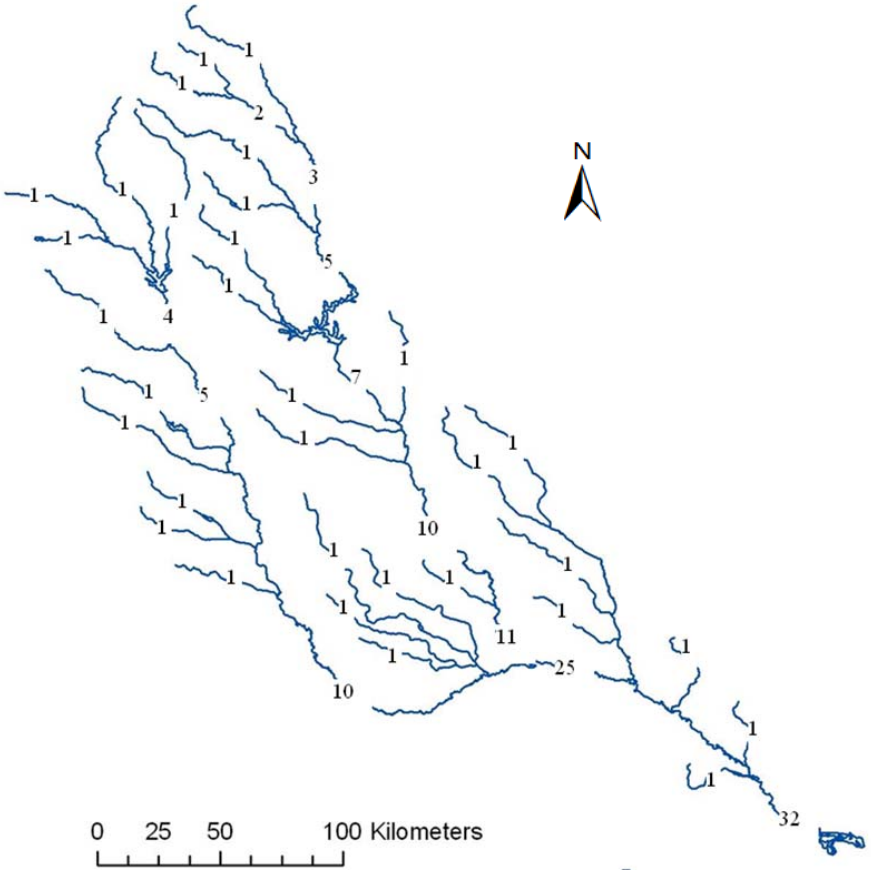


Figure 3.9 Upstreamness factors along Altamaha river system

Table 3.4 Geometric parameters used for the analysis.

Upstreamness Factor	b (ft)	z
1	73.00	10.00
2	80.00	11.25
3	87.00	11.88
4	94.00	12.50
5	101.00	13.13
6	108.00	13.75
7	115.00	14.38
8	122.00	15.00
9	129.00	15.63
10	136.00	16.25
11	143.00	16.88
14	164.00	18.75
25	241.00	25.63
29	269.00	28.13
30	276.00	28.75
31	283.00	29.38
32	288.00	30.00

In order to assign steady state inflow rates to each river reach, the data provided by USGS gauging stations are used. The steady state hydraulic system is calibrated for the

flow pattern in the river based on the data obtained from annual average flow rates measured in 1990 at 14 gauging stations by USGS. These gauging stations are distributed throughout the river network. The inflow discharges are assigned at the most upstream river reaches such that the flow rates at the USGS gauging stations obtained from the hydrodynamic simulation became consistent with the field measurements. At this calibration step, the rainfall runoff from the watershed into the river system was not considered, since the regionally heterogeneous rainfall data was not available for the period of analysis. For this reason it is impossible to obtain the exact flow rates that are measured by USGS at the gauging stations. The weighted average percent error between the measured and predicted flow rates was 4.5% for the fourteen gauging stations considered in the calibration analysis performed which is considered to be acceptable for the purposes of this study. The measured discharge at the most downstream gauging station in 1990 is reported as 14,430 cfs. This flow rate is also maintained by adjusting the discharges of the most upstream river reaches which would yield a minimum deviation in other gauging stations. A Manning's roughness coefficient of 0.02 is assumed for all the river system which is a typical value for the Altamaha region.

3.3.2.1 Selection of Potential Monitoring Stations and the Design of the Contamination Scenarios

For the locations of candidate monitoring stations, confluences and most upstream points are considered first. However, since this selection limits the set of potential monitoring

sites with the locations of confluences, additional nodes are added evenly along each river reach. As a result, number of candidate monitoring locations is extended to 100 as shown in Figure 3.10.

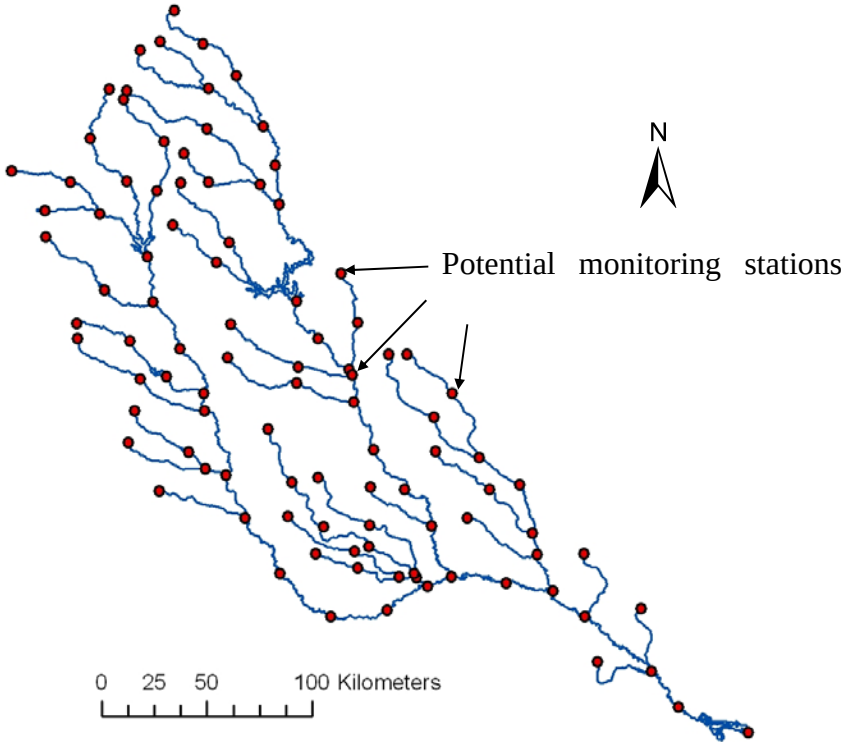


Figure 3.10 Selected locations for spill locations and candidate monitoring stations.

In this study, single instantaneous spills are considered for the contamination scenarios. Possible locations of these spills coincide with potential monitoring sites described in Figure 3.10. Although similar spill events are taken into account, 3 different scenario sets are generated in terms of timing of the spills and hydrodynamics of the river system. The peak contaminant concentration profile obtained for one of these instantaneous spills

is shown as a typical outcome profile along the river path in Figure 3.11. In this case the spill originates at junction one and along the way passes through the junctions 19, 80 and 100. In Figure 3.11 the horizontal axis is the junction index thus representing the distance from the spill. In this scenario the maximum concentration at the spill location is 100 g/l which is kept constant for all scenarios. Although conservative chemicals are considered in this study the dilution effects can be clearly observed from Figure 3.11 as the contaminant migrates in the downstream direction. Based on this observation, it is important to note that detection threshold selected in the analysis plays an important role. As can be seen from Figure 3.11, if the detection threshold is selected as 0.01 mg/L this spill will not be detected by any monitoring station to the downstream of junction 80. In addition to the analysis of the transient hydrodynamics within the watershed as described below, the algorithm proposed in this study takes the dilution effects observed into consideration in selecting the monitoring station locations.

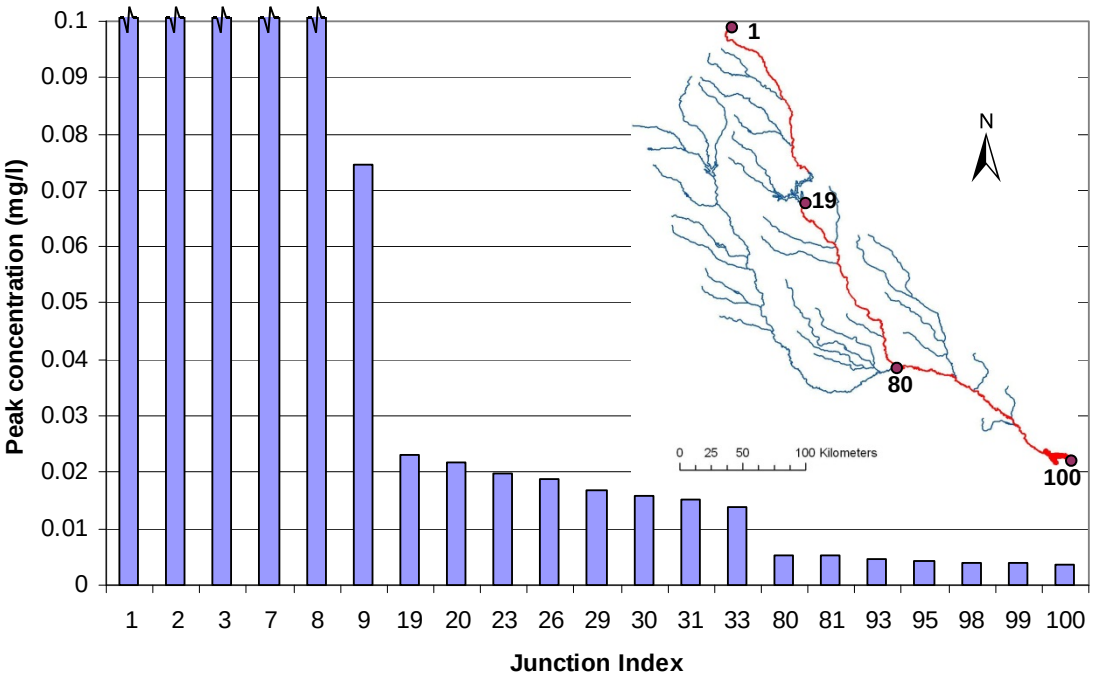


Figure 3.11 Transport and fate of peak concentration of a contaminant plume.

In the Scenario set A, single instantaneous spill events occurring at random junctions are considered and all contamination events are assumed to occur at the beginning of the simulation period which is selected as 15 days, i.e. $t_{s,inj} = 0$ and $t_{sim} = 21600\ min$ in Equation (3.1). Therefore, for this case the number of possible spill events to be simulated is *S=100*.

The Scenario set B consists of spill events in scenario set A repeated at different spill times. In this scenario set, the simulation time is reduced to 4 days consistent with the contaminant travel time in the river network and the contamination events are assumed to occur in a fraction of this simulation period indicated as t_{sp} as shown in Figure 3.12.

Thus t_{sp} is the maximum time period for possible spills to occur. For the Scenario set B, t_{sp} is selected as 2 days and 48 different injection times, $t_{s,inj}$ evenly distributed in this time period are considered. Therefore, number of scenarios to be simulated in the Scenario set B is *S=4800.*

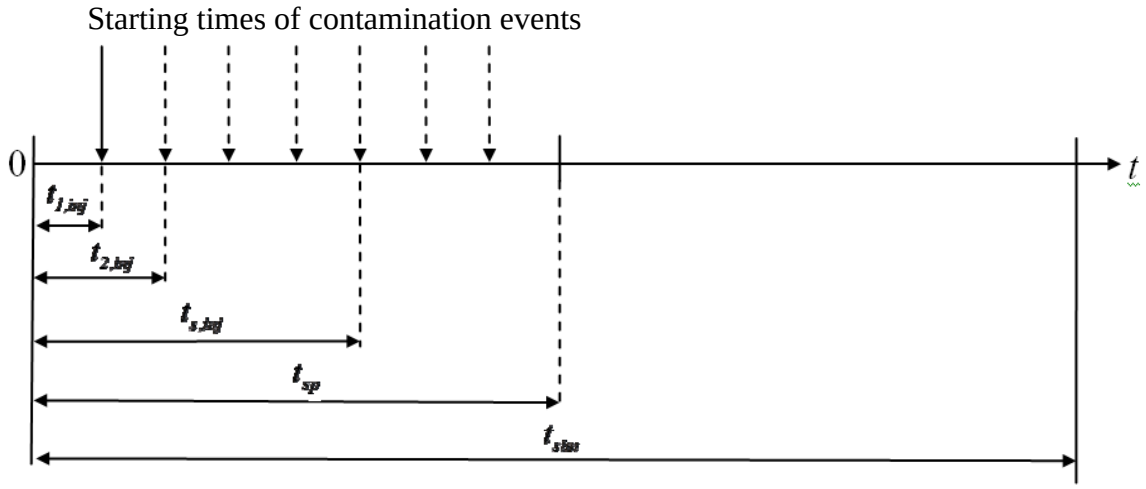


Figure 3.12 Spill times for scenario sets B and C.

The Scenario set C is designed to represent unsteady hydrodynamic behavior of Altamaha watershed. For this purpose, Altamaha basin is divided into 100 sub-catchments using digital elevation data as described in Figure 3.13 where a sample set of these sub-catchments are presented to indicate their selection process. These sub-catchments are determined such that each area drains rain water to one of the 100 nodes selected as possible spill locations shown in Figure 3.10. Among these 100 sub-catchments, 24 of are selected to receive rainfall. Thus, 2400 scenarios are generated such that each scenario has a contamination location randomly selected among 100 nodes, and a random contamination time selected between 0 and 48 hours i.e. $t_{sp} = 2\,days$ as

described in Figure 3.12. Also, in each scenario, there are random rain events occurring in 12 sub-catchments randomly selected among 24 sub-catchments. Each rain event also has a randomly assigned rainfall intensity as 1, 3 and 5 inch/hr and a random rainfall duration selected among 1, 2 and 3 hours. Furthermore, each rainfall event has a random start time selected among 0, 12, 24, 36, 48, 60 and 72 hours. The simulation time is also set to 4 hours in Scenario set C.

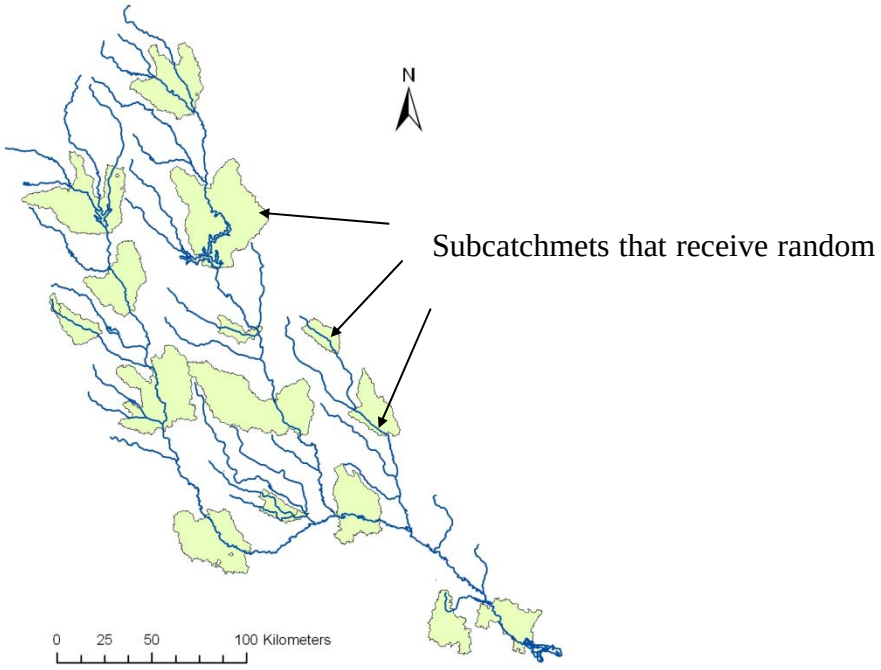


Figure 3.13 Watershed delineation for Altamaha basin.

3.3.2.2 Optimal Monitoring Locations for the Scenario Set A

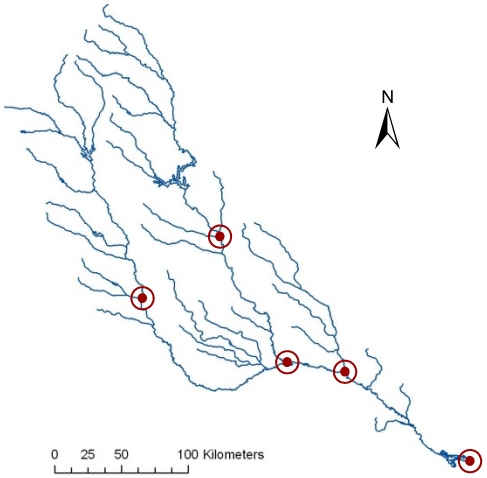
For this scenario set, optimal solution is obtained for several different numbers of monitoring stations such as $(M_o = 5, 6, 7, 10 \; and \; 20)$. For this analysis both GA and enumeration methods are used for the cases in which 7 or less number of monitoring

stations are to be located since the total number of possible placements for a monitoring network composed of 7 monitoring stations is $P_7 = C_7^{100} \cong 1.6 \cdot 10^{10}$ for which all possibilities can be searched in a reasonable computational time. However for 10 and 20 monitoring stations the number of possible placements would be very large (e.g. $P_{10} = C_{10}^{100} \cong 1.7 \cdot 10^{13}$). For those two cases GA is used to identify the best monitoring locations. In addition to the use of different numbers of monitoring stations, effect of detection threshold of the sensors on the optimal solution is also analyzed using two different threshold values of $C_{dt} = 0.0001$ mg/l and $C_{dt} = 0.01$ mg/l. This analysis is included to provide information on the effect of dilution on the selection of best locations of the monitoring stations.

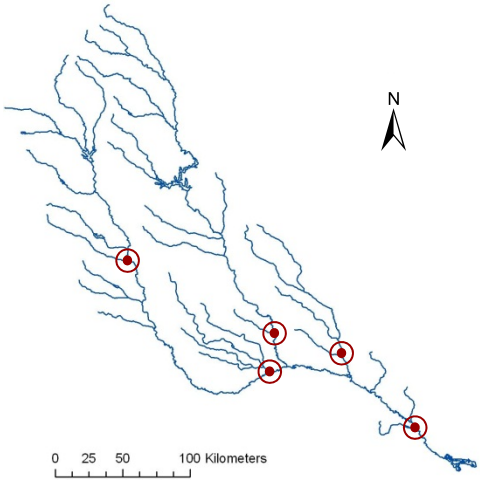
The best monitoring locations are determined by minimizing the average detection time and by maximizing the detection likelihood as given in Equation (3.6). In Figure 3.14, results of this optimization problem for a monitoring network of 5 stations are summarized. Figure 3.14(a) shows that for the low detection threshold, the optimization algorithm locates one of the monitoring stations at the outlet to reach 100% reliability. When the solutions for the 100%, 80% and 60% reliabilities are considered for the same detection threshold value, it can be seen that for lower reliability values, the optimization algorithm yields more upstream junctions. Figure 3.14(b) shows that the maximum reliability that can be obtained by 5 monitoring stations with high detection threshold value is 96%. This result indicates that in cases where dilution is important, the most downstream monitoring station may not necessarily be at the outlet. When Figure 3.14(a)

and (b) are compared, it can be seen that detection threshold value changes the optimum solution for maximum reliability values and for the 80% reliability. However, for 60% reliability solutions the best result for both detection threshold values are the same. The reason behind this outcome is that 60% reliability value is so low that the optimal solution is independent of the detection threshold of the monitoring system.

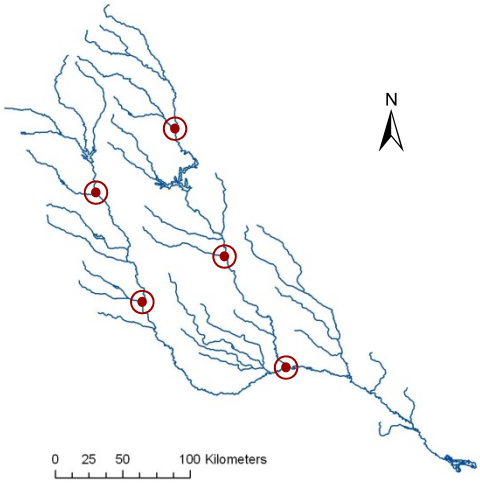
As the number of monitoring locations increase, the additional stations are placed in more upstream nodes as shown in Figure 3.15 and Figure 3.16. In Figure 3.15(b), it can be seen that the maximum reliability value that can be reached by a 6-station monitoring system with a detection threshold of $C_{dt} = 0.01$ mg/l is 98%. In a similar way, Figure 3.16(b) shows that 100% reliability can be reached, if the number of stations is increased to 7 for the same detection threshold value. Therefore, it can be concluded that for Altamaha River, at least 7 monitoring stations are required to maintain reliability of the system at 100% if the monitoring devices have a detection threshold of $C_{dt} = 0.01$ mg/l.



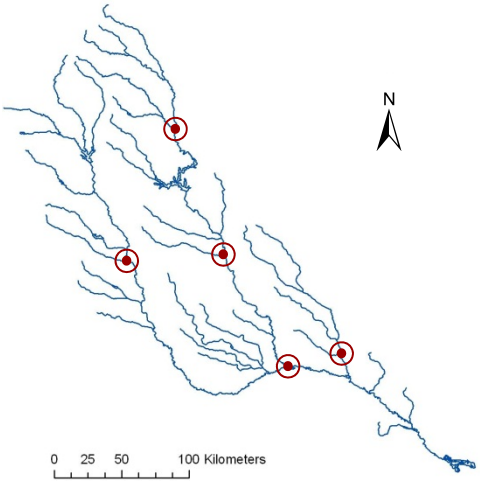
100% Reliability



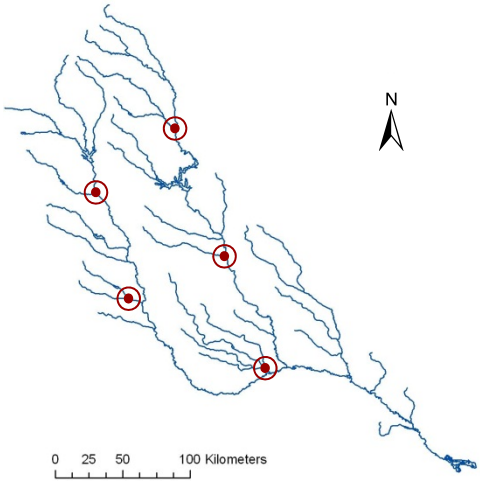
96% Reliability



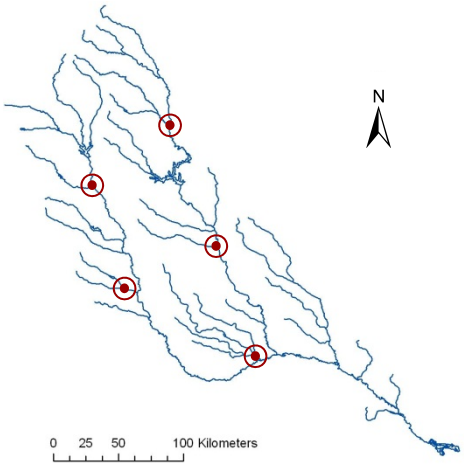
80% Reliability



80% Reliability



60% Reliability

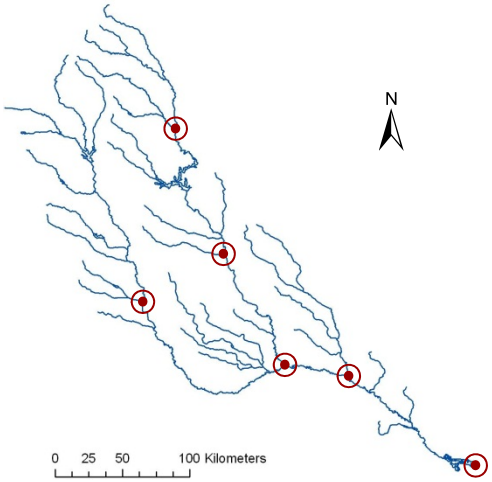


60% Reliability

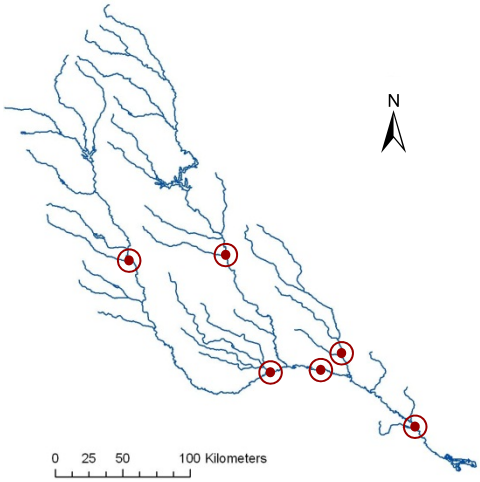
(a) Detection threshold: 0.0001 mg/l

(b) Detection threshold: 0.01 mg/l

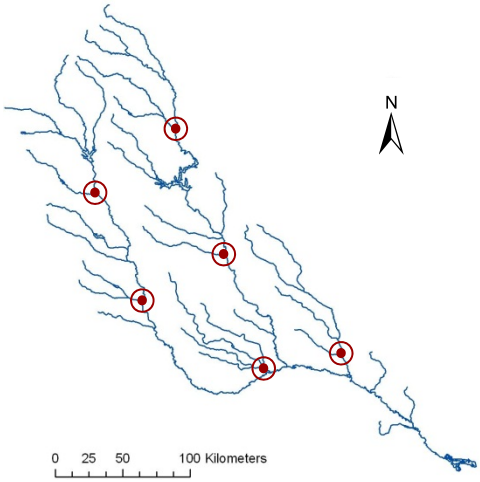
Figure 3.14 Optimum locations for 5 monitoring stations for Scenario set A.



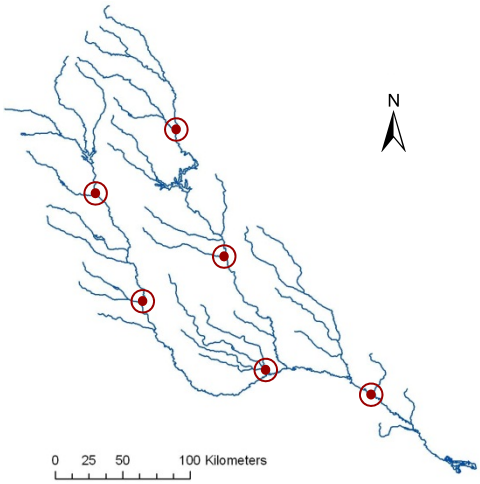
100% Reliability



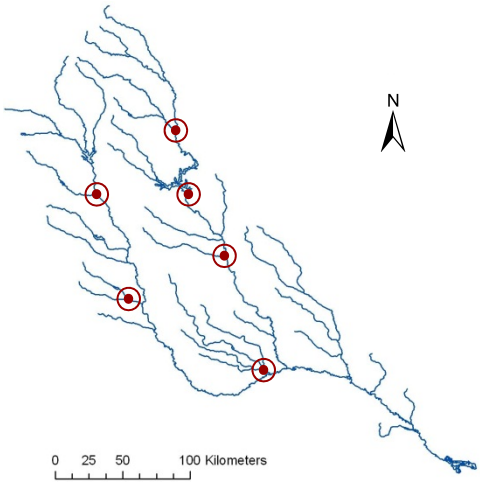
98% Reliability



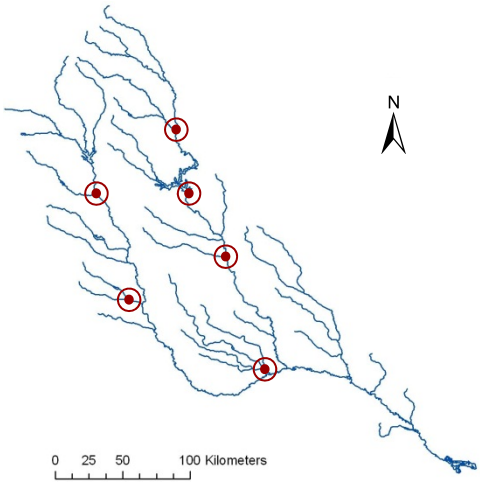
80% Reliability



80% Reliability



60% Reliability

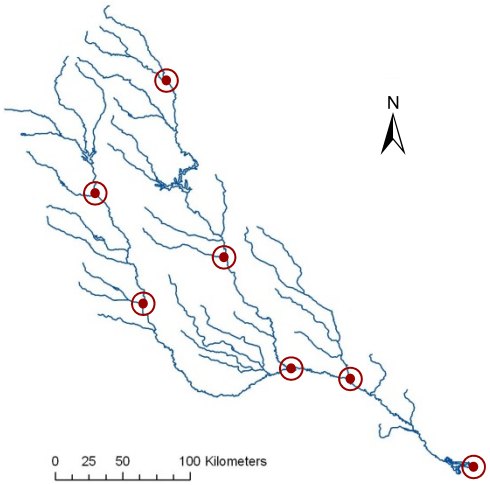


60% Reliability

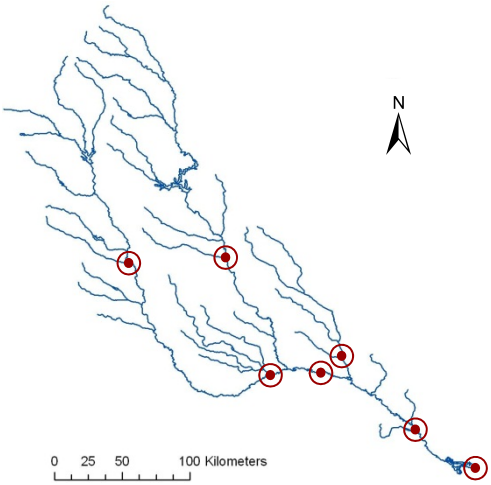
(a) Detection threshold: 0.0001 mg/l

(b) Detection threshold: 0.01 mg/l

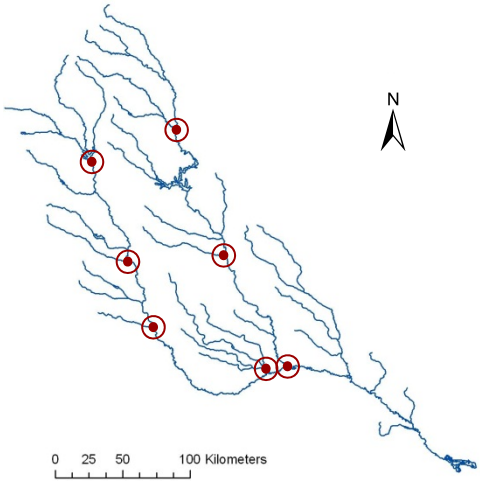
Figure 3.15 Optimum locations for 6 monitoring stations for Scenario set A.



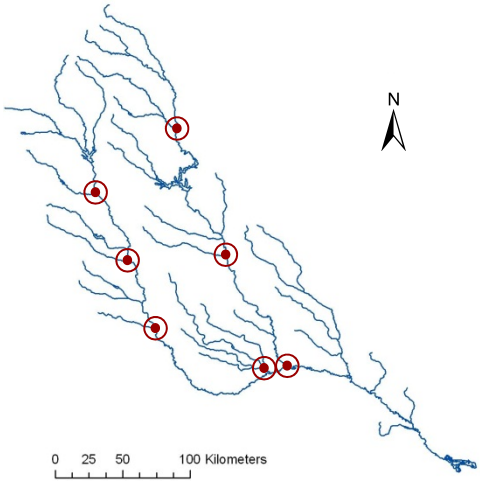
100% Reliability



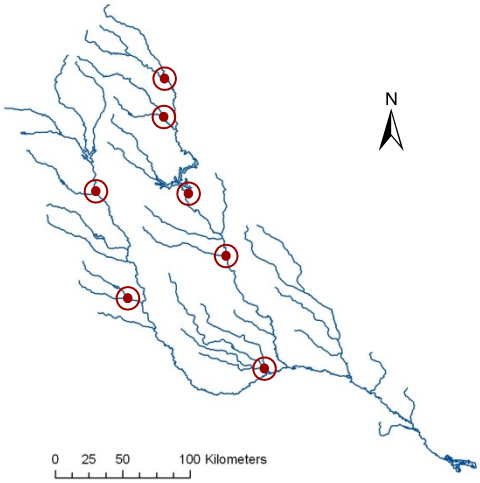
100% Reliability



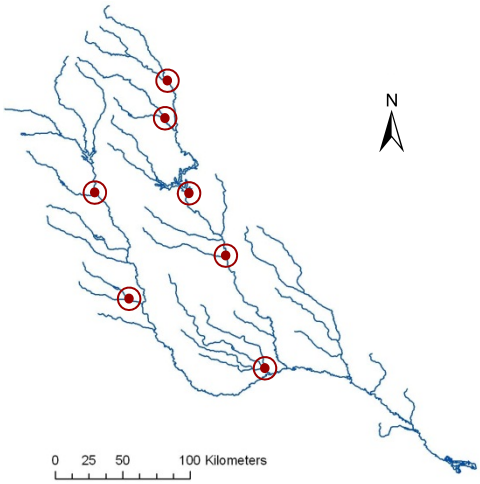
80% Reliability



80% Reliability



60% Reliability



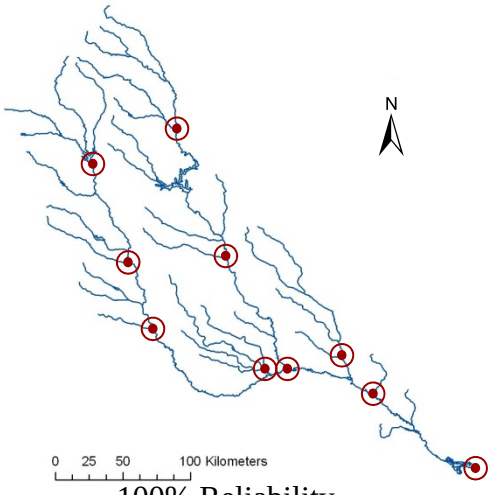
60% Reliability

(a) Detection threshold: 0.0001 mg/l

(b) Detection threshold: 0.01 mg/l

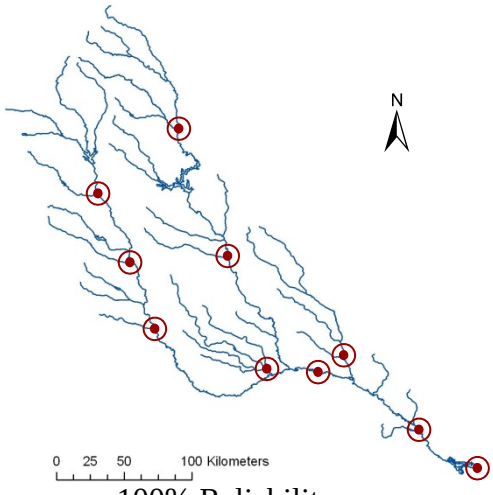
Figure 3.16 Optimum locations for 7 monitoring stations for Scenario set A.

Another interesting result that can be observed in Figure 3.15(b) and Figure 3.16(b) for the maximum reliability values is that the optimization algorithm does not always prefer the junctions at the confluences of the river reaches. This shows that internal nodes may also be selected as the monitoring locations as long as the average detection time is minimized. This fact can also be seen in cases where number of monitoring stations is further increased to 10 and 20 as shown in Figure 3.17 and Figure 3.18 in which optimum locations only for 100% reliability values are presented. When these figures are considered, it can be said that as the number of monitoring stations increases in a monitoring system, the effect of detection threshold (or dilution) on the optimum placement of the stations decreases as expected. This fact can easily be observed in Figure 3.18 where the location of only one monitoring station is changed when the detection threshold is increased from $C_{dt} = 0.0001$ mg/l (Figure 3.18(a)) to $C_{dt} = 0.01$ mg/l (Figure 3.18(b)).



100% Reliability

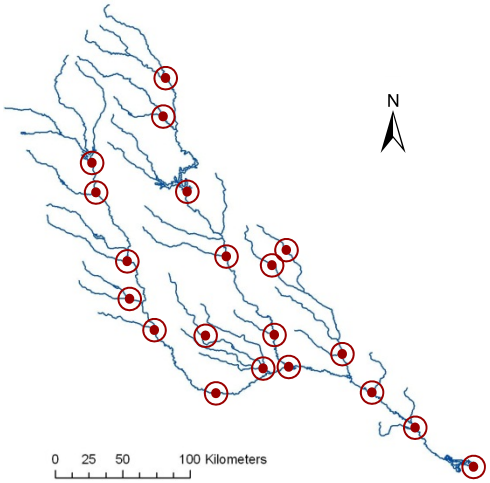
(a) Detection threshold: 0.0001 mg/l



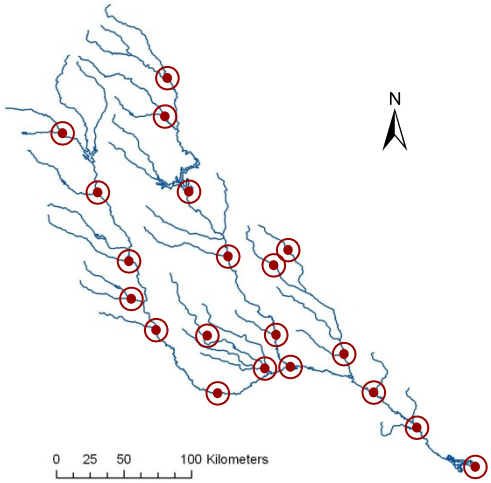
100% Reliability

(b) Detection threshold: 0.01 mg/

Figure 3.17 Optimum locations for 10 monitoring stations for Scenario set A.



(a) Detection threshold: 0.0001 mg/l



(b) Detection threshold: 0.01 mg/l

Figure 3.18 Optimum locations for 20 monitoring stations for Scenario set A.

In Table 3.5, minimum average detection time values obtained for all monitoring systems for different numbers of stations and for different reliabilities are listed. From this table, it can be easily observed that as the reliability of the monitoring network decreases, the minimum average detection time increases. This outcome comes from the fact that the penalty value used in the calculation of detection time for the non-detected spill scenarios is very high. Another outcome that can be deduced from Table 3.5 is the fact that as the detection threshold (or dilution) increases, the minimum average detection time that can be reached by a monitoring network also increases. This is an expected result because when the concentration distribution within a contaminant plume is considered, it is obvious that regions of the plume with higher concentrations will reach the sensor at later times than the lower concentration regions. However, as mentioned earlier, effect of detection threshold on the minimum average detection time decreases as the reliability

value of the monitoring system decreases and this effect become insignificant in determining best monitoring locations after some low reliability value.

Table 3.5 Summary of optimum solutions for Scenario Set A.

Number of Stations	Detection threshold=0.0001 mg/l		Detection threshold=0.01 mg/l	
	Reliability (%)	Average Detection time (min)	Reliability (%)	Average Detection time (min)
5	100	1323	96	2706
	80	4988	80	5181
	60	9017	60	9095
6	100	1120	98	1898
	80	4894	80	4999
	60	8972	60	9042
7	100	926	100	1471
	80	4836	80	4945
	60	8911	60	8971
10	100	695	100	905
20	100	366	100	456

Finally, change of average detection time with the number of monitoring stations with different detection thresholds for the highest reliability values is presented in Figure 3.19. This figure shows that for small number of monitoring stations, average detection time sharply decreases as new stations are added to the system. However, for monitoring systems with larger number of stations, additional stations result in more gradual decrease in average detection time. Thus, Figure 3.19 can be used as a guideline for decision makers in adjusting the budget for the monitoring system design. Another important outcome that can be deduced from Figure 3.19 is the observation that as the number of monitoring stations increases, the difference between the average detection times of the monitoring system with different detection threshold values decreases.

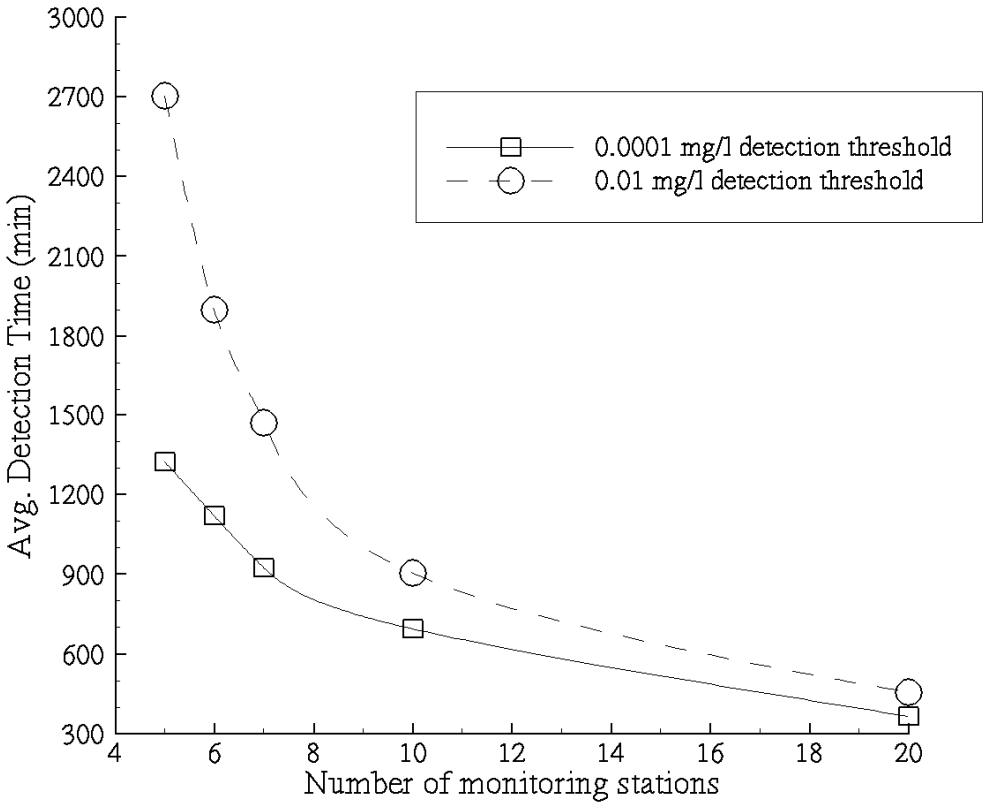


Figure 3.19 Change of average detection time with the number of monitoring stations for maximum reliability values for Scenario set A.

3.3.2.3 Optimal Monitoring Locations for the Scenario Set B

In the application discussed above it can be seen that optimization algorithm can be used to select a single best solution given the scenarios and the reliability levels considered. The reason behind this outcome is the fact that all contamination events in scenario set A occur at the beginning of simulation period with a constant penalty of a large simulation time in case of non detection. Since in Scenario set B, simulation time is smaller and

penalty values of the scenarios are variable due to the variable spill times, the optimization algorithm determined 5 non-dominated best placement schemes for a monitoring project composed of 5 stations with a detection threshold of $C_{dt} = 0.01$ mg/l. In Figure 3.20 we present some of the optimal results obtained in terms of average detection time and reliability. In this figure, it can be seen that all the solutions represented by empty diamonds are dominated by the solutions represented by shaded diamonds. However, no dominance can be claimed among the shaded diamonds which indicate that these solutions constitute a Pareto front. In Figure 3.21, locations of the monitoring stations that correspond to the optimum solutions on the Pareto front are shown.

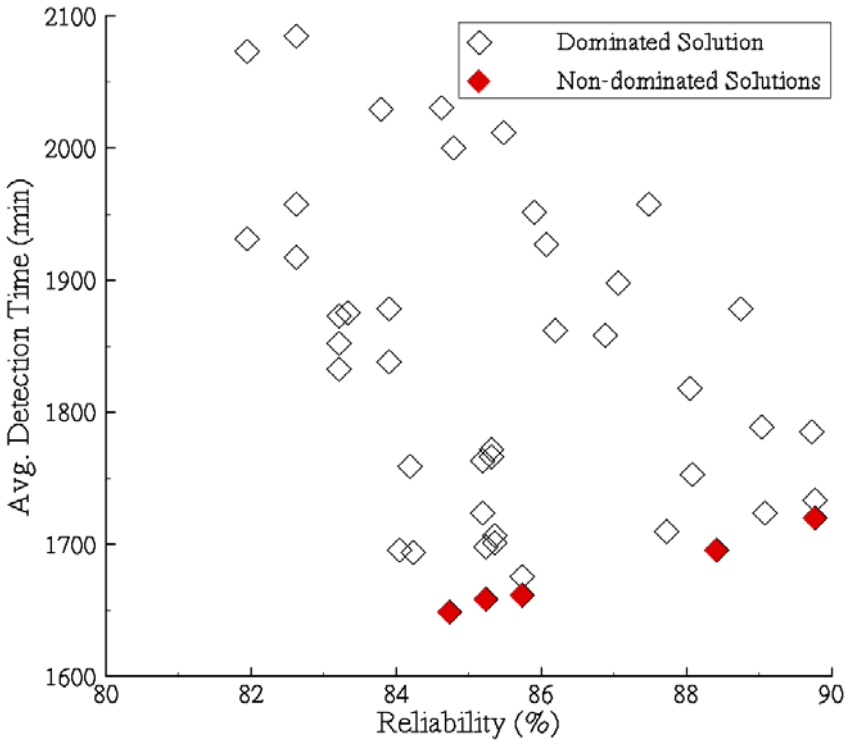


Figure 3.20 Pareto front for scenario set B.

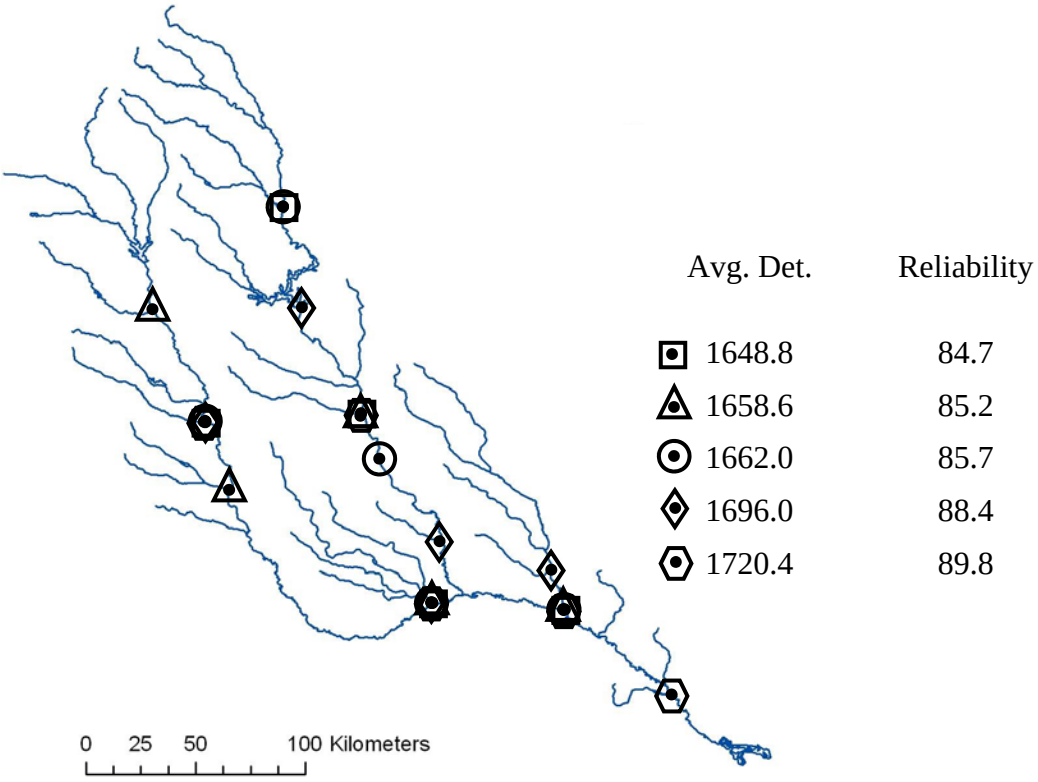


Figure 3.21 Locations of monitoring stations for the solutions on the Pareto front for the Scenario set B.

3.3.2.4 Optimal Monitoring Locations for the Scenario Set C

For scenario set C in which transient hydrodynamic behavior of Altamaha watershed is considered using random rainfall events on randomly selected watershed, the optimization algorithm determined 8 non-dominated best placement schemes for a monitoring project composed of 5 stations with a detection threshold of $C_{dt} = 0.01$ mg/l. In Figure 3.22 some of these results are presented for selected optimal outcomes in terms of average detection time and reliability (presenting all dominated solutions in this figure would yield a very crowded figure). In this figure, a Pareto front is formed by the shaded

diamonds since they both have a shorter average detection time and higher reliability than those of all other solutions represented by empty diamonds according to Pareto front definition. In Figure 3.23, the locations of the monitoring stations corresponding to the optimum solutions that are on the Pareto front are shown. In this figure, it can be seen that all of the 8 monitoring systems (solutions) have monitoring stations at the locations indicated by arrows. With this outcome, one can argue about the importance of these locations, since which optimal solution is selected from the Pareto set, there will be a station placed at these nodes.

When the outcomes of Scenario Sets B and C are compared, it can be observed that the number of solutions in Pareto front is increased from 5 (Scenario Set B) to 8 (Scenario Set C). The reason behind this result can be the increased level of randomness in scenario generation for Set C. When Figure 3.21 and Figure 3.23 are considered, it can be observed that the ranges of average detection time and reliability values in the Pareto set are similar for Scenario sets B and C (e.g. reliability values on Pareto front for Set B take values from 84.7% to 89.8% and they take the values between 84.3% to 90.5% for Scenario Set C). However, final monitoring system designs reached by these two sets of scenarios are quite different. As an example, the optimization utilizing from the Scenario Set B prefers more number of nodes on Ocumulgee River (Figure 3.7) than the optimization using the Scenario Set C.

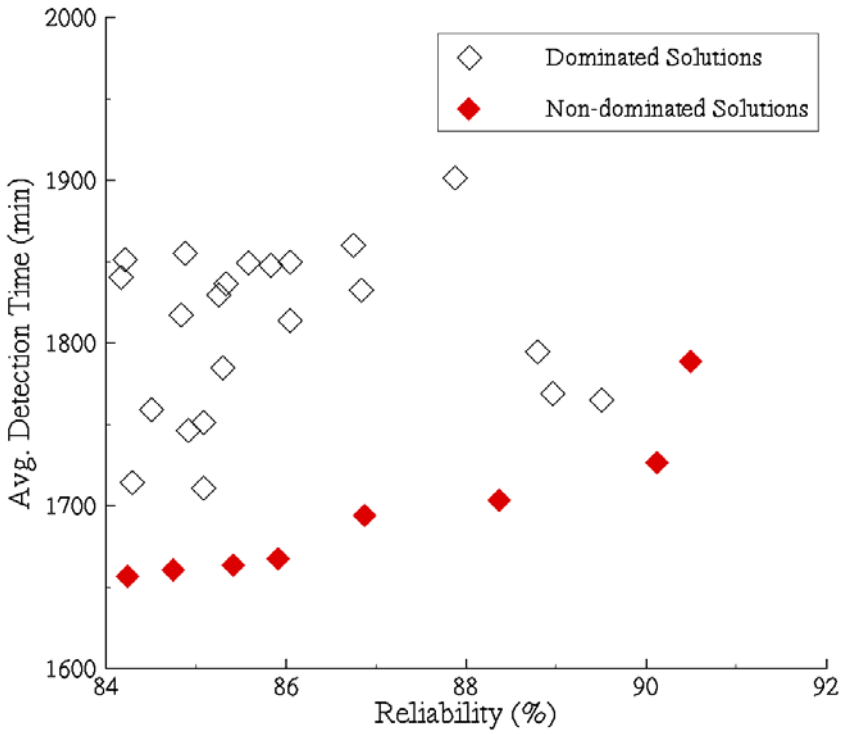


Figure 3.22 Pareto front for Scenario set C.

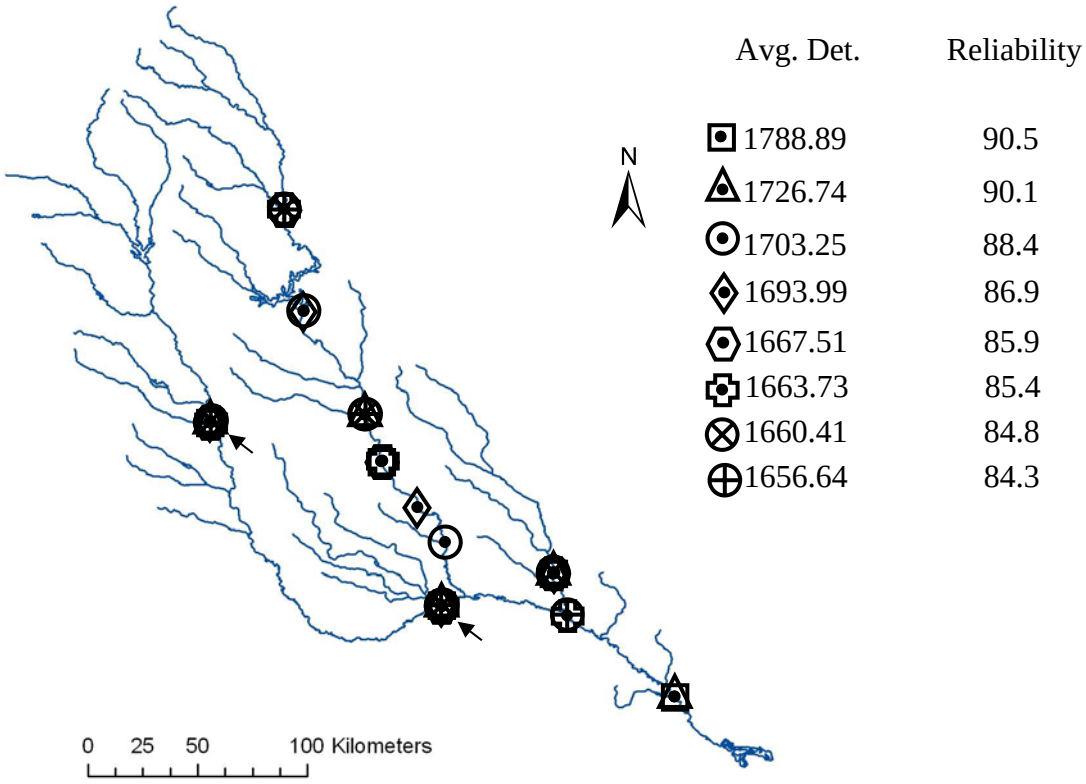


Figure 3.23. Locations of eight monitoring station sets for the solutions on the Pareto front for Scenario set C.

3.4 Conclusions

A methodology is proposed to identify the locations of the best monitoring stations in a river system. In the proposed method, hydrodynamics and the fate and transport of the contaminants play important role in the determination of the best monitoring locations. In optimization, genetic algorithms are used and results are verified by enumeration when possible. Optimization method is applied to the Altamaha river network and several monitoring systems are designed and compared. It is observed that optimization parameters such as simulation time and penalty values have important effects on the

diversity of the optimum solutions. It is demonstrated that in some cases non-dominated solution sets can be obtained where a Pareto front analysis would be necessary. The results indicate that, in the Altamaha river system application discussed in this thesis, certain common patterns can be observed in the optimal monitoring location sets obtained. This may or may not be the case for other river networks. However, in the Altamaha river system case these common patterns indicate that preferred locations of monitoring stations are apparent for the objectives that are considered in this study. Results are promising in terms of applicability of the proposed methodology to the monitoring network system design.

The analysis presented is based on two objectives: i.e. the detection time and reliability. The proposed method may also be extended to include other objectives such as minimal cost of operation of the system, the suitability of access to monitoring stations, cost of sensors, and use of non-conservative chemicals in the analysis. These are minor extensions of the overall methodology described here and are not considered in this case.

CHAPTER 4

IDENTIFICATION OF CONTAMINANT SOURCE LOCATION IN RIVER NETWORKS

4.1 Introduction

The improving real-time monitoring technologies introduce new tasks to these systems such as rapid identification of contamination source locations. In theory, this is an ill posed problem which has non-unique solutions due to the irreversible nature of contaminant transformation and transport processes. In this study, we propose a methodology that utilizes a classification routine which associates the observations on a contaminant spill to one or more of the candidate spill locations. This approach consists of a training step followed by a sequential elimination of the candidate locations which leads to the identification of spill locations. The training of the monitoring system although requires a significant simulation time, it is performed only once. The statistical elimination for the ranking of the candidate locations is a rapid process. The proposed methodology is applied to the Altamaha river system in the State of Georgia, USA. The results show that the proposed approach can be effectively used for the preliminary planning of the contaminant source investigation studies in complex river systems.

4.2 Methodology

The proposed methodology is based on three steps: (i) obtaining contaminant concentration breakthrough curves for the spill scenarios at the monitoring stations under random hydrodynamic behavior of the river system. Features of each of these breakthrough curves quantified and stored in this step are going to be utilized in the second step; (ii) estimation of conditional probability density functions of each feature given that the spill has occurred at a known location. These probability densities once obtained are generic for the river system under consideration and steps (i) and (ii) are performed only once. These first two steps can also be interpreted as the training process of the proposed methodology. Finally, in step (iii), possible spill locations are determined by applying adaptive sequential feature selection (ASFS) algorithm to the conditional probabilities obtained in step (ii).

4.2.1 Features of Breakthrough Curves

When a spill occurs at a random location in a river system, the contaminant mass undergoes transformation and transport processes such as advection, diffusion, dispersion and chemical reactions as it travels downstream. When the plume reaches to the monitoring station, its concentration can no longer be measured as a unit pulse concentration. Instead, due to the mixing effects of diffusion and dispersion processes, the monitoring station will observe a distribution of concentration over time as a temporal

breakthrough curve. Breakthrough curves can therefore be interpreted as and have the same properties with the probability density curves (Grubner 1971). For some monitoring stations, it is also possible to measure the average velocities or the flow rates. In this case, the monitoring system provides both temporal and volumetric distribution (breakthrough curve) of the contaminant concentration (Brooks and Wise 2005). Either temporal or volumetric, the breakthrough curves can be characterized by a series of statistical moments. The ordinary statistical moments are given in Equation (4.1).

$$m_k^{\Psi} = \frac{\int_0^{\infty} C(\Psi)\Psi^k d\Psi}{\int_0^{\infty} C(\Psi) d\Psi} \tag{4.1}$$

where Ψ is a generic variable which takes the values of time or cumulative volume, $C(\Psi)$ is the concentration measured at a given Ψ and m_k^{Ψ} is the k^{th} temporal or volumetric (depending on Ψ) ordinary statistical moment. Then, the central statistical moments can be estimated about the first ordinary statistical moment.

$$\overline{m_k^{\Psi}} = \frac{\int_0^{\infty} C(\Psi)\left(\Psi - m_1^{\Psi}\right)^k d\Psi}{\int_0^{\infty} C(\Psi) d\Psi} \tag{4.2}$$

The zeroth and the first ordinary moments (m_0^{Ψ} and m_1^{Ψ}) represents the area under the normalized distribution curve and the mean of this distribution respectively. More precisely, m_0^{Ψ} always has the value 1 and the value of m_1^{Ψ} depends mainly on the spill time. Therefore, these two ordinary moments do not give sufficient information to characterize breakthrough curves. The second central moment $\overline{m_2^{\Psi}}$ quantifies how wide the breakthrough curve is and represents the variance of the concentration distribution curve. The standard deviation of this curve, σ^{Ψ}, can be estimated by the square root of the second central moment.

$$\sigma^{\Psi} = \sqrt{\overline{m_2^{\Psi}}} \tag{4.3}$$

The third central moment is used to quantify the asymmetry of the concentration distribution curve relative to a normal distribution. This measure is called skewness and calculated as shown in Equation (4.4).

$$S^{\Psi} = \frac{\overline{m_3^{\Psi}}}{\left(\sigma^{\Psi}\right)^3} \tag{4.4}$$

A negative skewness value indicates that the concentration distribution curve has a tailing toward small values of Ψ . Another important feature that distinguishes a breakthrough curve from others is the excess. Excess gives information on the peaking of the concentration distribution curve compared to a Gaussian distribution and it is calculated from the fourth central moment.

$$E^{\Psi} = \frac{\overline{m_4^{\Psi}}}{\left(\sigma^{\Psi}\right)^4} - 3 \tag{4.5}$$

A positive excess value indicates that the distribution is more peaked relative to a Gaussian distribution which has an excess value of zero.

Since the fifth and higher statistical moments do not provide important information for the analysis of the probability curves, it can be said that breakthrough curves have three main features, which are obtained from the second, third and fourth statistical moments. Then, a spill may have six main features since a monitoring station generates a temporal and a volumetric breakthrough curve.

The real-time water quality monitoring stations provide time variation of the contaminant concentrations in the river water by electronic measurements with a certain time step. Therefore, the breakthrough curve is observed at some finite number of times and Equations (4.1) and (4.2) should be approximated accordingly. Although the time step between two consecutive observations may most of the time be constant, the volume that

passes the monitoring station at every time step will not be constant due to the unsteady nature of the river flow. Therefore, the generic equations for the moments can be approximated such that the difference in Ψ values is variable and trapezoidal rule is applied for the approximation.

$$m_k^{\Psi} \cong \frac{\sum_{i=1}^{n-1}\left[C(\Psi_i)\Psi_i^{\,k} + C(\Psi_{i+1})\Psi_{i+1}^{\,k}\right](\Psi_{i+1}-\Psi_i)}{\sum_{i=1}^{n-1}\left[C(\Psi_i)+C(\Psi_{i+1})\right](\Psi_{i+1}-\Psi_i)} \tag{4.6}$$

$$\overline{m_k^{\Psi}} \cong \frac{\sum_{i=1}^{n-1}\left[C(\Psi_i)\left(\Psi_i - m_1^{\Psi}\right)^k + C(\Psi_{i+1})\left(\Psi_i - m_1^{\Psi}\right)^k\right](\Psi_{i+1}-\Psi_i)}{\sum_{i=1}^{n-1}\left[C(\Psi_i)+C(\Psi_{i+1})\right](\Psi_{i+1}-\Psi_i)} \tag{4.7}$$

where n is the total number of observations in the breakthrough curve.

It should be noted that this study assumes that contaminants are conservative substances thus chemical reactions are neglected. The choice of conservative substances is not a limitation on the methodology proposed. However, in that case non-conservative contaminants should be used to develop the training scenarios as discussed below.

4.2.2 Training Process

The proposed methodology requires a large number of contamination scenarios with known source locations in order to construct the training sets for the monitoring stations. A training set is constructed by several spill scenarios with random spill mass, spill time and river hydrodynamic parameters at every candidate spill location. Each of these scenarios creates a breakthrough curve. The features of these breakthrough curves can be estimated as explained in the previous section. If each breakthrough curve has d features which are continuous random variables, then they form a random feature vector X. If we let there are M available spill locations (or classes) in the class set $\Omega = \{1,2,...,M\}$ then the class $Y \in \Omega$ is a discrete random variable. The spill location (or the class) of a vector X is known in the training process. Therefore, one can obtain the conditional probability density function of the feature vector X given that it belongs to a class of Y, $f(X|Y)$. However, in this study we are interested in the probability of occurrence of a spill at a candidate location given that its breakthrough curve has the observed features, i.e. $P(Y|X)$. This probability can be estimated using Bayes' theorem as described in Equation (4.8).

$$P(Y = j \mid X_i = x_i) = \frac{f(X_i = x_i \mid Y = j)P(Y = j)}{\sum_{n \in \Omega} f(X_i = x_i \mid Y = n)P(Y = n)} \tag{4.8}$$

where j is an integer specifying the class of the observed features, X_i and x_i represent the i^{th} feature and its value respectively. $P(Y = j)$ is the prior probability of class j to be the spill location. In this study, it is assumed that all candidate locations have the same prior probability of being the spill location, i.e. $P(Y = n) = \frac{1}{|\Omega|}, \forall n \in \Omega$.

4.2.3 Adaptive Sequential Feature Selection

Once a spill is detected at the monitoring station and its features described in the previous section are calculated, the next step is to decide the location of the spill. In this study, we formulate this question as a classification problem. This formulation requires a finite set of classes which are preselected nodes as the candidate spill sites located upstream of the monitoring station. Then, the spill monitored is classified into one of these candidate spill locations according to its observed features. For this classification an adaptive feature selection (ASFS) algorithm proposed by Jiang (2008) is implemented.

ASFS algorithm is a classification tool to label an arbitrary test data which has the features $x = [x_1, x_2, ..., x_d]^T$ with one of the M classes by optimizing a criterion associated with the statistical entropy of each class conditioned on observations described by x.

In information theory, entropy is a measure of uncertainty of a random variable or process (Shannon 1948). ASFS algorithm uses entropy to measure the uniformity of the probability distributions. The entropy of a probability distribution increases as it approaches to a uniform distribution. The entropy of a discrete random variable S which may have N different values with the probabilities $P(S = s_n)$, where $n = 1,2,...,N$ is defined in information theory as:

$$H(S) = -\sum_{n=1}^{N} P(S = s_n) log\left[P(S = s_n)\right] \tag{4.9}$$

Then, the entropy for the prior probability mass function for the classes Y and the class entropy conditioned on the feature x_i can be calculated as described in Equations (4.10) and (4.11) respectively.

$$H_0 = -\sum_{i=1}^{M} P(Y = i) log\left[P(Y = i)\right] \tag{4.10}$$

$$H_i = -\sum_{j=1}^{M} P(Y = j \mid X_i = x_i) log\left[P(Y = j \mid X_i = x_i)\right] \tag{4.11}$$

where $P(Y = j \mid X_i = x_i)$ is the conditional probability of Y being j given that the i^{th} feature has the value of x_i. The ASFS algorithm calculates the quantity $q(x_i) = H_0 - H_i$

to find the best feature to simplify the classification process. Thus, the feature, x_*, which maximizes the reduction in the class entropy can be determined by Equation (4.12).

$$x_* = arg\,max\; q(x_i) \tag{4.12}$$

ASFS algorithm provides a systematic process which iteratively eliminates one or more elements from a set of classes at each stage of the search and selects class of the tested sample as the final class remaining in the set. To summarize this iterative process, let Ω_k be the set of candidate classes survived until the k^{th} stage of the algorithm and $|\Omega_k|$ be the number of elements in this set, where k is an integer greater than zero. Then, as shown in Figure 4.1, ASFS algorithm first calculates the quantity $q(x_i)$ for all features using Equations (4.10) and (4.11). The next step is to determine the best feature, x_*, for the class elimination process by satisfying Equation (4.12). In the class elimination process, probability of each class $\Omega_k(i)$ conditioned on the selected feature, x_*, is tested against a threshold T and the classes which have a probability value greater than this threshold is kept in a set Ω' which was initially an empty set. Once all the classes are tested the new class set Ω_{k+1} takes the elements of Ω'. This procedure is repeated until Ω' has only one element which will be assigned as the class label of the sample. Here, Jiang (2008) recommends to use the median value of the probabilities $P(Y = j \mid X_i = x_i)$ as the threshold value which eliminates half of the classes at each stage. However, in this study, our aim is to rank the classes rather than a quick elimination. Thus, we use the

minimum value of these conditional probabilities as the threshold which eliminates only one class most of the time. In this thesis, we have provided a brief description of ASFS algorithm including the details that are essential for this study. Other details such as properties of this algorithm and its comparison with another pattern recognition method can be found in Jiang (2008).

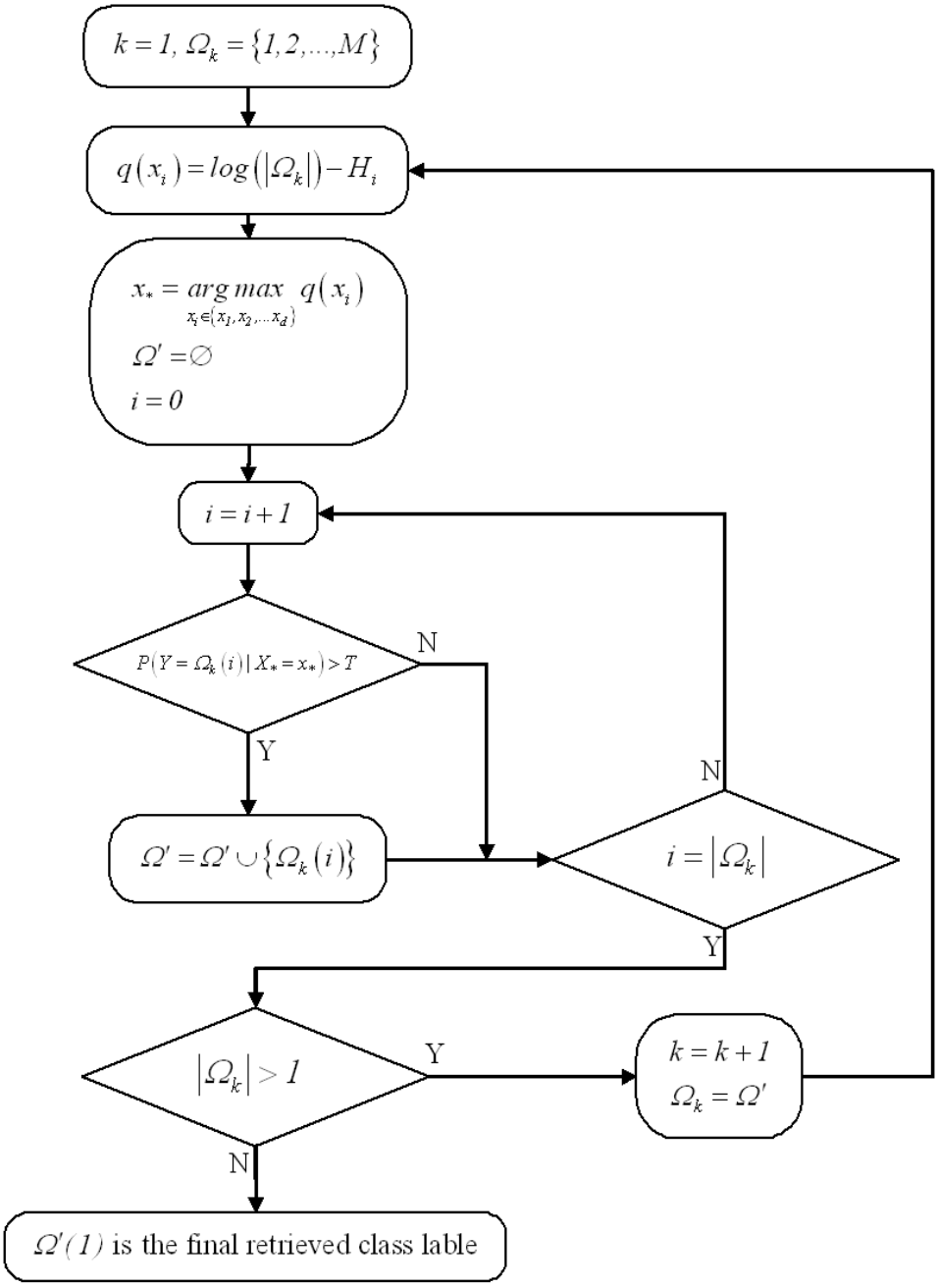


Figure 4.1 Flow chart for ASFS algorithm

4.3 Applications

The proposed methodology is tested for several spill realizations that might occur on Altamaha river system. For this purpose, first a monitoring system is designed for this large river system (Telci, Nam et al. 2008; Telci, Nam et al. 2009). Then, monitoring stations are trained for the ASFS algorithm by a large number of spill scenarios to obtain the required probability distributions of the features of breakthrough curves observed at these stations. Finally, the spill realizations which occur at random locations are used to evaluate the performance of the methodology. This evaluation is performed by taking into consideration the temporal and volumetric breakthrough curves.

4.3.1 Study Area

The methodology is applied to the Altamaha river network as described in section 3.2.2. However, in this part of the study, a variable roughness distribution with higher values (up to 0.077) for upstream river reaches is sequentially developed throughout the river network as given in Table 4.1 where upstreamness factor is defined in Figure 3.9 of Section 3.2.2.

Table 4.1 Manning's roughness coefficients used for the analysis

Upstreamness Factor	n	Upstreamness Factor	n
1	0.077	10	0.041
2	0.070	11	0.039
3	0.064	14	0.034
4	0.060	25	0.023
5	0.055	29	0.021
6	0.052	30	0.020
7	0.049	31	0.020
8	0.046	32	0.020
9	0.043		

Once the geometric parameters of the river system are determined, the next step is to estimate flow conditions in the river channels. For this purpose a two-step procedure is applied. In the first step, steady state flow rates also called the base flows in each river reach is estimated using the data provided by USGS gauging stations. The steady state hydraulic system is calibrated for the flow pattern in the river based on the data obtained from annual average flow rates measured in 2006 at 20 USGS gauging stations distributed throughout the river network. The inflow discharges are assigned at the most upstream sections of the river network such that the flow rates at the USGS gauging stations obtained from the steady state hydrodynamic simulation became consistent with field measurements.

In the second step of flow estimation, the random, unsteady hydrodynamic behavior of the river system is constructed for each scenario by superimposing the effect of the rainfall events on the base flows determined in the first step. For this purpose, rainfall measurements in the year 2006 from 10 different observation stations distributed throughout the watershed are obtained from National Climatic Data Center (NCDC 2009). A statistical analysis is performed on these measurements and it is observed that the daily rainfall measurements follow a log-normal distribution. Then, for the simulations, the random rainfall intensities are assigned from the same log-normal distributions for 10 different regions of the Altamaha watershed. Next, these random time series of rainfall intensities are assigned to 10 different catchments as shown in Figure 4.2. These 10 catchment areas are composed of 100 sub-catchments determined by watershed delineation such that the water accumulated on these sub-catchments is discharged to each of the 100 nodes indicated in Figure 4.2.

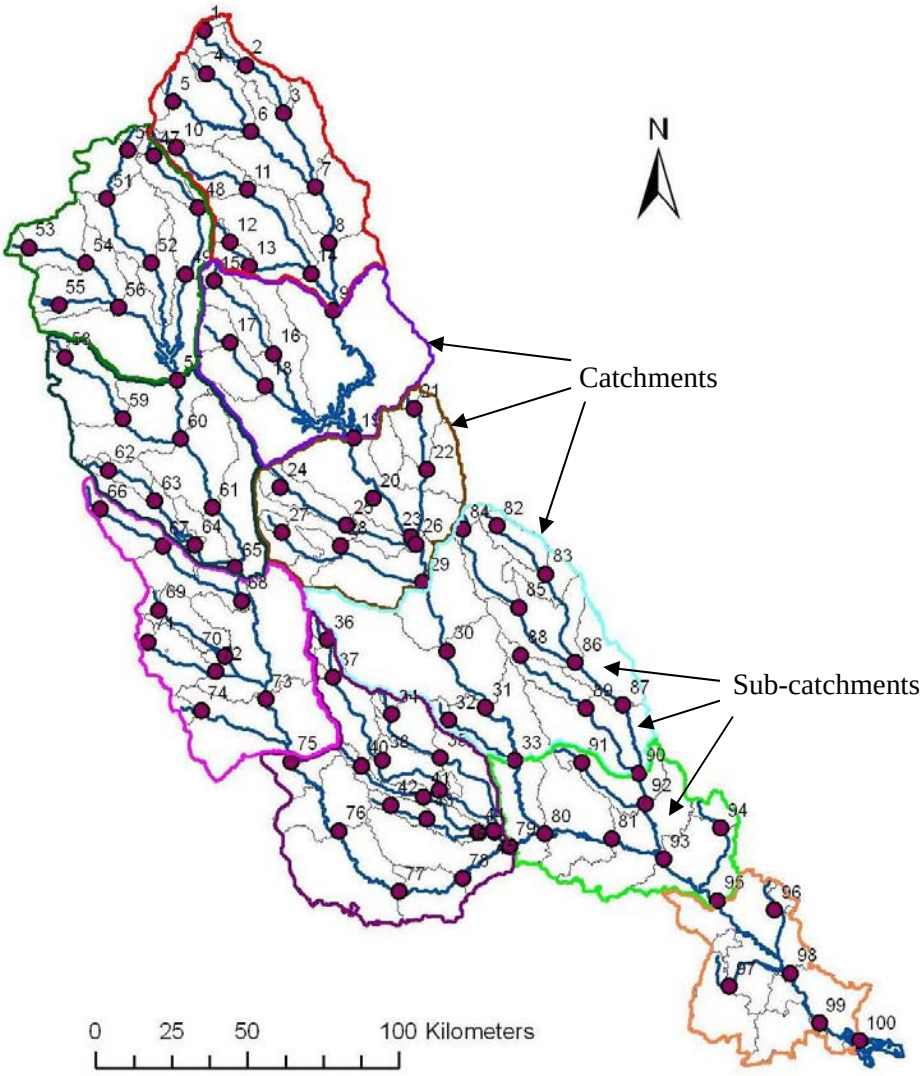


Figure 4.2 Watershed delineation for Altamaha basin.

4.3.2 Scenario Generation and Design of Water Quality Monitoring Network

In order to generate a training set for the proposed methodology, a large number of scenarios are designed such that the random, unsteady behavior of contaminant transport throughout the river network is represented. For this purpose, 1000 spill scenarios are produced for each of the 100 nodes indicated in Figure 4.2. In this study, only instantaneous spills are taken into account in this study. Each spill event has a random

spill time and spill mass. The random, unsteady hydrodynamic behavior of the river network is guaranteed by a random combination of unsteady rainfall intensities assigned to 10 different catchments described in Figure 4.2, which also includes no rainfall events. The training sets are generated only once for each monitoring station separately and the same training set can be used to determine different real spill locations.

Once these contamination scenarios are generated and simulated using SWMM, the results can be used to obtain breakthrough curve data for the training data set. Same breakthrough curves can be used to design an optimum real-time monitoring network for the river system. Telci, Nam et al. (2008) have developed a methodology for the optimal design of the monitoring systems in river networks. They determined the best locations for the monitoring stations by minimizing the average detection time and maximizing the expected reliability of the monitoring system. Then they applied this methodology to Altamaha river network (Telci, Nam et al. 2009). They showed that a Pareto optimal front made up of non-dominating solutions can be obtained for this optimization problem defined by two objective functions. In this study, a real-time monitoring system with 5 stations is designed for the Altamaha river using the same methodology described by Telci, Nam et al. (2009) but using different hydrodynamic data. The locations of the stations of this selected optimal monitoring system are shown in Figure 4.3. It should be noted that the proposed methodology is applicable to any real-time monitoring system independent of its design procedure.

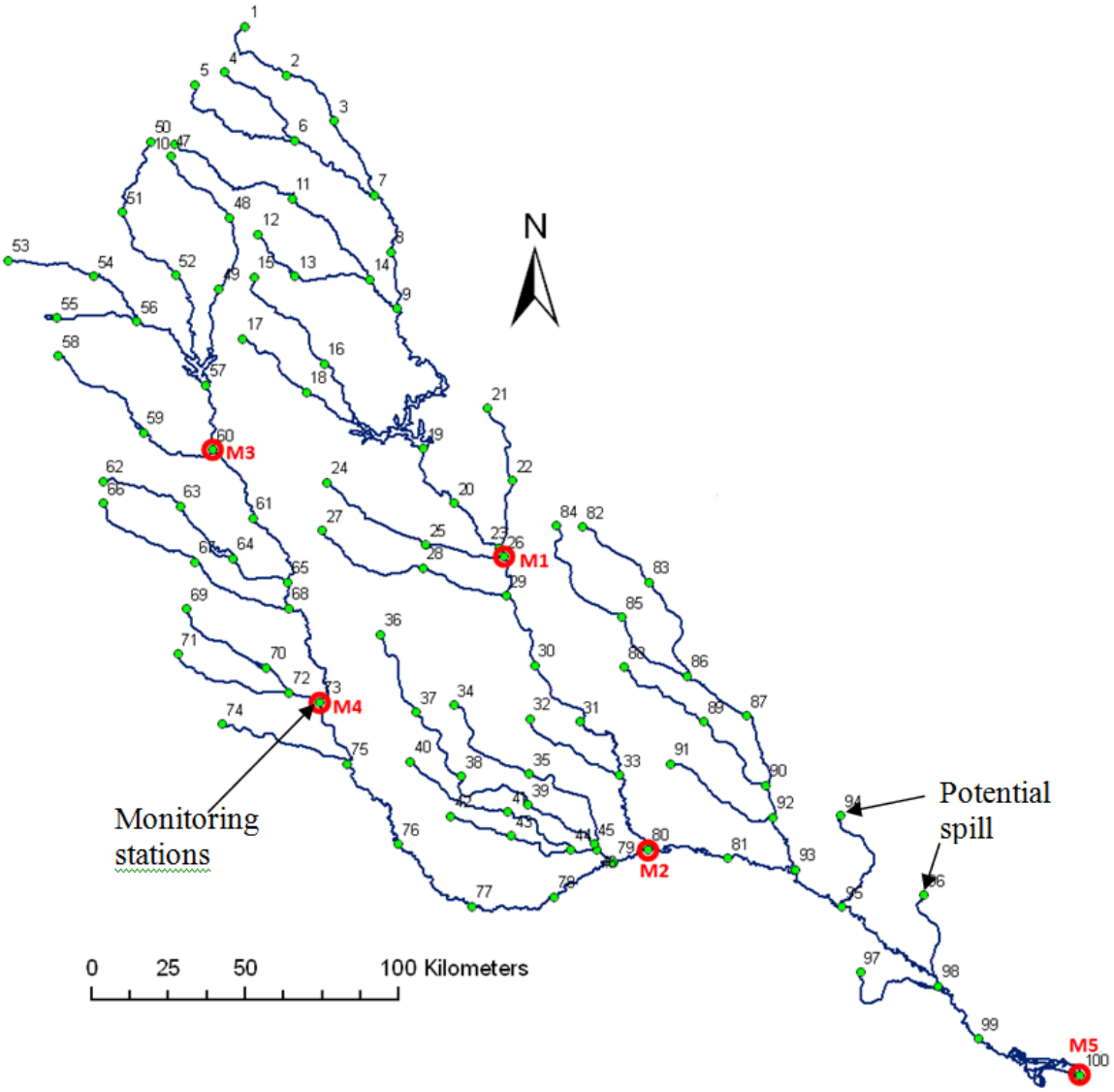


Figure 4.3 Selected locations of spills for the training process and designed water quality monitoring system.

4.3.3 Training of the Monitoring Stations

When a spill occurs in the river network, the contaminant plume travels downstream with the river flow. If the plume is detected by a monitoring station, then one can decide that

the spill location is in the upstream region of the first activated monitoring station. This region consists of the river reaches between the first activated station and the station located just upstream which was not activated. Therefore ASFS algorithm requires that each monitoring station is trained by a set of spill scenarios occurring in its own region. Training of a monitoring station is performed to estimate the probabilities of the candidate spill locations conditioned on the value of features of the observed breakthrough curve.

In this study, training process for the monitoring station M1 (Figure 4.3) is presented. Training set is a collection of large number of breakthrough curves observed at the monitoring station being trained. Each of these breakthrough curves have 6 features defined by Equations (4.3) to (4.5). For practical and consistency purposes, these features are collected in a feature vector X as described in Equation (4.13).

$$X = \left[\sigma^t \quad S^t \quad E^t \quad \sigma^V \quad S^V \quad E^V \right]^T \tag{4.13}$$

In this equation, X_1, X_2 and X_3 where $\Psi = t$ are temporal features of the breakthrough curves and similarly, X_4, X_5 and X_6 where $\Psi = V$ are volumetric features.

In this study, monitoring stations are trained by 1000 random spill scenarios occurring at each candidate spill location. Since, monitoring station M1 has 26 candidate spill locations in its region, 26000 breakthrough curves are used in its training process. Figure 4.4 presents how the features of these breakthrough curves are distributed relative to each

other. In this figure, the clusters of scenario spill locations are observable. The clusters of some nodes are distinctly separated from others. However, majority of the spill locations have cluster clouds located within the other clusters. This observation is consistent with the non-uniqueness of the problem where the breakthrough curves which have the same features may originate from different spill locations.

Once values of the features of all the scenarios in the training set are determined, the next step is to obtain conditional probability density function of a feature $X_i \in \{X_1, X_2, ..., X_6\}$ given that spill occurred at node $Y \in \Omega_k$, where Ω_k is the set of integer numbers used to label the candidate spill locations. For this purpose, a frequency analysis is applied to each feature of each of the spill locations. Figure 4.5 shows some examples of this analysis for arbitrarily chosen spill locations and features. Since the spill location is known and fixed in these frequency plots, the probability density function $f(X \mid Y)$ obtained from them will be conditioned on the spill location. This probability density function can be approximated using Kernel density estimation technique. Next, Bayes' theorem described in Equation (4.8) can be utilized to obtain posterior conditional probability of a candidate spill location given an observed feature, i.e. $P(Y = j \mid X_i = x_i)$. Then, this probability is used in ASFS algorithm to locate a real spill in the next step.

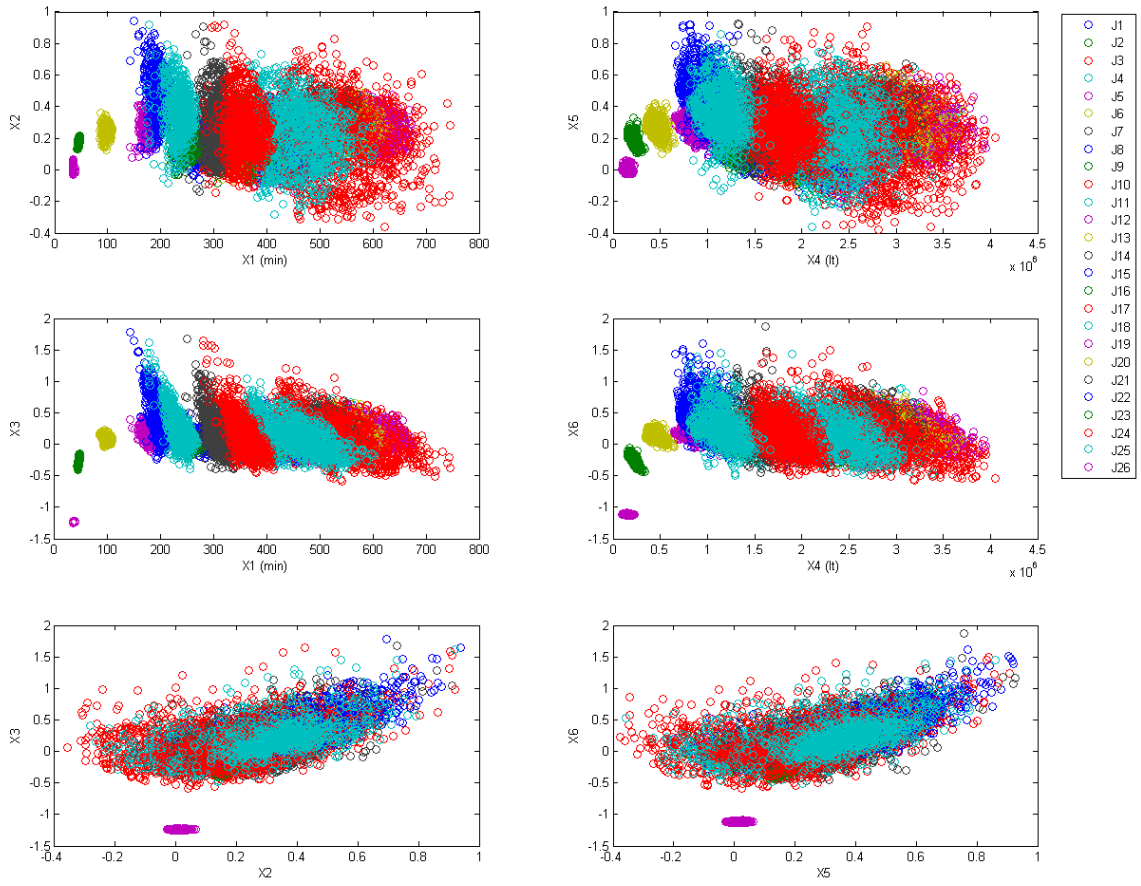


Figure 4.4 Feature plots for training of monitoring station M1.

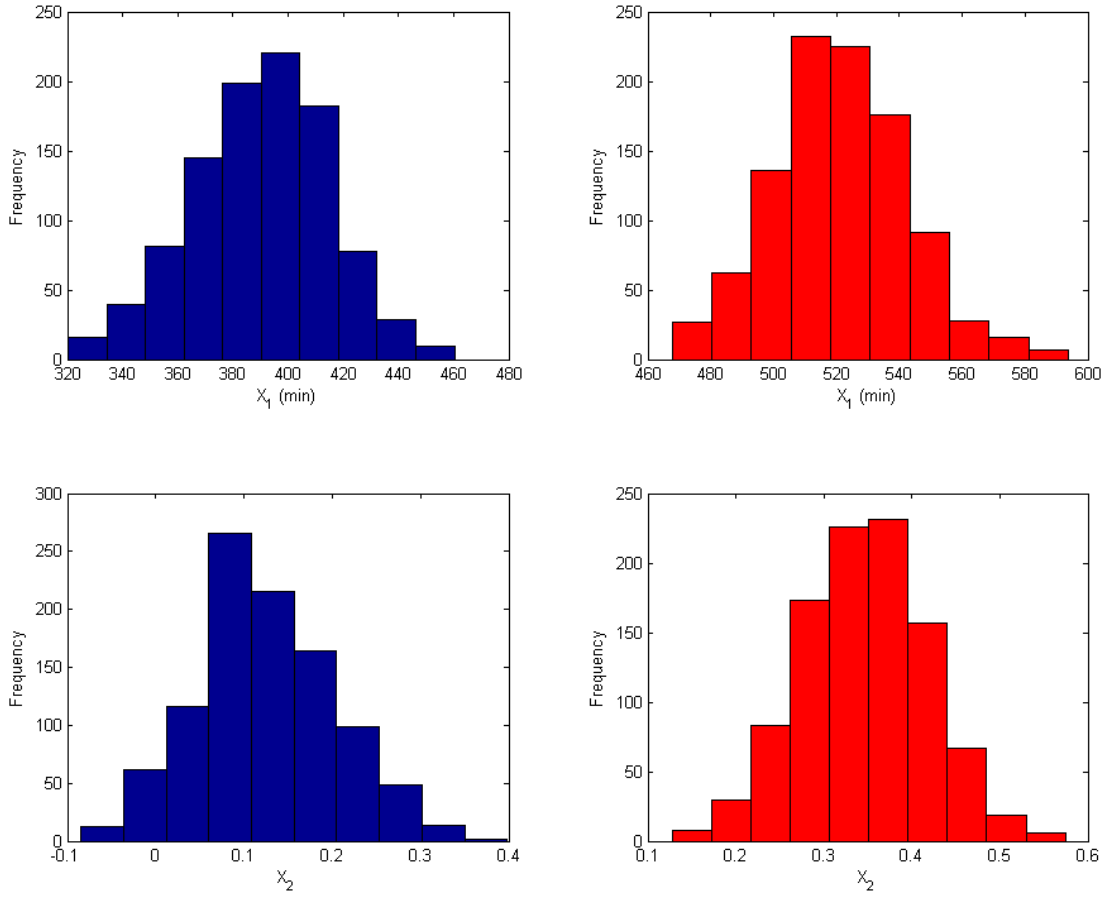


Figure 4.5. Examples for the frequency analysis.

4.3.4 Locating Spill Events

In the training of the monitoring stations, spill scenarios located at predefined candidate spill locations are used. However, a spill event may occur anywhere throughout the river system rather than these discrete candidate locations. These candidate locations indicate the upstream and downstream nodes of the river reaches where a spill event may occur. A spill also creates temporal and volumetric breakthrough curves at the downstream

monitoring stations. The features of the breakthrough curves of the spill event define a point in the six dimensional feature space described in Figure 4.4. The ASFS algorithm associates this point to one of the clusters provided by the training set. However, the proposed methodology provides a ranking of the clusters with respect to their associations to the spill point rather than a final selection of the spill location among the candidate locations due to non-unique nature of the solution. ASFS algorithm inherently offers this ranking approach while it is eliminating the classes from the candidate set.

Performance of the proposed methodology is tested by several spill event scenarios distributed over the regions of the monitoring stations M1 and M2 as shown in Figure 4.6. In this figure, arrows in south direction indicate the location of the spill events tested and these spills are located at arbitrary places along the river reaches. Each realization has a random spill time, spill mass and hydrodynamic configuration. Realizations R1 to R9 occur in the region of monitoring station M1 (Region 1) and realizations R10 to R19 represents the spills located in the region of the second monitoring station M2 (Region 2).

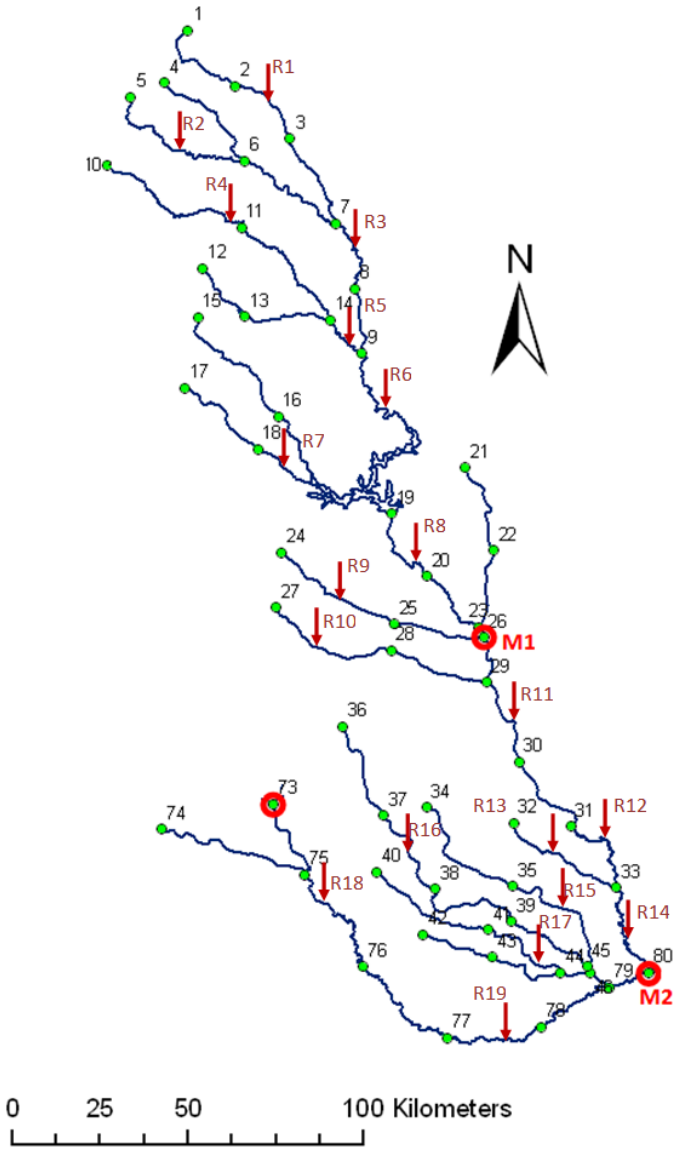


Figure 4.6 Spill realizations used to test the proposed methodology.

Table 4.2 demonstrates how ASFS algorithm eliminates candidate nodes for the spill of realization R7 shown in Figure 4.6. In this table, i^{th} column indicates the candidate nodes for the i^{th} step of the algorithm and first row indicates the features used for the elimination of candidate nodes in each step. As seen from this table, ASFS algorithm

selects the best feature for the elimination process automatically at each step of the algorithm. The shaded cells show the upstream (dark) and downstream (light) nodes of the river reach where spill of realization R7 is occurred. The algorithm starts with the set of all possible candidate nodes $\left(\left|\Omega_1\right| = 26\right)$. In this first step, the algorithm selects X_2 as the best feature which satisfies Equation (4.12) to eliminate the candidates from the class set. Since the nodes 20, 23 and 26 have the minimum (in this case zero) conditional probability given that feature X_2 has the value obtained from the breakthrough curve observed at monitoring station M1 due to the spill of realization R7, these three nodes are eliminated from the class set in the first step. In the following steps of the algorithm, since only one node has the minimum probability conditioned on the selected feature, only one node is eliminated from the class set. At the end, the algorithm selects one node as the spill location and in the case of realization R7, node 18 which is the upstream node of the reach where spill has occurred is selected. However, we propose to provide a ranking of the candidate nodes rather than a single selected node and this ranking process can be clearly observed in Table 4.2. According to ASFS algorithm, the spill has occurred most probably at node 18. The next possible place is node 16 and then nodes 24, 21, 3 and so on are other possible spill locations. Therefore, according to this algorithm, the investigation team will be sent first at region of node 18. This region can be defined as the upstream and downstream river reaches between node 18 and halfway to the other nearest nodes. If the investigation team cannot find the spill location in this first trial, they are going to investigate the regions of nodes 8, 21, 9, 5 and so on sequentially.

Table 4.2 Elimination process of ASFS algorithm for realization R7.

X_2	X_2	X_2	X_1	X_2	X_2	X_1	X_1	X_2	X_2	X_2	X_1	X_1	X_1	X_1	X_1	X_1	X_2	X_1	X_1	X_4	X_4	X_2
1	1	1	1	1	1	1	1	1	1	1	4	4	4	4	4	4	4	4	5	16	16	16	18
2	2	2	3	3	3	4	4	4	4	4	5	5	5	5	5	5	5	5	16	18	18	18	
3	3	3	4	4	4	5	5	5	5	5	6	6	6	6	6	6	6	16	18	21	24		
4	4	4	5	5	5	6	6	6	6	6	11	12	13	14	15	16	16	18	21	24			
5	5	5	6	6	6	7	7	7	8	11	12	13	14	15	16	17	18	21	24				
6	6	6	7	7	7	8	8	8	11	12	13	14	15	16	17	18	21	24					
7	7	7	8	8	8	10	11	11	12	13	14	15	16	17	18	21	24						
8	8	8	9	9	10	11	12	12	13	14	15	16	17	18	21	24							
9	9	9	10	10	11	12	13	13	14	15	16	17	18	21	24								
10	10	10	11	11	12	13	14	14	15	16	17	18	21	24									
11	11	11	12	12	13	14	15	15	16	17	18	21	24										
12	12	12	13	13	14	15	16	16	17	18	21	24											
13	13	13	14	14	15	16	17	17	18	21	24												
14	14	14	15	15	16	17	18	18	21	24													
15	15	15	16	16	17	18	21	21	24														
16	16	16	17	17	18	21	24	24															
17	17	17	18	18	21	24	25																
18	18	18	21	21	24	25																	
19	19	21	22	24	25																		
20	21	22	24	25																			
21	22	24	25																				
22	24	25																					
23	25																						
24																							
25																							
26																							

The spill realization R7 is selected first to discuss the elimination process since it provides a very good example of feature selection property of the algorithm by switching between different features. It should be noted that this is not always the case and the algorithm may select the same feature for the majority of the elimination process. Spill realization R3 can be a good example of this case. Table 4.3 shows how ASFS algorithm eliminates candidate nodes for the spill realization R3 given in Figure 4.6. In this example, ASFS algorithm has selected nodes 7 and 8, which are the upstream and downstream nodes of the river reach where the spill has occurred, as the final two possible spill locations and at the end, node 7 was selected as the final prediction. In order to visualize how candidate locations change with ASFS iterations for the realization

spill R3, Figure 4.7 and Figure 4.8 are plotted showing the candidate spill locations at 10^{th} and 20^{th} ASPFS iteration, respectively. In these figures shaded circles indicate candidate spill locations at a given ASFS iteration. In this case, we can rank the first five candidate nodes as 7, 8, 21, 9 and 3.

Table 4.3 Elimination process of ASFS algorithm for realization R3.

X1	*X1*	*X1*	*X1*	*X1*	*X1*	*X1*	*X1*	*X1*	*X1*	*X1*	*X1*	*X1*	*X1*	*X1*	*X1*	*X1*	*X1*	*X1*	*X1*	*X1*	*X4*
1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	3	3	3	7	7	7	7
2	2	2	2	2	2	2	2	2	2	2	2	2	3	3	3	7	7	7	8	8	8	
3	3	3	3	3	3	3	3	3	3	3	3	3	7	7	7	8	8	8	9	21		
4	4	4	4	4	4	4	4	4	4	6	6	7	8	8	8	9	9	9	21			
5	5	5	5	5	5	5	5	5	6	7	7	8	9	9	9	21	21	21				
6	6	6	6	6	6	6	6	6	7	8	8	9	17	18	21	24	24					
7	7	7	7	7	7	7	7	7	8	9	9	17	18	21	24	25						
8	8	8	8	8	8	8	8	8	9	16	17	18	21	24	25							
9	9	9	9	9	9	9	9	9	16	17	18	21	24	25								
10	10	10	12	12	13	15	15	16	17	18	21	24	25									
11	11	12	13	13	15	16	16	17	18	21	24	25										
12	12	13	14	15	16	17	17	18	21	24	25											
13	13	14	15	16	17	18	18	21	24	25												
14	14	15	16	17	18	21	21	24	25													
15	15	16	17	18	21	22	24	25														
16	16	17	18	21	22	24	25															
17	17	18	21	22	24	25																
18	18	21	22	24	25																	
19	21	22	24	25																		
20	22	24	25																			
21	24	25																				
22	25																					
23																						
24																						
25																						
26																						

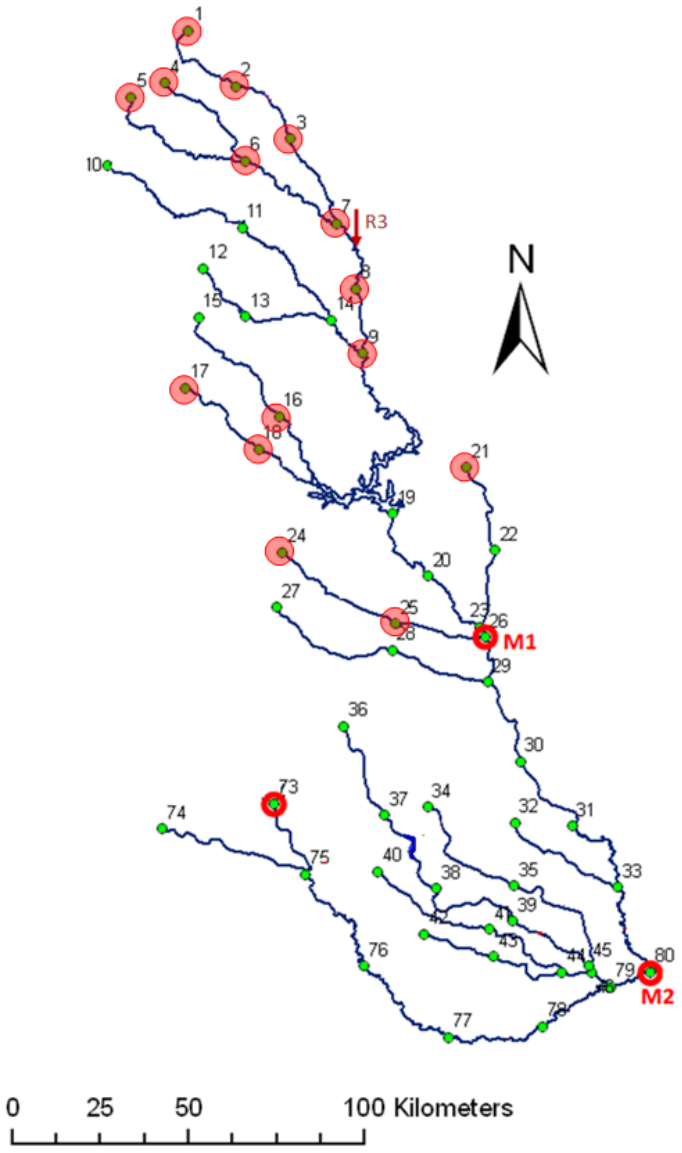


Figure 4.7 Candidate spill locations for realization R3 at 10th ASFS iteration.

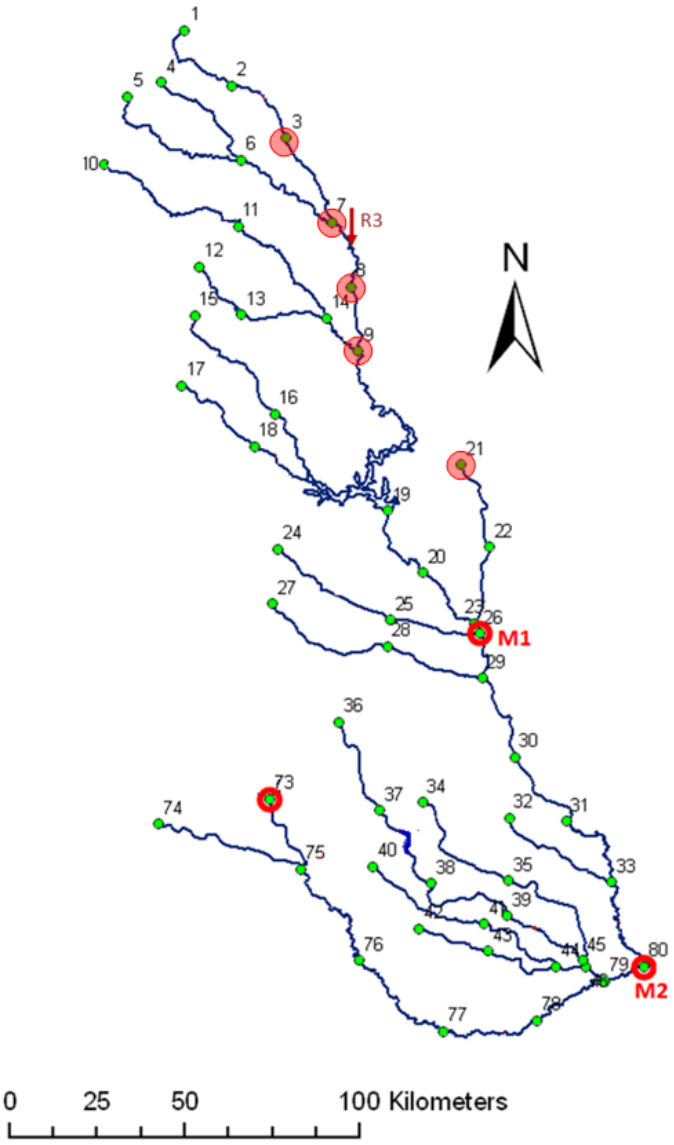


Figure 4.8 Candidate spill locations for realization R3 at 20th ASFS iteration.

The overall performance of the proposed methodology for the realizations in Figure 4.6 is presented in Figure 4.9. In this figure, horizontal axis represents the number of elements in the class set at the k^{th} step of ASFS algorithm $\left(\left|\Omega_k\right|\right)$ and the vertical axis, Reliability, represents the ratio of the number of realizations where correct spill location is in the set

Ω_k to the total number of realizations in percent. In this study, correct spill location is represented by either the upstream or the downstream node of the reach where the spill has occurred. Therefore, it is assumed that if the upstream or downstream node of the reach where the spill has occurred is kept in the set Ω_k, then Ω_k contains the correct spill location. The square and diamond markers in Figure 4.9 represent the reliability values calculated using the realizations only in Regions 1 and 2 respectively. The circles indicate the reliability values estimated by including all realizations in both regions. This figure shows that when ASFS algorithm is allowed to eliminate the candidate class set until one element is left $\left(\Omega_k = 1\right)$, then this selected node is a correct spill location for approximately 60% of the time. When 3 elements are left in the class set, the reliability is approximately 90%. If the set size is further increased to 5, the correct spill locations for all the realizations in Region 1 are included the class set. Figure 4.9 also indicates a limitation for the proposed methodology since the reliability value for Region 2 can become 100% only when the candidate class set size is as high as 16. This is due to realization R13 which occurs between nodes 32 and 33 (Figure 4.6). These two nodes are eliminated from the class set at different steps of ASFS algorithm before the set size is reduced to 15. The reason behind this problem is the fact that node 32 is a most upstream node where the base flow rate is much smaller than that at node 33 which discharges the water coming from a large portion of the watershed. Then, a realization which occurs between these nodes does not have the feature values which result in high probabilities using Equation (4.8). However, such reaches are not very common in river systems and 90% reliability at a class set size of 3 indicates that the proposed

methodology is a reliable way of ranking the candidate nodes in terms of their association with the real spill.

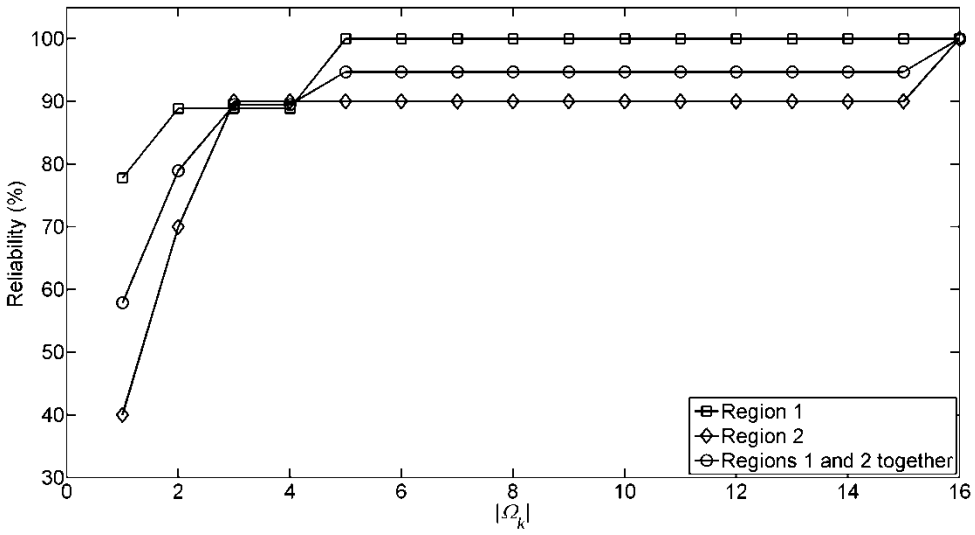


Figure 4.9 Overall performance of the proposed methodology.

4.4 Conclusions

A methodology is proposed to identify the location of a contaminant spill in a river system utilizing the measurements obtained from a water quality monitoring system. The first step of the method is training the monitoring stations with a large number of spill scenarios. Then the proposed method provides a rapid way of ranking the candidate locations using a classification routine. Although the training process requires a significant amount of simulation time, once it is done for a monitoring system, the method works without additional simulations. The classification routine requires only the features of the breakthrough curves measured at the monitoring stations to rank the

candidate spill locations. The results show that methodology ranks the correct spill location in the first 3 candidates for the 90% of the test cases. This result reveals that the proposed methodology can be used effectively for the preliminary planning of investigation studies for the spill locations.

The analysis presented is based on 6 features obtained from statistical analysis of temporal and volumetric breakthrough curves. The proposed methodology may also be utilized without any change for completely different features other than these statistical parameters. However, the performance of the method with the new feature set may be different than that presented in this thesis. The technique presented in this study is developed for instantaneous spill and conservative contaminant cases only. It may also be extended to include non-instantaneous and non-conservative spills.

CHAPTER 5

RENEWABLE ENERGY PRODUCTION FROM WATER DISTRIBUTION SYSTEMS

5.1 Introduction

Water distribution systems are designed to satisfy the consumer demands at the outlet nodes. To achieve this goal, adequate pressures need to be maintained throughout the network. As the complexity of a water distribution network increases, maintaining target pressures becomes difficult. This yields excess pressures at several locations of the network. The conventional solution to this problem is to install pressure reducing valves which adjust the local head loss to lower the downstream pressure. This approach dissipates significant amount of energy that can be recovered and used for the benefit of the community. This energy recovery is possible by the utilization of micro hydroelectric turbines as an alternative means of pressure reduction. In this study, an optimization approach for the design of energy recovery systems in water distribution networks is proposed. This methodology is based on finding the best locations for micro hydroelectric plants in the network to recover the excess energy. Due to the unsteady nature of flow in water distribution networks, the proposed methodology also determines an optimum operation schedule for the micro turbines. The objective of the optimization algorithm is determined as the net annual energy gained by the energy recovery system. As a test case, the proposed methodology is applied to the water distribution system

serving the Dover Township area in New Jersey, which can be considered to be a typical small town in USA. The results show that this study is an effective tool for the assessment of renewable energy potential in water distribution systems.

5.2 Methodology

5.2.1 Hydrodynamic Simulation of Water Distribution System

For the unsteady hydrodynamic analysis of the water distribution network EPANET 2.0 is used. EPANET is a simulation tool for extended period modeling of hydraulic and water quality behavior of pressurized pipe networks (Rossman 2000). Elements of a network defined in EPANET are pipes, nodes (pipe junctions), pumps, valves and storage tanks or reservoirs. EPANET estimates the flow of water in each pipe, the pressure at each node and the height of water in each tank throughout the network at each time step during a simulation period. EPANET's capability of considering multiple demand categories at nodes, each with its own time variation makes it a strong simulation tool for unsteady hydraulic behavior in water distribution systems. Although EPANET does not have a predefined turbine object as a network element, turbines can be simulated using a general purpose valve (GPV) defined in EPANET. In order to simulate a turbine as a GPV, the user needs to supply a special flow-head loss relationship.

5.2.2 Optimization Model

In this study, the primary objective of the energy recovery system is defined as harvesting any available energy in the water distribution network at any given time while guaranteeing consumer demands without violating pressure constraints set by the management throughout the service period. The first step of this design procedure is deciding candidate location(s) and types of micro turbines that will be utilized in the energy recovery system. In addition to this decision, the highly unsteady nature of the water distribution system dictates an operation schedule for each micro turbine proposed to be installed. A proper energy recovery system design is possible with an operation schedule which enables the turbine to generate as much energy as possible without violating pressure constraints while satisfying consumer demands. For this purpose, the micro-turbine-bypass valve combination demonstrated in Figure 5.1 is proposed. This design enables us to adjust the flow passing through the turbine at any given time by adjusting the bypass valve opening. A completely closed bypass valve would force all the flow to pass through the turbine while a completely open bypass valve would let all the flow pass through the bypass pipe. Therefore, the operational schedule of a turbine can be determined by the degree of opening of the bypass valve.

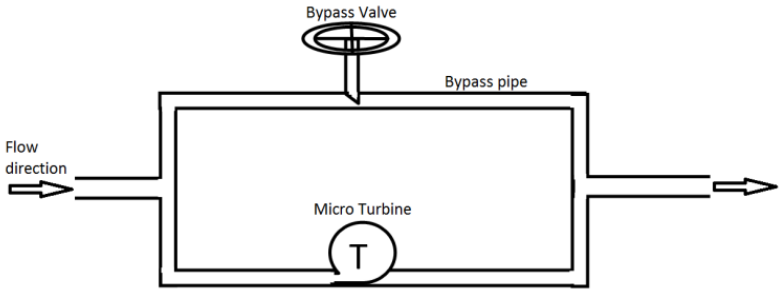


Figure 5.1 Micro turbine-bypass valve combination

If the energy recovery system has N micro turbines located at N candidate locations, then we can represent these turbines with a vector $\Psi = [\psi_1, \psi_2, ..., \psi_N]^T$, where ψ_i represents candidate turbine proposed to be installed at the i^{th} candidate location and let the vector $\Lambda = [\lambda_1, \lambda_2, ..., \lambda_N]$ stands for the pipes where turbines, Ψ are proposed to be installed. The operation schedule of each of these turbines can be represented by another vector $\Omega = [\omega_1, \omega_2, ..., \omega_N]^T$, where ω_i stands for the operational schedule of the turbine located at the i^{th} candidate location and therefore, ω_i is a vector of a variable δ_k which takes a real value between 0 and 1 representing the degree of opening of the bypass valve of the turbine, i.e. $\omega_i = [\delta_1, \delta_2, ..., \delta_{N_{ST}}]^T$ where N_{ST} is the number of time steps in the analysis period and $k = \{1, 2, 3, \ldots, N_{ST}\}$.

For a given set of operation schedule of the candidate turbines, the system may not satisfy pressure constraint. In order to describe this mathematically, let $P_k = [p_{1,k}, p_{2,k}, ..., p_{N_J,k}]$

denote the vector of pressures $p_{j,k}$ at the j^{th} junction of the water distribution network at the k^{th} time step of the analysis period where N_J is the total number of junctions in the network. Then, one can determine the first pressure constraint failure time, $t_f(\Psi,\Omega)$, such that $\min(P_k) < P_{\min}$, where $P_{\min}$ is the minimum pressure limit set by the management for safe operation. According to this definition, when $t_f(\Psi,\Omega)$ is equal to the analysis period, one can say that for the given configuration of turbines (Ψ) and respective operation schedules (Ω), pressure constraint is satisfied at all locations, at all times. At this point we can define a dimensionless failure time, $\bar{T}(\Psi,\Omega)$ as the ratio of the pressure constraint violation time $t_f(\Psi,\Omega)$ to the total time of the analysis period, T_A, as in Equation (5.1). It can be seen that $\bar{T}(\Psi,\Omega)$ can have values between 0 and 1. While $\bar{T}(\Psi,\Omega)=0$ indicates that pressure constraint is violated at the very beginning of the analysis, $\bar{T}(\Psi,\Omega)=1$ implies that pressure constrained is satisfied throughout the analysis period.

$$\bar{T}(\Psi,\Omega) = \frac{t_f(\Psi,\Omega)}{T_A} \tag{5.1}$$

The energy produced by the energy recovery system until the first pressure constraint failure time can be denoted by the scalar $E(\Psi,\Omega)$. One can also calculate the total energy of the flow passing through the pipes Λ without any turbines installed as shown in Equation (5.2).

$$E_M = \sum_{k=1}^{N_{ST}} \sum_{i=1}^{N} \left[\gamma Q(\lambda_i, k) H(\lambda_i, k) \Delta t \right] \tag{5.2}$$

where γ is the specific weight of water, Δt is the time step, and $Q(\lambda_i, k)$ and $H(\lambda_i, k)$ are the flow rate and the average total head in the pipe λ_i at the k^{th} time step of the analysis, respectively. A non-dimensional energy measure, $\bar{E}(\Psi, \Omega)$, can be defined as the ratio of the energy produced, $E(\Psi, \Omega)$ to the total energy of the flow, E_M as described in Equation (5.3).

$$\bar{E}(\Psi, \Omega) = \frac{E(\Psi, \Omega)}{E_M} \tag{5.3}$$

Since the main aim of the energy recovery system is defined as achieving as high energy production by the micro turbines as possible without violating the pressure constraint, the design of this system can be formulated as an optimization problem that can be mathematically expressed as

$$\begin{aligned} f &= \underset{\Omega}{\text{maximize}} \left\{ \bar{E}(\Psi, \Omega) \right\} \\ s.t.\ &\bar{T}(\Psi, \Omega) = 1 \\ &\Psi = \Psi_o \end{aligned} \tag{5.3}$$

where Ψ_o is vector of predefined micro turbines used in the analysis.

Because any decision on the micro turbine operation at a given time step k affects the flow conditions in future time steps, this is a very complex and nonlinear optimization problem. In this study, Genetic Algorithm which imitates the evolutionary natural selection is utilized to find the best operation schedules (Ω) for the candidate micro turbines. In this case a configuration of operation schedules of the turbines in the system is represented by an individual and a population of a number of individuals needs to be simulated. Each individual will have an objective function fitness value represented by the total energy obtained from the energy recovery system configuration. In a genetic algorithm a common way of dealing with candidate solutions that violate the constraints is to generate potential solutions without considering the constraints and then penalizing them by decreasing the goodness of the fitness function. In this respect, the fitness function of the genetic algorithm, f_{GA} is determined as in Equation (5.4).

$$f_{GA} = A\left(\bar{T}\left(\Psi,\Omega\right)-1\right)+\bar{E}\left(\Psi,\Omega\right) \tag{5.4}$$

where A is a large penalty coefficient which is selected as 1000 in this study. According to Equation (5.4), the fitness value of an individual which violates the pressure constraint (i.e. $\bar{T}\left(\Psi,\Omega\right)<1$) will have a lower fitness value compared to an individual which satisfies pressure constraint at all times (i.e. $\bar{T}\left(\Psi,\Omega\right)=1$). The GAs will improve the initial population by reproducing new generations applying natural selection and population genetics such as, selection, cross over and mutation. In every generation, individuals are rated according to their fitness values obtained from the simulations. The

optimal configuration which has the highest energy production is reached in a finite number of generations.

5.2.3 Smart Seeding of the Genetic Algorithm

Genetic algorithms (GAs) are stochastic in nature and every step in a GA from generation of initial population to mutation has a random characteristic. Although this property helps GA get out of local optima, it creates individuals with irregular chromosomes which make it difficult to reach global optimum solution for some problems. Many researchers suggest that the solution to this problem lies at the initial population of the GA (Ponterosso and Fox 1999; Karci 2004; Maaranen, Miettinen et al. 2004; Maaranen, Miettinen et al. 2007; Saavedra-Moreno, Salcedo-Sanz et al. 2011). They suggest that good feasible solution(s) can be seeded in the initial population of the elitist GA where a number of individuals which has the highest fitness value are always selected for the next generation. By this way the GA is initially supported by the information of a strong individual and it is forced to find fitter individuals. This process helps gain significant amount of time and it considerably improves the final best solution found by the GA using a reasonable population size.

The most important step of this seeding process is to find a good feasible solution to the problem. Thus, the individual operational schedule of the turbine of this seeding process needs to satisfy pressure constraints all the time. In that sense, schedule which has a

completely open bypass valve throughout the simulation is a feasible solution for this problem since the micro turbine does not affect the system. However, this solution is not a good feasible solution for seeding purpose because the fitness value that is the energy generated by the micro turbine is zero for this individual. Therefore the seeding individual should have a fitness value as high as we can possibly find using a deterministic approach. Because the high energy generation increases the quality of our seed, we prefer longer turbine operation. At the same time, we have to satisfy pressure constraints in order to save the feasibility of the seed. Therefore we have to open the bypass valve whenever it is necessary. Sometimes we can eliminate a pressure violation in the system by simply opening the bypass valve at the pressure violation time. However this not always the case since the system may require a longer period of bypass valve opening in order to satisfy the pressure constraints. Thus an iterative procedure is proposed to find a good individual to seed the GA as described below.

i. Assume full turbine operation
ii. Run the simulation
iii. Determine the first pressure failure time.
iv. Open Bypass valve at this pressure failure time
    - If bypass valve is already open at this time, open the bypass valve one time step before.
v. Update the operation schedule of the turbine.
vi. Repeat steps 2 to 5 until a feasible solution which does not violate pressure constraint is reached.

This procedure is based on delaying the first pressure violation time by opening the bypass valve at or before that time. The final output of this seed generation algorithm is an operation schedule where the bypass valve is either completely open or completely closed. Then this individual operation schedule is seeded into the initial population of the GA and it will be further improved enabling us to reach the best solution that can be obtained in a reasonable amount of computational expense.

5.3 Applications

5.3.1 Study Area

In this study, the proposed methodology is applied to the water distribution system serving the Dover Township area in New Jersey (Figure 5.2), which can be considered to be a typical small town in USA. This water distribution system is selected to test the proposed approach since it has been extensively studied and well documented (Maslia, Sautner et al. 2000; Maslia, Sautner et al. 2001; Aral, Guan et al. 2004; Aral, Guan et al. 2004). This water supply network is composed of 16048 pipes connecting 14945 junctions. 20 underground wells located at 8 inlet points serve as the main water supply of the system (Figure 5.2). Therefore, the energy required for the flow in the system is provided by the pumps located at these wells.

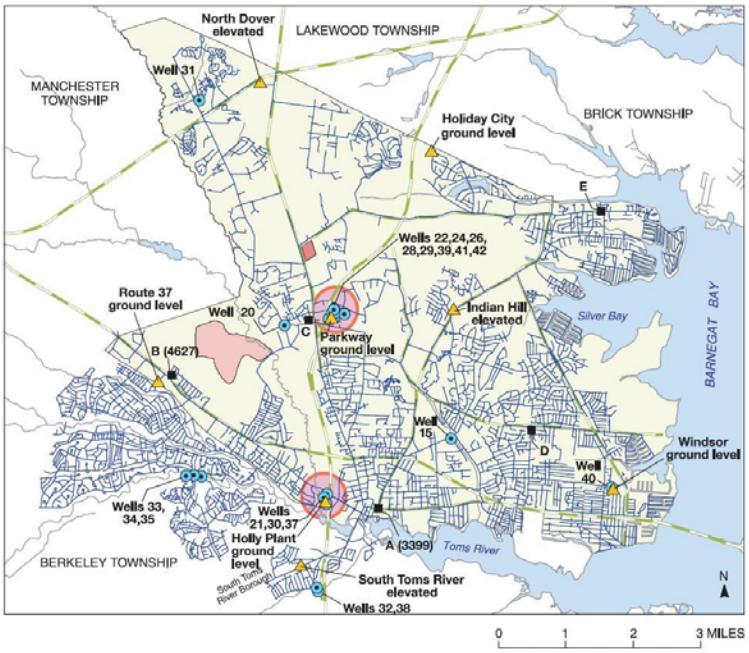


Figure 5.2 Dover Township water distribution system, Toms River, New Jersey.

Previous studies have reported the average daily demand pattern for each month of the year and in this study we use these demand patterns consecutively to represent the yearly demand pattern as shown in Figure 5.3. Since one month is represented by a 24-hour time period, the simulation time required to represent one year is 288 hours. Similarly, the operational schedule to be determined for a micro turbine is composed of 288 hours (i.e. $N_{ST} = 288$). Every node in the water distribution system has a different base demand which will yield a different time series of demand flow rate when multiplied with the demand pattern in Figure 5.3.

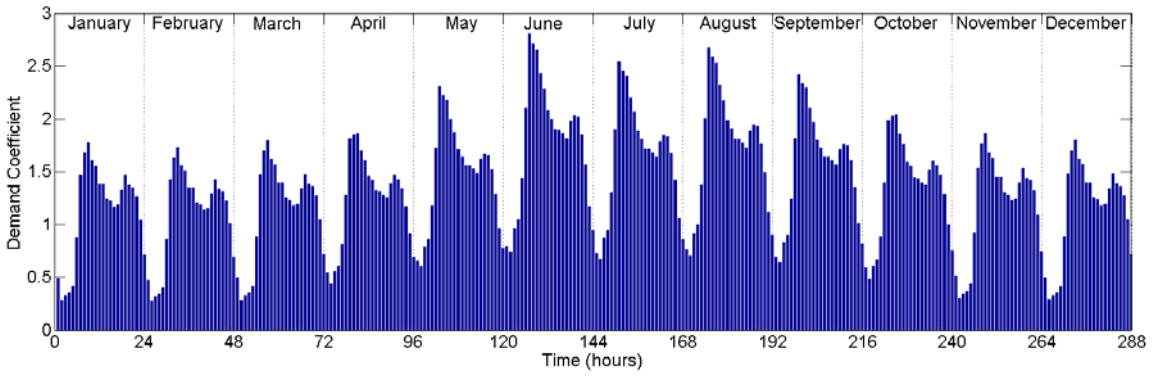


Figure 5.3 Hourly demand pattern representing one year.

Since the main goal of this study is to harvest the available excess energy in the system, it is important to assess the excess energy potential of the water distribution network. If the excess energy at a node is defined as the energy due to the pressure above a minimum limit, $P_{\min}$, then one can make a rough estimate of total excess energy dissipated at a demand node i using Equation (5.5).

$$EE_i = \int \gamma q \Delta h dt = \sum_{k=1}^{N_{ST}} \gamma q_{i,k} \Delta h_{i,k} \Delta t \tag{5.5}$$

where $q_{i,k}$ and $\Delta h_{i,k}$ are the demand and pressure head above $P_{\min}/\gamma$ at node i at k^{th} time step respectively. The total excess energy dissipated in the system can be calculated as $TEE = \sum_{i=1}^{N_N} EE_i$ where N_N is the number of nodes. The total excess energy dissipated in the Dover Township water distribution system is estimated as 1.4 GWh/y. We can also calculate excess energy input to the system at the 8 inlet nodes utilizing Equation (5.5)

using the discharge passing through the inlet pipe and average pressure head above $P_{\min}/\gamma$ along the inlet pipe for $q_{i,k}$ and $\Delta h_{i,k}$ respectively. The total excess energy input at the inlet pipes of the Dover Township water distribution system is estimated as 1.6 GWh/y. Figure 5.4 shows the distribution of this energy at every inlet pipe. In this figure percentages indicate the excess energy contribution of the corresponding inlet location divided by the total excess energy input to the system. The numbers below the percentages indicate the value of annual excess energy input in GWh/y at the corresponding inlet location by setting the pressure limit as $P_{\min} = 20$ psi. This figure indicates that the two locations of highest excess energy supply are locations 3 and 4, and these two inlet points are investigated in this study as candidate micro turbine locations. The locations other than these inlet points in Figure 5.4 do not provide better energy recovery sites. The reason behind this is the fact that the micro turbine introduces a significant head loss at the pipe it is installed behaving as an obstacle to the flow and the flow prefers other routes with lower head losses in the network to reach its final destination distributing itself so that the flow at the turbine is very low. This result in a very low energy generation at the micro turbine since the energy produced at the turbine is an increasing function of flow rate. At the inlet locations however, since the flow does not have alternative routes, the flow through the turbine does not decrease as much as an ordinary location in the network resulting in a higher energy recovery.

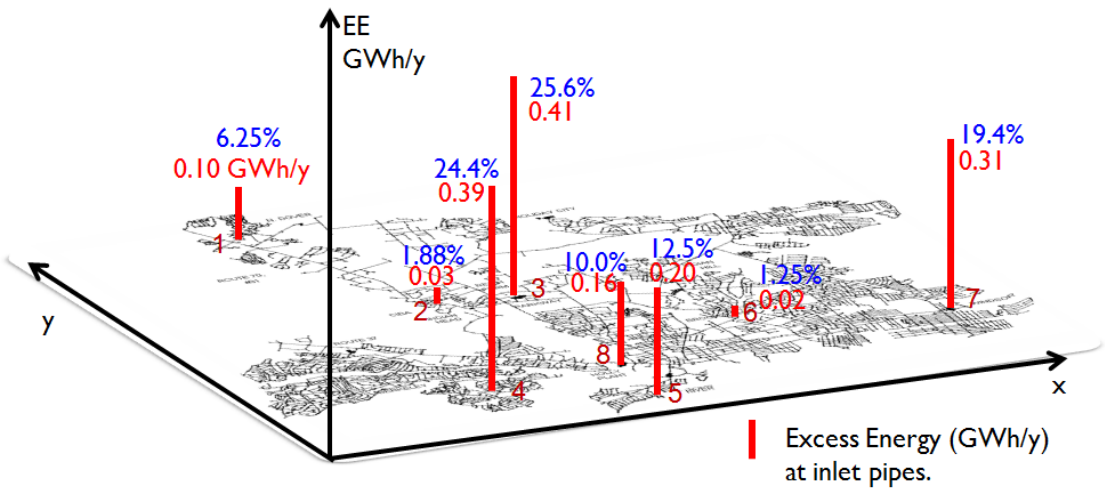


Figure 5.4 Excess energy input distribution.

5.3.2 Micro Turbines Used

In this study, the two micro turbines which are pumps used as turbines (PATs) listed in Table 5.1 (Giugni, Fontana et al. 2009) are used. In this table, Q_{tb}, H_{tb}, η_{tb} and P_{tb} stands for the flow rate, turbine head, efficiency and power generated at the best efficiency point in turbine mode, respectively. Derakhshan and Nourbakhsh (2008) have developed Equation (5.6) for turbine head and Equation (5.7) for turbine power estimations. These equations are utilized to plot Figure 5.5 and Figure 5.6 to show how the turbine head and turbine power changes with the flow rate for the two micro turbines. According to these figures, although NC 100-200 produces higher power for a given flow rate, it causes a higher head loss in the flow than NC 150-200. This can be interpreted as a higher possibility of pressure violation. The head curves in Figure 5.5 are used in the hydrodynamic simulation as input to EPANET providing head loss vs. flow rate

information for the general purpose valve representing the micro turbine. The power curves in Figure 5.6 are used to calculate the energy produced by the micro turbine at a given time step.

Table 5.1 Characteristics of micro turbines used.

PAT	$Q_{tb}\left(m^3/s\right)$	$H_{tb}\left(m\right)$	$\eta_{tb}\left(\%\right)$	$P_{tb}\left(kW\right)$
NC 100-200	0.05	19.81	79	7.82
NC 150-200	0.13	18.22	80	18.27

$$\frac{H_t}{H_{tb}} = 1.0283\left(\frac{Q_t}{Q_{tb}}\right)^2 - 0.5468\left(\frac{Q_t}{Q_{tb}}\right) + 0.5314 \tag{5.6}$$

$$\frac{P_t}{P_{tb}} = -0.3092\left(\frac{Q_t}{Q_{tb}}\right)^3 + 2.1472\left(\frac{Q_t}{Q_{tb}}\right)^2 - 0.8865\left(\frac{Q_t}{Q_{tb}}\right) + 0.0452 \tag{5.7}$$

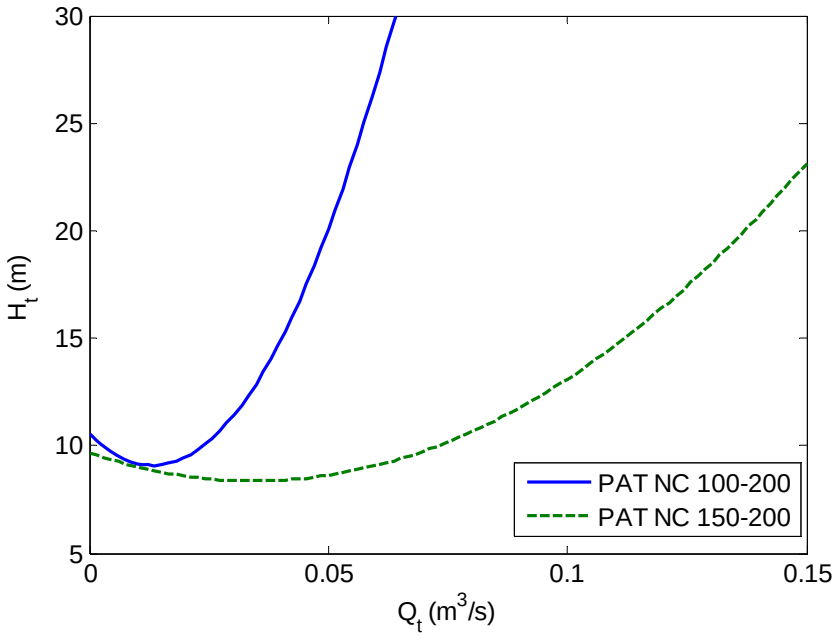


Figure 5.5 Turbine head vs. flow rate curves of the micro turbines.

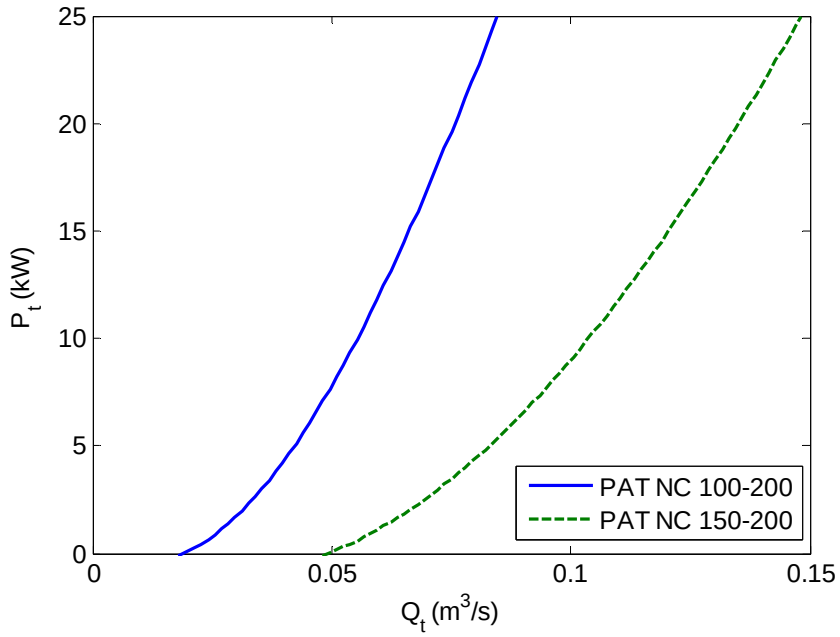


Figure 5.6 Power vs. flow rate curves of the micro turbines.

5.4 Gravity Driven Water Distribution System

In its original form, the Dover Township water distribution system is a pump driven network where the water is supplied by 20 pumps located at 8 pumping stations (Figure 5.4). For comparison purposes, a hypothetical gravity driven network is modeled by removing all the pumping stations in Dover Township water distribution system and connecting 3 constant head reservoirs at locations 3, 4 and 7 with total heads of 320.4 ft, 320.4 ft and 306 ft, respectively.

The same optimization algorithm is applied to this hypothetical gravity driven water distribution system and the results are reported after the results of the original pump driven network in the coming sections.

5.5 Results

In this section, we first present the results of a preliminary analysis performed to decide the population size of the GA and demonstrate the effect of the smart seeding process on the final optimal result of the GA. Then, the optimization algorithm is applied to several energy recovery system configurations in pump driven network and energy budget of each is discussed and their economic and environmental impacts are reported. Next, the same analysis is performed for the gravity driven network.

5.5.1 Smart Seed vs. Non-Seeded GA Solutions and the Population Size

In order to understand the decision process for the population size of the GA, we need to recall that the decision variable of our optimization algorithm is the hourly operation schedule of the micro turbine. Since our model simulates each month by a representative 24 hour period, the operational schedule of the micro turbine is composed of 288 hours. Therefore, for a single turbine problem, each individual in the GA has a chromosome size of 288 bytes. This large chromosome size may require a large number of individuals, i.e. the population size. In order to observe the effect of population size on the final output of the GA, several population sizes with number of individuals 80, 160, 400 and 1000 are tested for the case NC 150-200 installation at location 3. Moreover, a smart seed for the GA is found and its fitness value is compared with the result of non-seeded GA results. In these test runs, the minimum pressure that needs to be satisfied is set as 20 psi.

Figures 5.7, 5.8, 5.9 and 5.10 demonstrates the best operational schedules of the bypass valve of a single NC 150-200 installed at location 3 found by the GA with a population size of 80, 160, 400 and 1000 respectively. In these figures, a bypass valve opening of 1 indicates a fully open bypass valve. A partially open bypass valve is indicated by a bypass valve opening value between 0 and 1, and a completely closed bypass valve is represented by the value 0. Figure 5.11 shows the operational schedule of the bypass valve found as the smart seed for the GA as explained in Section 5.2.3.

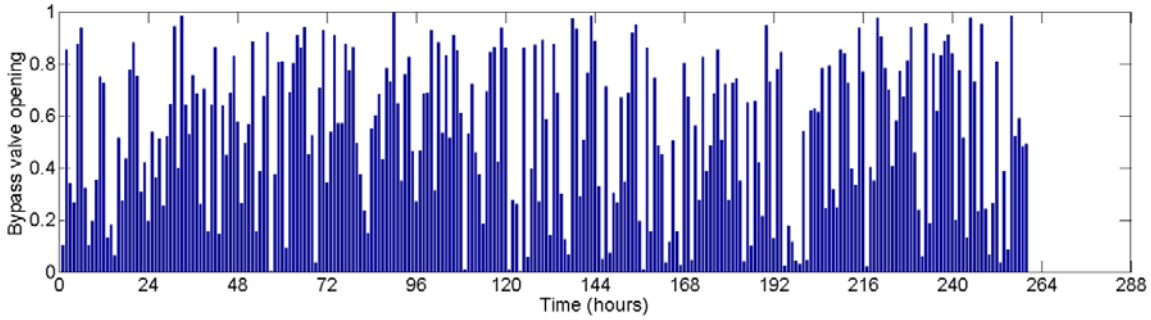


Figure 5.7 Best operational schedule found by the GA with a population size of 80 for the bypass valve of single NC 150-200 installed at location 3.

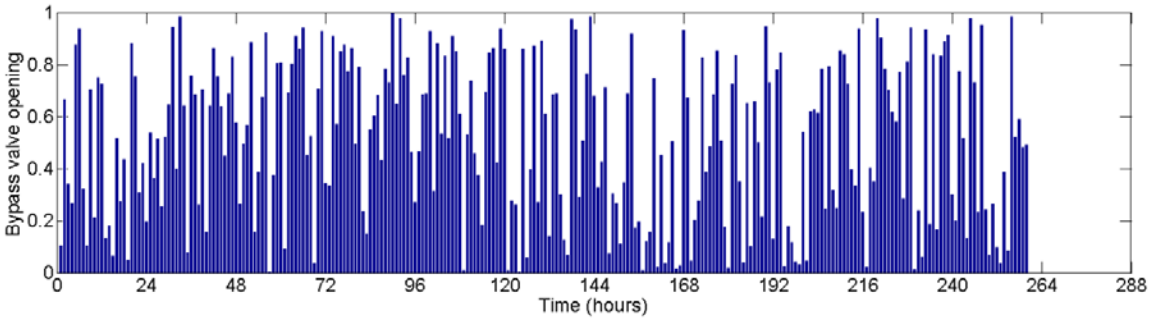


Figure 5.8 Best operational schedule found by the GA with a population size of 160 for the bypass valve of single NC 150-200 installed at location 3.

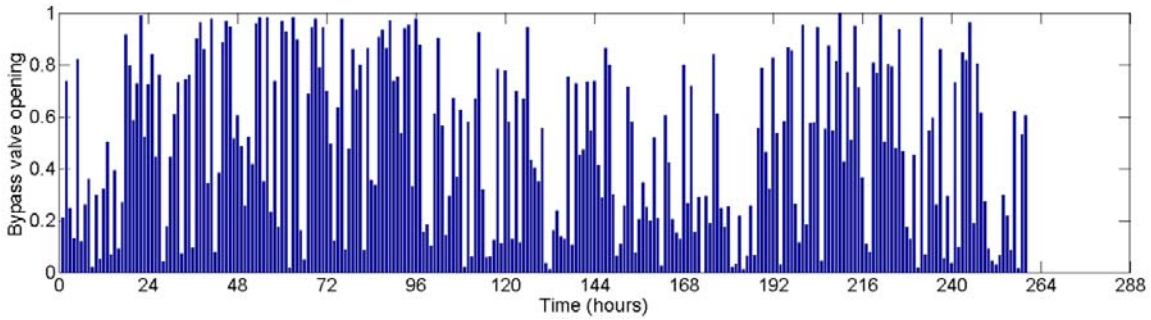


Figure 5.9 Best operational schedule found by the GA with a population size of 400 for the bypass valve of single NC 150-200 installed at location 3.

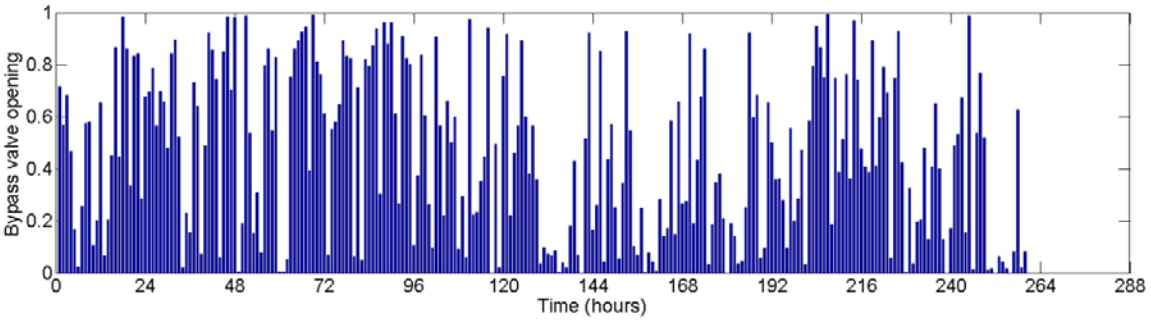


Figure 5.10 Best operational schedule found by the GA with a population size of 1000 for the bypass valve of single NC 150-200 installed at location 3.

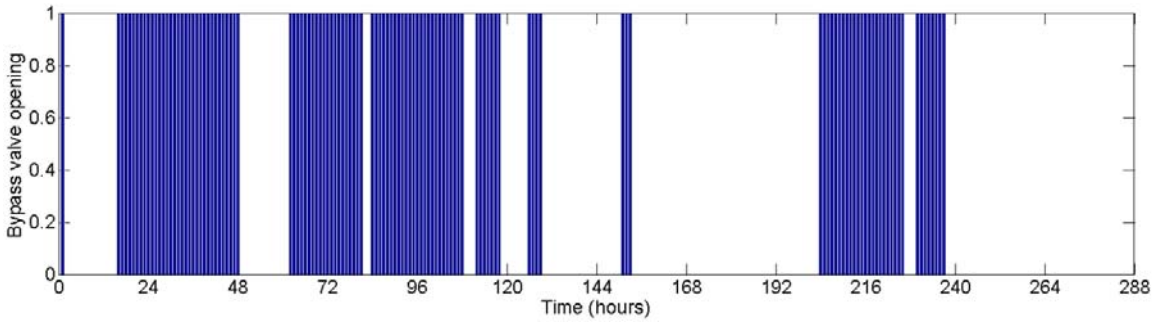


Figure 5.11 The operational schedule found as the smart seed for the bypass valve of single NC 150-200 installed at location 3.

Figure 5.12 compares the fitness values of the individuals shown in Figures 5.7-5.11. In this figure the blue circles demonstrate how the best solution found by the non-seeded GA improves as the number of individuals in the population is increased and the red line indicates the fitness value of the individual found by the iterative process as the smart seed to the GA. It is clear from this figure that the population size needs to be increased significantly in order to find a solution which has a fitness value higher that the smart seed. As a result of these test runs, it is decided that smart seeded genetic algorithm with a population size of 1000 is used for the optimization of the operational schedules. By this way, the genetic algorithm starts with an individual which has considerably good fitness and it may improve this fitness value with a lower computational effort.

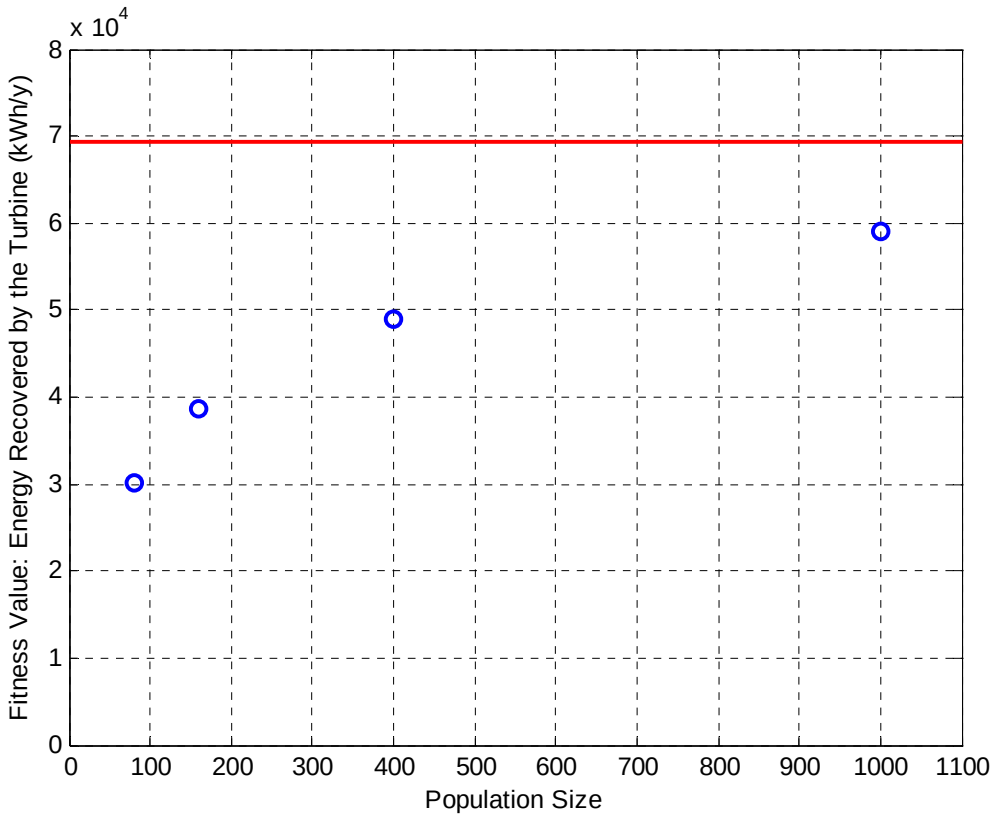


Figure 5.12 Fitness values of the smart seed and the GA outputs for different population sizes.

5.5.2 Pump Driven Network

In this study, 5 different configurations of energy recovery systems are proposed by installing single micro turbine (either NC100-200 or NC 150-200) at locations 3 and 4 (Figure 5.4) and NC 150-200 at both locations. The optimal operational schedule which satisfies consumer demands without pressure violations and resulting energy gain are determined for each scenario. The minimum pressure that needs to be satisfied is set as 20 psi.

Figures 5.13 and 5.14 show the optimal bypass valve operation schedules found for the single turbine case at location 3 for the micro turbines NC 100-200 and NC 150-200, respectively. When Figures 5.13 and 5.14 are compared, the effect of head curve of a turbine (Figure 5.5) on the optimal operational schedule can easily be seen since it is obvious that NC 100-200 which is a higher head turbine needs more bypass valve openings than NC 150-200 which is a lower head turbine.

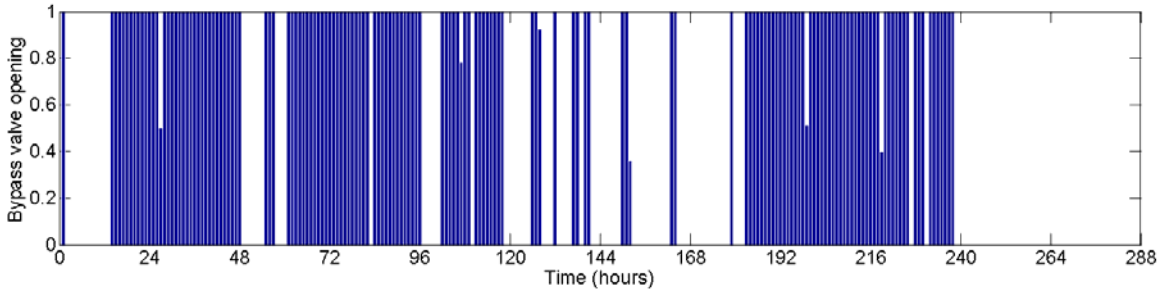


Figure 5.13 Operational scheduling for the bypass valve of single NC 100-200 installed at location 3.

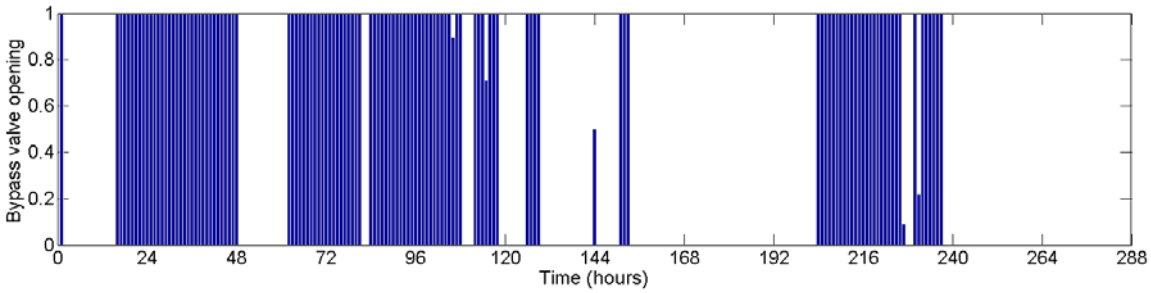


Figure 5.14 Operational scheduling for the bypass valve of single NC 150-200 installed at location 3.

Similar trend for the turbine head can be observed when Figures 5.15 and 5.16 which show the optimal operational schedules found for the single turbine case at location 4 for the micro turbines NC 100-200 and NC 150-200, respectively, are compared. These figures also provide the opportunity to compare the candidate locations of the micro turbines. For example, when Figures 5.13 and 5.15 are compared, one can say that since location 3 needs more bypass valve opening than location 4 for the same micro turbine, it is a more critical point for the operation of the water distribution system. Similar conclusion can be drawn when Figures 5.15 and 5.16 are compared.

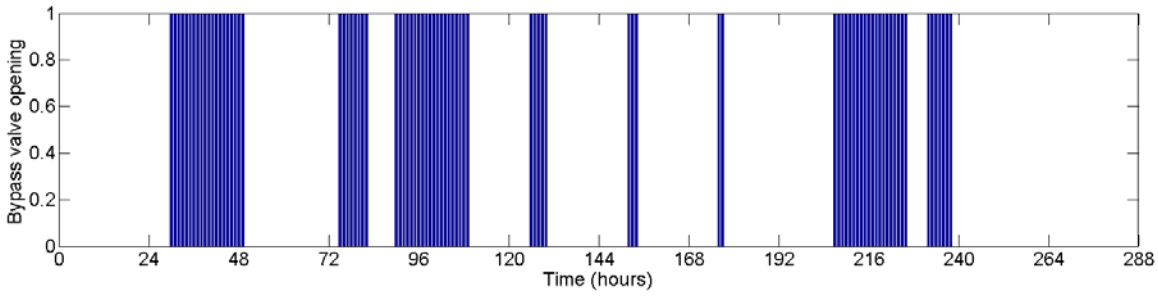


Figure 5.15 Operational scheduling for the bypass valve of single NC 100-200 installed at location 4.

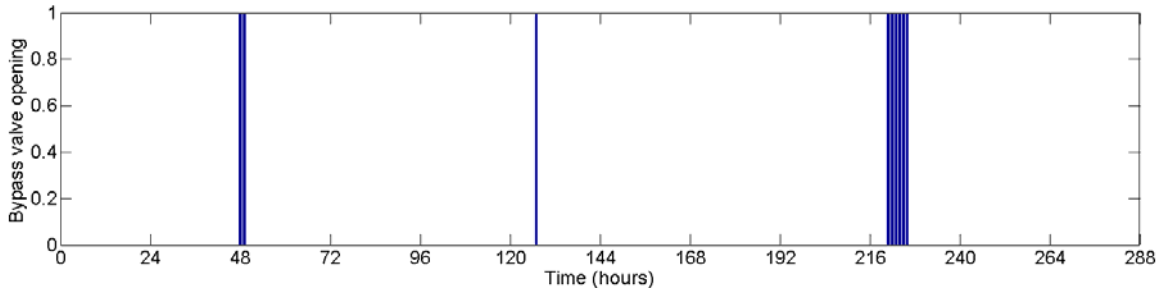
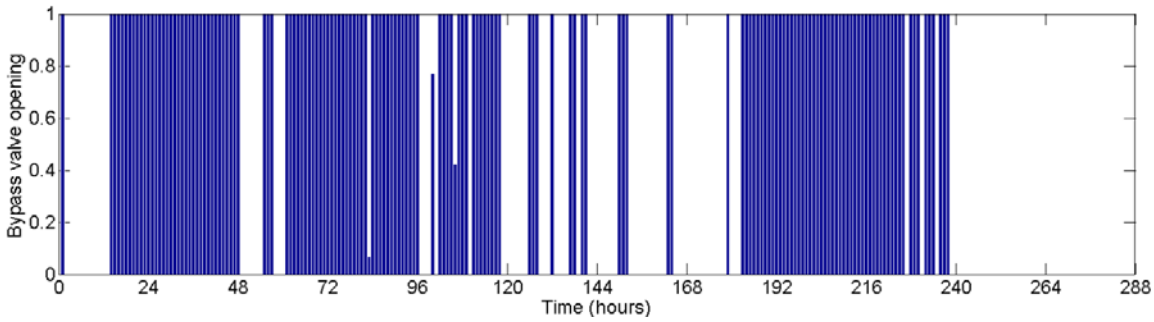


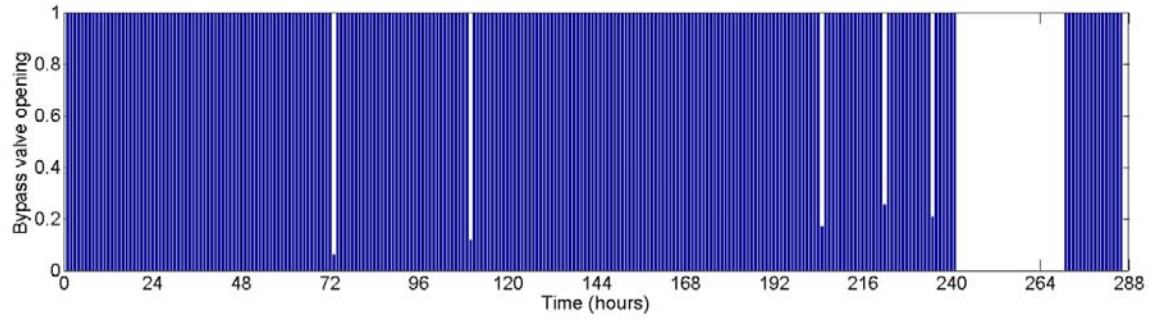
Figure 5.16 Operational scheduling for the bypass valve of single NC 150-200 installed at location 4.

The operating schedules of the bypass valves of individual micro turbines found for the double turbine case where one NC 100-200 is installed at location 3 and one NC 100-200 is installed at location 4 (Figure 5.4) are reported in Figure 5.17. The operational schedules of the bypass valves for the similar case with the double turbine NC150200 are shown in Figure 5.18. In these figures the operating schedules are obtained by the Genetic Algorithm seeded with an individual found by applying the iterative procedure to

find a seed for the bypass valve at Location 4 while the schedule of the bypass valve at Location 3 is the same as the smart seed found for the single turbine installed at this location. When Figures 5.17 and 5.18 are compared one can see that the high head micro turbines (NC100-200's in Figure 5.17) require more bypass valve openings than the low head micro turbines (NC150-200's in Figure 5.18).

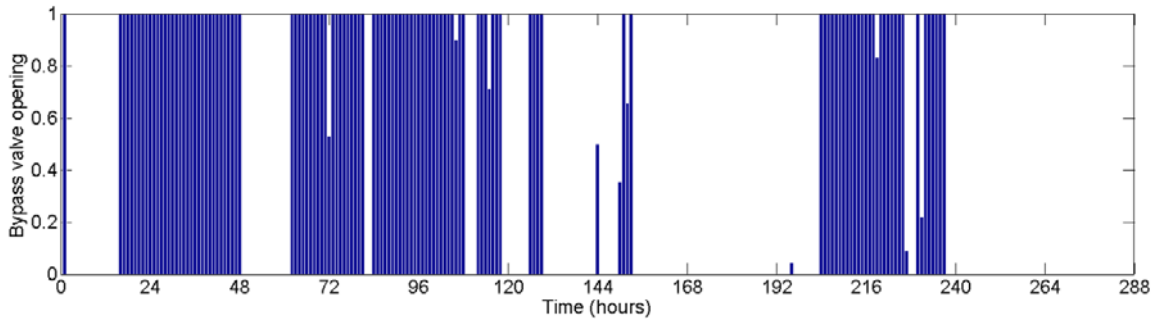


(a) Location 3

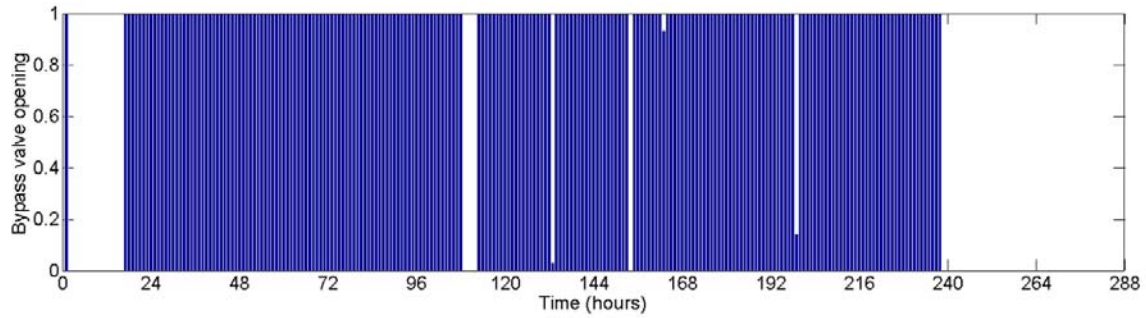


(b) Location 4

Figure 5.17 Operational scheduling for the bypass valves of two NC 100-200's installed at locations 3 and 4.



(a) Location 3



(b) Location 4

Figure 5.18 Operational scheduling for the bypass valves of two NC 150-200's installed at locations 3 and 4.

In order to find the best energy recovery system among the 5 configurations described above, it is necessary to estimate the energy gained by each system. At this point, it should be noted that because Dover Township water distribution system is driven by pumps, it is important to track the amount of annual energy consumed by the pumps before and after the energy recovery system is installed. In its original state when there are no turbines installed, the pumps of this water distribution system consume

approximately 3.5 GWh/y of energy to maintain the flow. Depending on the configuration of the energy recovery system, the amount of this energy consumption may decrease or increase when the micro turbines are introduced to the water distribution network. Table 5.2 summarizes the energy budgets for the energy recovery system configurations considered. In this table, it can be seen that the energy used by the pumps has decreased after the energy recovery system is installed in all cases except for the NC 150-200 installation at location 4 (Figure 5.4) which resulted in an increased energy consumption. Thus, in most of the cases, energy recovery system installation resulted in not only energy production by the micro turbine(s) but also energy savings at the pumps. This energy savings is most significant at location 3 where even though the energy recovered by the micro turbines are lower than that at location 4, the net energy gain is considerably higher due to the high energy savings at the pumps. Also, the installation of NC 150-220 at locations 3 and 4 resulted in the highest energy production by the turbines (89,209 kWh/y). However, this configuration did not save very high pumping energy and net energy gain was not as high as a single turbine installed at location 3. As a result, for this pump driven water distribution system, the best configuration which has the highest net energy gain is the double NC 100-200 installed at locations 3 and 4 with a net energy gain of 274,990 kWh/y.

Table 5.2 Energy budgets for the candidate energy recovery system configurations in pump driven network.

Energy Recovery System Configuration	Energy savings at the pumps (kWh/y)	Energy recovered by the turbines (kWh/y)	Net Energy Gain (kWh/y)
NC 100-200 at Location 3	223,869	30,061	253,930
NC 150-200 at Location 3	152,989	70,211	223,200
NC 100-200 at Location 4	13,080	80,744	93,824
NC 150-200 at Location 4	-27,162	63,597	36,435
NC 100-200 at Locations 3 and 4	228,464	46,526	274,990
NC 150-200 at Locations 3 and 4	132,681	89,209	221,890

According to U.S. Energy Information Administration (USEIA 2012), the average annual electricity consumption of a U.S. residential utility customer was 11,496 kWh in 2010. This average energy consumption value can be used to estimate the economic significance of the candidate energy recovery systems as shown in Figure 5.19. In this figure, the blue, wider bars indicate the number of average American homes that can be

fed by the energy gained by utilizing a given energy recovery system. As expected from Table 5.2, the double NC 100-200's installed at locations 3 and 4 support electricity consumption of the highest number (25) of average American homes. Besides its economic effects, this energy recovery has important environmental benefits. U.S. Environmental Protection Agency (USEPA 2012) suggests that an emission factor of 6.8956×10^{-4} metric tons CO_2/kWh can be used to calculate the equivalencies for emissions reductions from energy efficiency or renewable energy programs. This emission factor is utilized to estimate the annual equivalent CO_2 emissions reduction for each of the candidate energy recovery system as indicated by the thinner, red bars in Figure 5.19. The double NC 100-200's installed at locations 3 and 4 result in the highest CO_2 emissions reduction (190 tons/y).

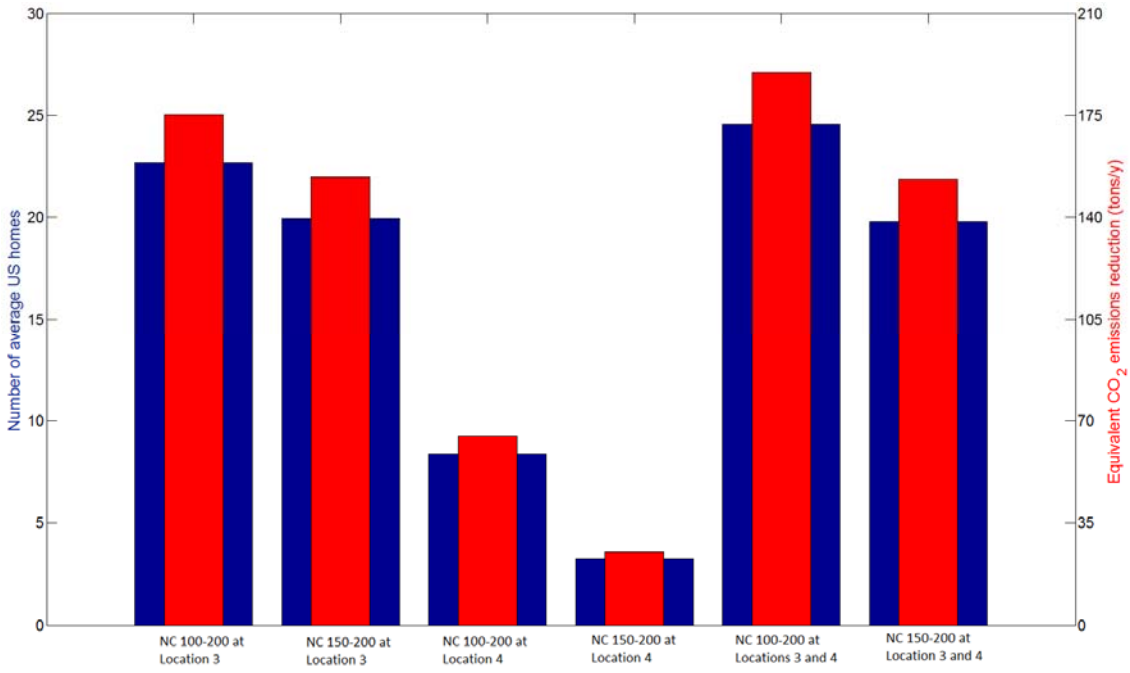


Figure 5.19 Economic and environmental impacts of the energy savings in pump driven network.

5.5.3 Gravity Driven Network

The same analysis is performed on the gravity driven network as explained in section 1.4. Therefore all six energy recovery system configurations are tested in this this new hypothetical water distribution system. The constant head reservoirs supply steady source of energy to the network lowering the chance of pressure violations in case of micro turbine installation. As a result, several configurations such as single NC100200 or NC150200 and double NC150200 installed at locations 3 and 4 do not require operational schedules for the micro turbines. However, double NC100200 installed at locations 3 and 4 cause pressure violations when operated continuously without scheduling. Thus an optimal operation schedule for the micro turbines is found for this configuration. In this operational schedule, NC100200 installed at location 3 can operate with bypass valve fully closed for the entire service period. However, NC100200 installed at location 4 requires bypass valve openings as shown in Figure 5.20.

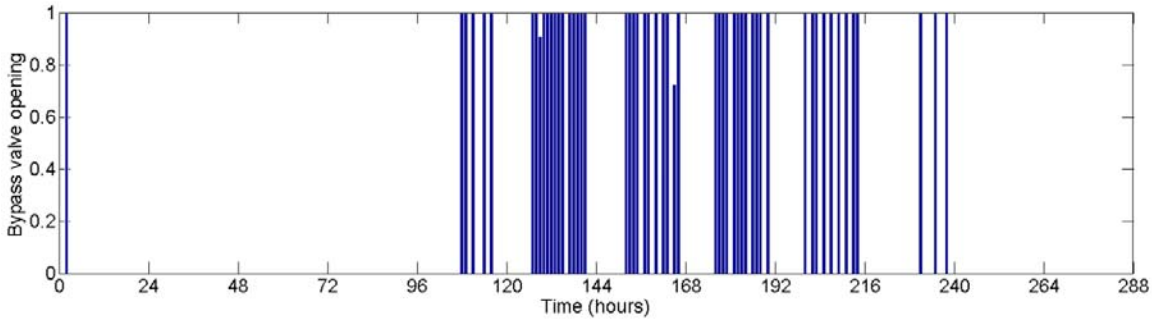


Figure 5.20 Operational scheduling for the bypass valve of NC 100-200 installed at location 4 for the case of double NC100200 installed at locations 3 and 4.

In order to find the best energy recovery system configuration among the 6 alternatives tested in this study, the energy recovered by the micro turbines needs to be calculated. Table 5.3 presents the results of this energy recovery calculation. Here it needs to be noted that since the water distribution system has been converted into a hypothetical gravity driven network without pumps, the net energy gained by the recovery system is equal to the energy produced by the turbines. When Table 5.3 is examined it can be clearly seen that energy recovery by a single turbine installation in this three-reservoir system is not very high when compared with a double turbine energy recovery system. The reason behind this is the fact that a turbine installation at an inlet pipe increases the energy loss for the flow through that pipe causing a significant decrease in the pipe if the system can acquire the necessary flow from other inlet points to satisfy the demands. This significant decrease in the flow rate is reflected as significant decrease in the energy production. In a double turbine situation, since 2 of the 3 inlet pipes have significant increase in the head loss due to the micro turbines installed. This might have resulted in a significant decrease in the flow rate at the pipes where the turbines are installed if the system could acquire the necessary flow from one inlet pipe which is intact. However, one constant head reservoir is not enough to compensate the required flow rate. As a result, the conservation of mass principle necessitates that the flow rate through the inlet pipes do not decrease significantly. Thus, double turbine configurations have considerably high net energy gain as shown in Table 5.3.

Table 5.3 Energy budgets for the candidate energy recovery system configurations in gravity driven network.

Energy Recovery System Configuration	Net Energy Gain (kWh/y)
NC 100-200 at Location 3	66,669
NC 150-200 at Location 3	96,457
NC 100-200 at Location 4	42,757
NC 150-200 at Location 4	42,359
NC 100-200 at Locations 3 and 4	278,870
NC 150-200 at Locations 3 and 4	376,830

When we look at the economic and environmental impacts of these energy recovery configurations (Figure 5.21), we can again see that energy recovery systems with two turbines can supply energy for significantly higher number of average American homes which results in higher reduction in carbon dioxide emissions when compared to single turbine energy recovery systems. In our best case scenario which is double NC150-200 installed at locations 3 and 4, the energy production by the micro turbines is equal to the energy consumption of 33 average American homes and by utilizing this energy we can reduce the carbon dioxide emissions by 260 metric tons annually.

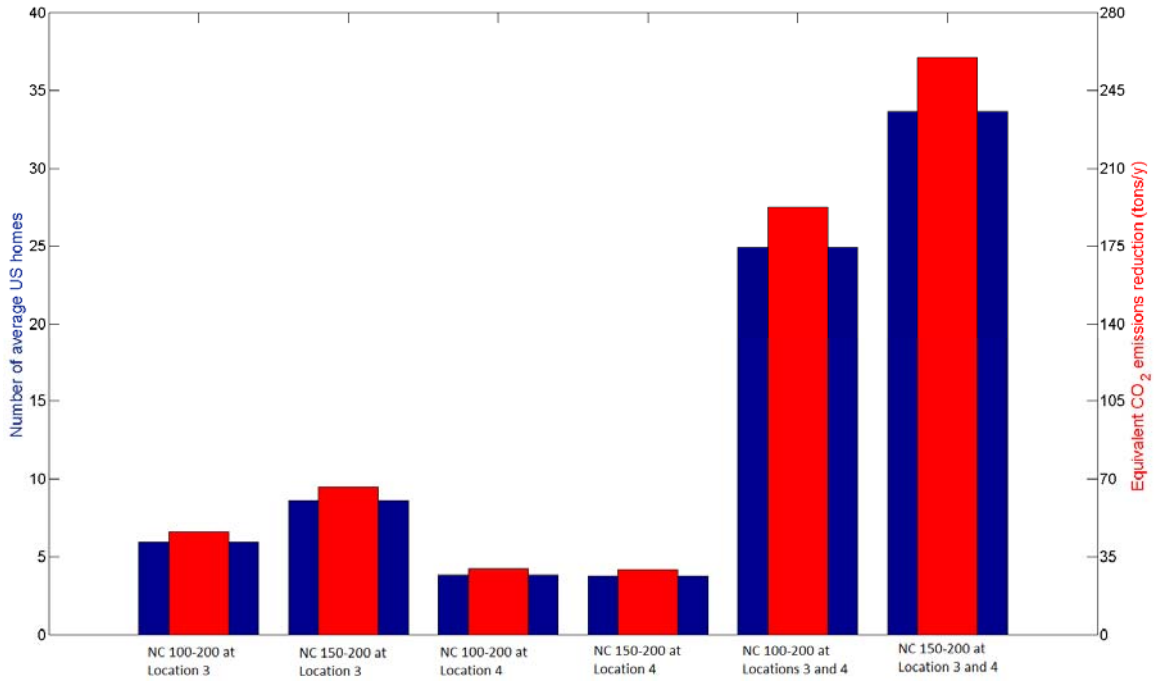


Figure 5.21 Economic and environmental impacts of the energy savings in pump driven network.

5.6 Conclusions

In this study, Genetic algorithms are used to find the optimal operating schedule of an energy recovery system in water distribution systems. The objective function of this optimization problem is formulated as the maximization of energy recovery at the micro turbines such that the consumer demands are always satisfied without pressure violations. The candidate locations for the energy recovery system are determined from the excess energy input distribution of the water supply system. For this study, two locations which have the highest excess energy inputs are selected. In order to demonstrate the effect of

the micro turbine used, two different turbines are tested. The proposed methodology is applied for different combinations of locations and turbine types.

The optimization algorithm has successfully determined operational schedules for the energy recovery systems which never violate the pressure constraint. The results show that location and turbine type have significant effects on the optimal operation schedules. A high head turbine requires more bypass valve openings than a low head turbine and one location can be more critical in satisfying the pressure constraint than another location.

Another important outcome of this study is the fact that the energy recovery system installed in pump driven networks may not only produce energy at the micro turbines but also decrease the energy consumption at the pumps depending on the location and the type of the micro turbines used. While some energy recovery systems may decrease energy consumption, some recovery systems may result in increased pump energy. In pump driven networks, the energy savings at the pumping stations constitute a significant portion of the net energy gain which eventually reveals the best configuration for the energy recovery system.

Another conclusion can be the fact that results of this study demonstrate the important economic and environmental impacts of the energy recovery systems in water distribution networks. Even in the pump driven water distribution network serving a typical small town in USA, the energy savings can support the electricity consumption of more than 20 average U.S. homes corresponding to a reduction of 177 tons of CO_2

emissions annually. In the gravity driven system, these numbers are increased to 33 average American homes and 260 tons of CO_2 reduction. These economic and environmental impact numbers will obviously increase for larger water distribution networks.

The available excess energy potential of a water distribution system may vary significantly depending on the complexity and type of the driving force of the network. A large network may have higher consumer demands which require higher flow rates at the inlet locations increasing the energy generation at the times of turbine operation. In addition to the complexity, the driving force of the network significantly affects the energy gain from a water distribution system. In a gravity driven water distribution system, where the nature supplies the required energy for the flow, the available excess energy potential may be much higher than that of a pump driven network. It should also be noted that since the gravity driven network does not require any pumps, the only energy that need to be tracked is the energy produced at the turbines.

The results of this study are comparable with the results of another work (Giugni, Fontana et al. 2009) which is based on a gravity driven network where three energy recovery systems were proposed. The energy production of these three options ranged from 418.8 kWh/d to 821.6 kWh/d. In the current study, the highest net energy gain was found for NC 100-200 at location 3 as 695.7 kWh/d in pump driven network (Table 5.2).

In gravity driven network, the highest energy production was 1032.4 kWh/d (Table 5.3) by 2 NC150-200 installed at locations 3 and 4.

CHAPTER 6

CONCLUSIONS AND FUTURE DIRECTIONS

This thesis provides optimal design methodologies for real-time water quality monitoring systems in rivers and energy recovery systems in water distribution networks. Moreover, it presents a novel solution approach for the ill-posed problem of contaminant source location identification. Therefore, it focuses on the analysis of protection of fresh water resources and utilization of renewable energy sources using optimization algorithms. These objectives are the two of the most important and challenging engineering problems today and in the future. In sections below, general conclusions and directions for each research study covered in this thesis are presented.

6.1 Real-Time Water Quality Monitoring Networks for River Systems

In this study, a methodology has been developed for the optimal design of real-time monitoring networks in river systems (Telci, Nam et al. 2008; Telci, Nam et al. 2009). The design objectives are determined as the early detection time and the reliability of the monitoring system designed. These two criteria are essential to protect humans from adverse effects of exposure to harmful contaminants. Focusing on these two objectives, this study proposes a methodology that is based on the transient behavior of a random contamination event or multiple events in a river network. This model is based on the

hydrodynamics and the contaminant fate and transport characteristics of the river system that is under study. The information gathered from this analysis is used in an optimization model to identify the best monitoring locations in the river network in real time which would satisfy the two objectives identified above. In this study, enumeration and Genetic Algorithms were used as the optimization tools.

As an important clue for the future direction of this research, this work has inspired a recent study (Park, Kim et al. 2010) which proposes another optimization tool (Nested Partitions Method) for the analysis of this problem which was the topic of another PhD thesis. This new research shows that the subject of water quality monitoring design has many aspects open to interdisciplinary research. In addition to application of new optimization tools, new objectives can be developed for these monitoring systems as many monitoring projects may serve for different purposes. For example, an optimal monitoring system can be designed to observe the effect of climate change and land use on hydrology and pollutant transport. It should be noted that any new objective function can be easily implemented in the optimization algorithm presented in this thesis. It is possible to extend this optimization problem to other environments such as estuaries as one of my future research topics. In this case, the contaminant transport model used in this thesis (EPA SWMM) needs to be replaced with another model developed for estuaries.

6.2 Identification of Contaminant Source Locations in River Systems

Once the real-time monitoring stations are optimally located in the river system, they provide continuous information about the quality of the river water. This information can be used for long term management of the river system. In this research (Telci and Aral 2011), it is shown that continuous observations at the monitoring stations can be used for rapid identification of contaminant source locations. The methodology proposed in this study parameterizes the breakthrough curve of the contamination observed at a monitoring station and utilizes a probabilistic classification routine which associates these observations with one or more of the candidate spill locations in the river network.

In future studies, this research topic can be extended to other environments such as groundwater, lakes, estuaries and air where a real-time monitoring system is providing continuous observations. Moreover, the methodology proposed in this thesis is so flexible that new parameters can be easily included in the classification algorithm which may increase the performance of the approach.

6.3 Renewable Energy Production from Water Distribution Systems

This research described in Chapter 5 is based on harvesting the available excess energy in water distribution systems inevitably produced while maintaining adequate pressures

throughout the network. Therefore the energy produced by the proposed system is pure green energy free from any environmental side effects. This energy recovery is possible by the utilization of mini and micro hydroelectric turbines. In this study, an optimization approach for the design of energy recovery systems in water distribution networks is developed. This methodology is based on finding the best locations for micro hydroelectric plants in the network to recover the excess energy. Due to the unsteady nature of flow in water distribution networks, the proposed methodology also determines optimum operation schedules for the micro turbines. The objective of the optimization algorithm is determined as the net annual energy gain from the energy recovery system. Genetic Algorithm supported by a "Smart Seeding" procedure is used to solve this complex optimization problem. In this approach, which significantly reduces the computational effort, a good feasible solution is supplied to the genetic algorithm as an individual of the initial population. The energy production from water distribution systems can vary significantly as the size, complexity and driving mechanism of the flow change. The proposed methodology is tested in pump and gravity driven networks and the results are compared and it is concluded that the gravity driven networks are more promising for energy production than the energy production. This comparison also showed that the proposed approach is not only provides an optimal design for the energy recovery system but also evaluates the available excess energy potential of a given water distribution system.

One of the essential inputs required by the methodology proposed in this thesis are the hourly demand patterns of the outlet nodes during the service period. This information

may not be available for some water distribution systems. Therefore as a future step, the design approach outlined in this thesis can be modified such that one or more measureable parameters such as pressures at specific locations in the water distribution networks are implemented to find the real-time optimal operation patterns of the micro turbines. In this case, the demand patterns can be regarded as stochastic parameters and a probabilistic approach outlined in Chapter 4 can be implemented to solve this new design problem. In addition, this work can be extended to other environments such as river networks where micro hydro power applications are recently growing. In addition to these micro hydro power applications, a future research topic can be optimal integration of new renewable energy sources to the existing power grid. Since these renewable energy sources depend on natural inputs such as wind and solar energy, their energy output will have a stochastic nature. The design of the adaptation tools of existing power grid to these new energy sources such as pump storage units will require stochastic optimization techniques and has not been studied in an integrated manner.

REFERENCES

Afshar, A., F. Benjemaa, et al. (1990). "OPTIMIZATION OF HYDROPOWER PLANT INTEGRATION IN WATER-SUPPLY SYSTEM." Journal of Water Resources Planning and Management-Asce **116**(5): 665-675.

Allan, I. J., G. A. Mills, et al. (2006). "Strategic monitoring for the European Water Framework Directive." Trac-Trends in Analytical Chemistry **25**(7): 704-715.

Aral, M. M. and J. Guan (1996). Genetic algorithms in search of groundwater pollution sources. Advances in Groundwater Pollution Control and Remediation. M. M. Aral. Kluwer, Dordrecht, The Netherlands. **9:** 346-371.

Aral, M. M., J. Guan, et al. (2004). "Optimal reconstruction of historical water supply to a distribution system: A. Methodology." Journal of Water and Health **2**(3): 123-136.

Aral, M. M., J. Guan, et al. (2004). "Optimal reconstruction of historical water supply to a distribution system: B. Applications." Journal of Water and Health **2**(3): 137-156.

Aral, M. M., J. B. Guan, et al. (2001). "Identification of contaminant source location and release history in aquifers." Journal of Hydrologic Engineering **6**(3): 225-234.

Bagtzoglou, A. C. and S. A. Baun (2005). "Near real-time atmospheric contamination source identification by an optimization-based inverse method." Inverse Problems in Science and Engineering **13**(3): 241-259.

Beckers, C. V., S. G. Chamberlain, et al. (1972). Quantitative methods for preliminary design of water quality surveillance systems. Socioeconomic Environmental Studies Series. Washington, DC., US EPA. **Report No. EPA-R5-72-001**.

Bieri, M., J.-L. Boillat, et al. (2010). Economic Evaluation of Turbining Potential in Drinking Water Supply Networks Hidroenergia 2010. Lausanne, CH**:** CD Rom.

Boano, F., R. Revelli, et al. (2005). "Source identification in river pollution problems: A geostatistical approach." Water Resources Research **41**(7).

Brooks, M. C. and W. R. Wise (2005). "Quantifying uncertainty due to random errors for moment analyses of breakthrough curves." Journal of Hydrology **303**(1-4): 165-175.

Chadalavada, S. and B. Datta (2008). "Dynamic optimal monitoring network design for transient transport of pollutants in groundwater aquifers." Water Resources Management **22**(6): 651-670.

Cieniawski, S. E., J. W. Eheart, et al. (1995). "Using Genetic Algorithms to Solve a Multiobjective Groundwater Monitoring Problem." Water Resources Research **31**(2): 399-409.

Datta, B. and R. C. Peralta (1986). Expert pattern recognition for pollution source identification. Water Forum '86. M. Karamouz, G. R. Baumli and W. J. Brick. New York, ASCE. **1:** 195-202.

Derakhshan, S. and A. Nourbakhsh (2008). "Experimental study of characteristic curves of centrifugal pumps working as turbines in different specific speeds." Experimental Thermal and Fluid Science **32**(3): 800-807.

Directive 2000/60/EC of The European Parliament and of The Council of 23 October 2000 establishing a framework for Community action in the field of water policy, Off. J. Eur. Comm. L327 (2000) 1.

Dixon, W. and B. Chiswell (1996). "Review of aquatic monitoring program design." Water Research **30**(9): 1935-1948.

Dixon, W., G. K. Smyth, et al. (1999). "Optimized selection of river sampling sites." Water Research **33**(4): 971-978.

Dudhani, S., A. K. Sinha, et al. (2006). "Assessment of small hydropower potential using remote sensing data for sustainable development in India." Energy Policy **34**(17): 3195-3205.

Farre, M., R. Brix, et al. (2005). "Screening water for pollutants using biological techniques under European Union funding during the last 10 years." Trac-Trends in Analytical Chemistry **24**(6): 532-545.

Gerlach, R. W., R. J. White, et al. (1997). "Field evaluation of an immunoassay for benzene, toluene and xylene (BTX)." Water Research **31**(4): 941-945.

Gevaert, V., F. Verdonck, et al. (2009). "Evaluating the usefulness of dynamic pollutant fate models for implementing the EU Water Framework Directive." Chemosphere **76**(1): 27-35.

Giugni, M., N. Fontana, et al. (2009). Energy saving policy in water distribution networks. International Conference on Renewable Energies and Power Quality (ICREPQ'09). Valencia, Spain.

Goldberg, D. E. (1989). Genetic Algorithms in Search, Optimization, and Machine Learning. MA, Addison-Wesley, Reading.

Gorelick, S. M., B. Evans, et al. (1983). "IDENTIFYING SOURCES OF GROUNDWATER POLLUTION - AN OPTIMIZATION APPROACH." Water Resources Research **19**(3): 779-790.

Grubner, O. (1971). "INTERPRETATION OF ASYMMETRIC CURVES IN LINEAR CHROMATOGRAPHY." Analytical Chemistry **43**(14): 1934-1937.

Guan, J. B. and M. M. Aral (1999). "Progressive genetic algorithm for solution of optimization problems with nonlinear equality and inequality constraints." Applied Mathematical Modelling **23**(4): 329-343.

Guan, J. B., M. M. Aral, et al. (2006). "Identification of contaminant sources in water distribution systems using simulation-optimization method: Case study." Journal of Water Resources Planning and Management-Asce **132**(4): 252-262.

Guan, J. B., M. M. Aral, et al. (2006). Optimization Model and Algorithms for Design of Water Sensor Placement in Water Distribution Systems. ASCE Water Distribution System Analysis Symposium (Battel of the Water Sensor Networks). Cincinnati, OH.

Gunduz, O. and M. M. Aral (2003). Hydrologic Modeling of the Lower Altamaha River Basin. 2003 Georgia Water Resources Conference, Athens, Georgia, University of Georgia.

Gunduz, O. and M. M. Aral (2003). A Simultaneous Solution Approach for Coupled Surface and Subsurface Flow Modeling. Atlanta, GA, MESL, CEE, Georgia Institute of Technology. **MESL-02-03:** 98pp.

Harmancioglu, N. B. and N. Alpaslan (1992). "Water-Quality Monitoring Network Design - a Problem of Multiobjective Decision-Making." Water Resources Bulletin **28**(1): 179-192.

Harmancioglu, N. B. and N. Alpaslan (1994). "Basic Approaches in Design of Water-Quality Monitoring Networks." Water Science and Technology **30**(10): 49-56.

Harmon, T. C., R. F. Ambrose, et al. (2007). "High-resolution river hydraulic and water quality characterization using rapidly deployable networked infomechanical systems (NIMS RD)." Environmental Engineering Science **24**(2): 151-159.

Holland, J. H. (1975). Adaptation in Natural and Arti®cial Systems, University of Michigan Press, Ann Arbor.

Hudak, P. F., H. A. Loaiciga, et al. (1995). "Regional-Scale Ground-Water Quality Monitoring Via Integer Programming." Journal of Hydrology **164**(1-4): 153-170.

Icaga, Y. (2005). "Genetic algorithm usage in water quality monitoring networks optimization in Gediz (Turkey) river basin." Environmental Monitoring and Assessment **108**(1-3): 261-277.

IPCC (2007). Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. S. Solomon, D. Qin, M. Manninget al. Cambridge, United Kingdom and New York, NY, USA**:** 996.

Jiang, H. (2008). Adaptive Feature Selection in Pattern recognition and ultra-wideband radar signal analysis. Pasadena, California, California Institute of Technology. **PhD:** 10-31.

Johansson, T. B., K. McCormick, et al. (2004). The Potentials of Renewable Energy. International Conference for Renewable Energies, Bonn, Germany.

Kanase-Patil, A. B., R. P. Saini, et al. (2010). "Integrated renewable energy systems for off grid rural electrification of remote area." Renewable Energy **35**(6): 1342-1349.

Karamouz, M., A. K. Nokhandan, et al. (2009). "Design of on-line river water quality monitoring systems using the entropy theory: a case study." Environmental Monitoring and Assessment **155**(1-4): 63-81.

Karci, A. (2004). Novelty in the generation of initial population for genetic algorithms. Knowledge-Based Intelligent Information and Engineering Systems, Pt 2, Proceedings. M. G. Negoita, R. J. Howlett and L. C. Jain. Berlin, Springer-Verlag Berlin. **3214:** 268-275.

Keats, A., E. Yee, et al. (2007). "Bayesian inference for source determination with applications to a complex urban environment." Atmospheric Environment **41**(3): 465-479.

Kwiatkowski, R. E. (1991). "Statistical Needs in National Water-Quality Monitoring Programs." Environmental Monitoring and Assessment **17**(2-3): 253-271.

Liu, X. and Z. Q. J. Zhai (2008). "Location identification for indoor instantaneous point contaminant source by probability-based inverse Computational Fluid Dynamics modeling." Indoor Air **18**(1): 2-11.

Liu, X. and Z. Q. J. Zhai (2009). "Prompt tracking of indoor airborne contaminant source location with probability-based inverse multi-zone modeling." Building and Environment **44**(6): 1135-1143.

Lo, S. L., J. T. Kuo, et al. (1996). "Water quality monitoring network design of Keelung River, northern Taiwan." Water Science and Technology **34**(12): 49-57.

Maaranen, H., K. Miettinen, et al. (2004). "Quasi-random initial population for genetic algorithms." Computers & Mathematics with Applications **47**(12): 1885-1895.

Maaranen, H., K. Miettinen, et al. (2007). "On initial populations of a genetic algorithm for continuous optimization problems." Journal of Global Optimization **37**(3): 405-436.

Maslia, M. L., J. B. Sautner, et al. (2001). Historical reconstruction of the water distribution system serving the Dover Township area, New Jersey: January 1962—December 1996. Atlanta, Agency for Toxic Substances and Disease Registry.

Maslia, M. L., J. B. Sautner, et al. (2000). "Using water-distribution system modeling to assist epidemiologic investigations." Journal of Water Resources Planning and Management-Asce **126**(4): 180-198.

Monition, L. (1984). Micro hydroelectric power stations. Chichester [West Sussex] ;, J. Wiley.

Nam, K. (2008). Optimization of Paths and Locations of Water Quality Monitoring Systems in Surface Water Environments. School of Civil and Environmental Engineering. Atlanta, GA, Georgia Institute of Technology. **PhD**.

Nam, K. and M. M. Aral (2007). Optimal Placement of Monitoring Sensors in Lakes. University of Georgia Water Resources Conference, Athens, GA.

Nam, K. and M. M. Aral (2007). Optimal Sensor Placement for Wind-Driven Circulation Environment in a Lake. ASCE, Proceedings of the World Water and Environmental Resources Congress. Tampa, FL**:** [CD ROM].

NCDC. (2009, July 16, 2010). "NNDC Climate Data Online." 2009, from http://www7.ncdc.noaa.gov/CDO/cdo.

Neupauer, R. M. and R. H. Lin (2006). "Identifying sources of a conservative groundwater contaminant using backward probabilities conditioned on measured concentrations." Water Resources Research **42**(3).

Neupauer, R. M. and J. L. Wilson (1999). "Adjoint method for obtaining backward-in-time location and travel time probabilities of a conservative groundwater contaminant." Water Resources Research **35**(11): 3389-3398.

Neupauer, R. M. and J. L. Wilson (2001). "Adjoint-derived location and travel time probabilities for a multidimensional groundwater system." Water Resources Research **37**(6): 1657-1668.

Neupauer, R. M. and J. L. Wilson (2004). "Numerical implementation of a backward probabilistic model of ground water contamination." Ground Water **42**(2): 175-189.

Ning, S. K. and N. B. Chang (2004). "Optimal expansion of water quality monitoring network by fuzzy optimization approach." Environmental Monitoring and Assessment **91**(1-3): 145-170.

Ning, S. K. and N. B. Chang (2005). "Screening the relocation strategies of water quality monitoring stations by compromise programming." Journal of the American Water Resources Association **41**(5): 1039-1052.

Nunes, V. and J. L. Genta (1996). "Micro and mini hydroelectric power assessment in Uruguay." Renewable Energy **9**(1-4): 1235-1238.

Ouyang, H. T., H. Yu, et al. (2008). "Design optimization of river sampling network using genetic algorithms." Journal of Water Resources Planning and Management-Asce **134**(1): 83-87.

Paish, O. (2002). "Micro-hydropower: status and prospects." Proceedings of the Institution of Mechanical Engineers Part a-Journal of Power and Energy **216**(A1): 31-40.

Park, C., S.-H. Kim, et al. (2010). Designing Optimal Water Quality Monitoring Networks for River Systems and Application to a Hypothetical Case. Proceedings of the 2010 Winter Simulation Conference. Austin, TX.

Park, S. Y., J. H. Choi, et al. (2006). "Design of a water quality monitoring network in a large river system using the genetic algorithm." Ecological Modelling **199**(3): 289-297.

Pokharel, G. R., A. B. Chhetri, et al. (2008). "Decentralized micro-hydro energy systems in Nepal: En route to sustainable energy development." Energy Sources Part B-Economics Planning and Policy **3**(2): 144-154.

Ponterosso, P. and D. S. J. Fox (1999). "Heuristically seeded Genetic Algorithms applied to truss optimisation." Engineering with Computers **15**(4): 345-355.

Ramos, H., D. Covas, et al. (2005). Available energy assessment in water supply systems. XXXI IAHR Congress. Seoul, Korea.

Rodriguez-Mozaz, S., M. J. L. de Alda, et al. (2005). "Biosensors for environmental monitoring - A global perspective." Talanta **65**(2): 291-297.

Rogers, S. W. (2009). Multiobjective Optimization of Contaminant Sensor Locations in Drinking Water Distribution Systems Using Nodal Importance Concepts. School of Civil and Environmental Engineering. Atlanta, GA, Georgia Institute of Technology. **PhD**.

Rogers, S. W., J. B. Guan, et al. (2007). Nodal Importance Concept for Computational Efficiency in Optimal Sensor Placement in Water Distribution Systems. ASCE, Proceedings of the World Water and Environmental Resources Congress. Tampa, FL.

Rossman, L. A. (2000). EPANET 2 users manual. Cincinnati, OH, National Risk Management Research Laboratory, U.S. EPA.

Rossman, L. A. (2007). Storm Water Management Model User's Manual Version 5.0. Cincinnati, OH, National Risk Management Research Laboratory, U.S. Environmental Protection Agency.

Saavedra-Moreno, B., S. Salcedo-Sanz, et al. (2011). "Seeding evolutionary algorithms with heuristics for optimal wind turbines positioning in wind farms." Renewable Energy **36**(11): 2838-2844.

Sadik, O. A. and J. M. Van Emon (1996). "Applications of electrochemical immunosensors to environmental monitoring." Biosensors and Bioelectronics **11**(8): I-XI.

Shannon, C. E. (1948). "A MATHEMATICAL THEORY OF COMMUNICATION." Bell System Technical Journal **27**(4): 623-656.

Sharp, W. E. (1971). "Topologically Optimum Water-Sampling Plan for Rivers and Streams." Water Resources Research **7**(6): 1641-&.

Singh, R. M. and B. Datta (2006). "Identification of groundwater pollution sources using GA-based linked simulation optimization model." Journal of Hydrologic Engineering **11**(2): 101-109.

Singh, R. M. and B. Datta (2007). "Artificial neural network modeling for identification of unknown pollution sources in groundwater with partially missing concentration observation data." Water Resources Management **21**(3): 557-572.

Singh, R. M., B. Datta, et al. (2004). "Identification of unknown groundwater pollution sources using artificial neural networks." Journal of Water Resources Planning and Management-Asce **130**(6): 506-514.

Soffia, C., F. Miotto, et al. (2010). Hydropower potential from the drinking water systems of the Piemonte region (Italy). 4th International Conference on Sustainable Energy & Environmental Protection. Bari, Italy.

Steele, T. D. (1987). "WATER-QUALITY MONITORING STRATEGIES." Hydrological Sciences Journal-Journal Des Sciences Hydrologiques **32**(2): 207-213.

Strobl, R. O. and P. D. Robillard (2008). "Network design for water quality monitoring of surface freshwaters: A review." Journal of Environmental Management **87**(4): 639-648.

Sun, A. Y., S. L. Painter, et al. (2006). "A robust approach for iterative contaminant source location and release history recovery." Journal of Contaminant Hydrology **88**(3-4): 181-196.

Telci, I. and M. Aral (2011). "Contaminant Source Location Identification in River Networks Using Water Quality Monitoring Systems for Exposure Analysis." Water Quality, Exposure and Health **2**(3): 205-218.

Telci, I. T., K. Nam, et al. (2009). "Optimal water quality monitoring network design for river systems." J Environ Manage **90**(10): 2987-2998.

Telci, I. T., K. Nam, et al. (2008). Real time optimal monitoring network design in river networks. World Environmental & Water Resources Congress 2008. Honolulu, Hawaii.

Telci, I. T., K. Nam, et al. (2009). "Optimal water quality monitoring network design for river systems." Journal of Environmental Management **90**(10): 2987-2998.

Tschmelak, J., G. Proll, et al. (2004). "Verification of performance with the automated direct optical TIRF immunosensor (River Analyser) in single and multi-analyte assays with real water samples." Biosensors & Bioelectronics **20**(4): 743-752.

USEIA. (2012). "How much electricity does an American home use?" Retrieved April 26, 2012, from http://www.eia.gov/tools/faqs/faq.cfm?id=97&t=3.

USEPA (2009). National Water Quality Inventory: Report to Congress 2004 Reporting Cycle Washington, DC United States Environmental Protection Agency Office of Water

USEPA. (2012). "Calculations and References." Retrieved April 26, 2012, from http://www.epa.gov/cleanenergy/energy-resources/refs.html.

USGS. (2008, July 17, 2008). "The National Map Seamless Server." 2008, from http://seamless.usgs.gov/.

Vieira, F. and H. M. Ramos (2009). "Optimization of operational planning for wind/hydro hybrid water supply systems." Renewable Energy **34**(3): 928-936.

Yuksek, O. and K. Kaygusuz (2006). "Small hydropower plants as a new and renewable energy source." Energy Sources Part B-Economics Planning and Policy **1**(3): 279-290.