2007

Production well operations optimization in water distribution system using genetic algorithm

Vineet Katiyar
Louisiana State University and Agricultural and Mechanical College, vkatiy1@lsu.edu

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_theses

Part of the Civil and Environmental Engineering Commons

Recommended Citation
https://digitalcommons.lsu.edu/gradschool_theses/3225

This Thesis is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Master's Theses by an authorized graduate school editor of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.
PRODUCTION WELL OPERATIONS OPTIMIZATION IN WATER DISTRIBUTION SYSTEM USING GENETIC ALGORITHM

A Thesis

Submitted to the Graduate Faculty of the Louisiana State University and Agricultural and Mechanical College in partial fulfillment of the requirements for the degree of Master of Science in Civil Engineering

in

The Department of Civil and Environmental Engineering

By
Vineet Katiyar
B.T., Indian Institute of Technology Roorkee, 2005
May 2007
To my Parents

Dr. Vinod Kumar Katiyar

&

Mrs. Vimla Katiyar
Acknowledgements

First and foremost, I would like to thank my major advisor Dr. Frank T-C Tsai for his direction, guidance and long time support for this research work and thus making it possible for me to pursue M.S in Water Resources at LSU. I equally, owe a special debt of gratitude to Dr. Vijay. P. Singh for giving me an opportunity to study in the United States. It is hard to find enough words to express my appreciation for their constant encouragement throughout my master’s studies.

I would like to extend my deepest thanks to Dr. Donald Dean Adrian and Dr. Clinton S. Willson for taking their valuable time out and agreeing to be part of my thesis committee and also helping me out in times of need.

I would not like to miss this opportunity, to thank the great contribution of my family, and teachers in my career growth and for every advancement that I make in my life. I am mightily beholden to my friends Mr. Vikas Singh and Mr. Imtiaz Hossain for their technical guidance in this thesis along with their constant encouragement and help throughout my stay here in the United States. I am also thankful to the Shell Coastal Environmental Modeling Laboratory for providing accessibility to their cluster for parallel computation.

I thank Doug Toy and Robert Goff at the Municipal Utilities Department, Chandler for providing the Chandler water distribution system model and William Greenslade and Nathan Miller at Southwest Ground-water Consultants, Inc. for valuable comments on groundwater modeling.

Finally, I want to thank the Almighty for healthy living and fervently pray for His infinite grace and benevolence.
# Table of Contents

DEDICATION .................................................................................................................. ii

ACKNOWLEDGEMENTS.................................................................................................... iii

ABSTRACT ........................................................................................................................ vi

CHAPTER 1: INTRODUCTION AND LITERATURE REVIEW ........................................... 1
  1.1 Introduction ........................................................................................................... 1
  1.2 Literature Review ................................................................................................... 5
    1.2.1 Water Resources Management Models and Optimization ................... 5
      1.2.1.1 Non-Linear Programming Approach ............................................. 5
      1.2.1.2 Simulated Annealing ................................................................. 8
    1.2.2 Water Distribution System Optimization ............................................... 10
    1.2.3 Reliability, Resilience and Vulnerability of Water Resources Systems ........................................................................................................ 15
    1.2.4 Genetic Algorithm for Water Distribution Systems ............................ 15
    1.2.5 Genetic Algorithm and Linear Programming for Water Resources Management Models .......................................................... 19
    1.2.6 Application of Genetic Algorithm in Groundwater Management Models ........................................................................................................ 19

CHAPTER 2: SIMULATION MODELS ................................................................................. 23
  2.1 Water Distribution System Model: EPANET ....................................................... 23
    2.1.1 Introduction ................................................................................................. 23
    2.1.2 Hydraulic Modeling Capabilities .............................................................. 23
    2.1.3 Chandler Water Resource System .............................................................. 24
    2.1.4 Chandler Water Distribution Model .......................................................... 25
  2.2 Groundwater Management Model: MODFLOW ................................................. 29
    2.2.1 Introduction ................................................................................................. 29
    2.2.2 Groundwater Flow Equation .................................................................... 29
    2.2.3 Chandler Groundwater Model ................................................................. 31
    2.2.4 Chandler MODFLOW Model ................................................................. 31

CHAPTER 3: GENETIC ALGORITHM .................................................................................. 33
  3.1 Introduction ........................................................................................................... 33
  3.2 Genetic Algorithm for Single-Objective Optimization Problems .................. 35
    3.2.1 Coding ........................................................................................................... 36
    3.2.2 Evaluation ................................................................................................. 37
    3.2.3 Selection ....................................................................................................... 38
    3.2.4 Crossover ....................................................................................................... 39
    3.2.5 Crossover for Binary Strings .................................................................... 39
    3.2.6 Mutation ........................................................................................................ 39
    3.2.7 Mutation for Binary Strings ........................................................................ 40
Abstract

The City of Chandler, Arizona lies in semi-arid area and the water resources management has been a major concern by the State citizen. There is always a need of optimizing the available water resources when the rate of demand constantly beats the rate of replenishment. In this study we have developed a management model to link the Chandler water distribution model (EPANET) with the Chandler groundwater flow model (MODFLOW) under an optimization framework (Genetic algorithm). The EPANET model supplied by the Chandler city municipal authority has several improper setting (pumping schedules, valves settings, pressure zones etc.). The original model has several warnings about system being disconnected and unbalanced since it lacks efficient set of constraints. The pumping values present in the well package for the Chandler water planning area do not commensurate with their corresponding values in the EPANET model. In the study when the models are linked it is made sure that the sub-surface and surface models have a proper connection and accurate data transfer, this is the main objective for linking the two models. The overriding goal of the management model is to optimize the well pump operations such that the pump energy cost is minimized. The optimization problem was solved using a genetic algorithm (GA) to search for the optimal 24-hour real-time pumping patterns for all well pumps while conducting 7-day EPANET and MODFLOW simulations. The main optimization function has several objectives like energy, pressure violation, drawdown, and reliability etc., each having a weight or penalty (w or r) in the function. These weights are first determined by trial-and-error approach. Using Parallel Genetic Algorithm (PGA) we have managed to determine the optimized weights and penalties for the function by running several simulations in a
very short time, which was not possible otherwise with a normal PC. In the later part of the study we have also developed a user interface in EXCEL to display the optimization results for different cases, with the weights and penalties as user input from EXCEL worksheet.
Chapter 1: Introduction and Literature Review

1.1 Introduction

Conjunctive management of limited water resources in fast growing semi-arid cities is crucial in order to sustain economic development. However, conjunctive management is increasingly challenging because growing population complicates the water distribution system in scale and operation. Finding cost-effective operating strategies is far from simple in a large-scale system. The challenges stem from several aspects; both the electricity tariff and water demand can vary greatly through a typical operating cycle; tank water have to be stored sufficiently to ensure reliability of the supply without deteriorating water quality; and the stressed aquifer system has to remain sustainable. These aspects render the facts that hydraulic behavior in a water distribution system is highly non-linear, interaction between the water distribution system and aquifer system is complex, and optimization problems for conjunctive management of the water distribution system and aquifer system is computationally challenging. Nevertheless, finding cost-effective operating strategies is more favored than expanding the existing water distribution system due to high construction cost and land constraints.

Fundamental concepts of system optimization and various optimization techniques have been extensively developed to multi-reservoir distribution systems in past forty years (Yeh 1985; Labadie 2004). Similar techniques for pressurized water distribution system optimization include linear programming (Crawley and Dandy 1993; Jowitt and Germanopoulos 1992; Burnell et al. 1993), non-linear programming (Chase and Ormsbee 1993; Yu et al. 1994; Sakarya and Mays 2000; Liu 2003), dynamic programming (Yeh et al. 1992; Lansey and Awumah 1994; Nitivattananon et al. 1996;
Marino and Mohammadi 1983); and mixed-integer nonlinear programming (Biscos et al. 2003). A favorable trend of using global search algorithms in solving water resources optimization problems is recognized in past ten years because most of global search methods are easy to be implemented to cope with highly nonlinear mixed integer objective functions and have a potential to reach to global or near global solutions. Moreover, hybrids of the global search methods with traditional gradient methods have been proven to be efficient to solve original optimization problems (Tu et al. 2005), which are commonly altered through linearization, relaxation, decomposition in order to be solved solely by gradient methods. The global methods include non-linear heuristic optimization (Ormsbee and Reddy 1995; Leon et al. 2000), e.g., simulated annealing (McCormick and Powell 2004; Goldman and Mays 2005), and genetic algorithms (Savic and Walters 1997).

Flow simulation in the pressurized water distribution system is much different from the multi-reservoir distribution system due to required energy conservation in the network. Without considering hydraulic constraints in the water distribution system, direct optimal controls on water quantities and flow directions in individual pipes are normally resulted (Sun et al. 1995; Tu et al. 2005). The time step is usually large and long-term system optimization is usually implemented when flows are only governed by the continuity equation. However, hydraulic constraints are necessary to ensure correct hydraulic-driven flow simulation in the pressurized water distribution system, which tremendously increase computation demand. The time step is normally in the unit of hour. Short-term to real-time simulation is commonly considered. The optimization in the pressurized water distribution system usually focuses on pumping energy cost reduction.
and water quality transport problems (Goldman and Mays 2005; Boccelli et al. 1998; Sakarya and Mays 2000; Constans et al. 2003).

With available modeling resources, many federal/state agencies and private sectors recognize the importance of utilizing simulation models as a decision-making tool to facilitate the development of water resources management strategic plans. Developing regional-specific water distribution model and groundwater model become crucial, especially for the city’s Master Plan. Although not initially for the research purpose, these models are so important to their communities such that they are developed in a full scale to represent the reality. For example, the water distribution model normally includes hundreds of thousands of junctions and pipes. Integration of the existing simulation models under the management framework become the next necessary step.

This study focuses on the production well management model development and uses the water resources system of City of Chandler, Arizona as a real-world case study. Production well operations optimization has been extensively investigated in the groundwater management aspect. However, the joint consideration on both water distribution system and aquifer system to determine optimal well pump operations is rarely discussed due to the complexity in the aforementioned hydraulic constraints. The actual well pumping rate and efficiency are governed by the dynamic groundwater head and the network system. Moreover, the integration of the real-world large-scale water distribution model with three-dimensional groundwater flow model through production wells is seldom conducted in literature. The computation for such a large simulation-optimization problem is challenging.
In this study, the Chandler’s large-scale pressurized water distribution model and three-dimensional groundwater flow model will be integrated under the optimization framework. Chandler is one of the fastest growing cities in the U.S. Management of water resources in Chandler is crucial due to its geographic location in the semi-arid area. Groundwater currently supplies 8% of city’s water demands through 23 production wells; however, groundwater has a “phase-in allowance” to produce 80% surface water shortfall from Salt River Project (SRP) and Central Arizona Project (CPA) during a drought. Potentially, groundwater will be a crucial water resource to Chandler at its build-out population in the next 15 years. From the continuing economic development standpoint, the Chandler water distribution system and groundwater system have to be integrated as a whole to manage its water resources.

The proposed well pump management model incorporates multiple objectives, where the well pump energy cost minimization and pressure violation minimization are two overriding competing objectives. Multi-objective optimization problems in water resources have been discussed by Vamvakeridou-Lyroudia et al. 2006, Kapelan et al. 2006, Farmani et al. 2006, and Cheung et al. 2003. The weighting method and the constraint method are commonly used to the multi-objective optimization problem. This study adopts the weighting coefficient method and develops a Pareto curve with respect to energy cost and pressure violation to obtain the non-inferior solutions. A parallel genetic algorithm using Message Passing Interface (MPI) was developed to solve the optimization problem over a cluster of computers, which makes the management model computationally feasible to be implemented for real-time operations.
1.2 Literature Review

1.2.1 Water Resources Management Models and Optimization

In the past few decades there has been a major development in the field of optimization of water resources management models. Several new techniques and methods have been evolved and tested over time to produce satisfactory results. Water resources management problems have been always very complex and large for simple and straightforward method e.g. linear programming; hence there has been a constant need for a better, faster and accurate method. Nonlinear programming, dynamic programming, simulated annealing, nonlinear integer programming, genetic algorithm and many other techniques have evolved with time.

1.2.1.1 Non-Linear Programming Approach

An optimization approach which has been highly adopted for multi-reservoir and multi-source water distribution systems is the non-linear programming approach. Construction of the objective function and constraints has been a common methodology for such techniques. In the area of water resources planning this has been a successful attempt to get rid of ultra complex problems with high dimensions. In recent years the water utility industry has begun to investigate the use and integration of on-line computers and control technology in improving the daily operation of water distribution systems. This has been motivated by the desire to reduce operational costs and provide more reliable operations and one of the greatest potential areas for cost savings is in the scheduling of daily pump operations. Using non-linear programming a perfect model can be easily developed for obtaining least cost pump operation policy for a multi-source, multi-tank water distribution system. The algorithm links a minimum-cost-constraint
identification methodology with a network simulation model. The algorithm has the advantage of being computationally efficient while incorporating the non-linear characteristics of the water distribution network. In addition, the algorithm has the advantage of providing several feasible solutions to the control problem, which then provides the system operator with increased flexibility.

The objective of the optimal pump operation is to minimize the energy cost while satisfying the hydraulic operational requirements of the system. For most water utilities, the pumping cost is composed of an energy-consumption charge and demand charge. In a typical water distribution system, the energy-consumption cost incurred by the pumping facility depends on the rate at which the water is pumped, the associated pump head, the duration of pumping, and the unit cost of electricity. Mathematically the objective function may be expressed as:

Minimize \( Z = \sum_{t=1}^{T} \sum_{i=1}^{I} \frac{\gamma O_i H_i}{e_i} X_{it} r_t \)

Where \( Z \) is the total cost to be minimized; \( T \) is the number of time intervals which contribute to the operating horizon; \( i \) is the number of pumps in the system; \( Q \) is the average flow rate associated with the pump; \( H \) is the head associated with the pump; \( X \) is the duration of the time a particular pump is operating for a time \( t \); \( r \) is the electric rate during time \( t \) and \( e \) is the average wire to water efficiency associated with pump \( t \) and time \( t \). The objective function expressed above has three different kinds of constraints

1. A set of implicit system constraints: This type of constraint deals with the conservation of energy and conservation of mass relationships.
a. The nodal conservation of mass requires that the sum of the flows into or out of any junction node minus any external demand must be equal to zero.

b. The conservation of energy constraint requires the sum of the line losses and the minor losses over any distinct pipe path or loop minus any energy added into the liquid by the pump, minus the difference in the energy grade between the end points of the paths is equal to zero.

2. A set of implicit bound constraints: These types of constraints in the problem may include constraints on the nodal pressure, pipe flow rate or velocities, and tank water levels.

3. A set of explicit bounds constraints: These set of bound constraints consists of explicit bounds on the decision variables. For example in this case the decision variable for each pump station for a particular time interval will be restricted between a lower value of zero and an upper value equal to the maximum number of pump combinations available for that pump station.

Previous attempts to develop optimal control algorithms for water distribution systems have typically focused on the development and use of implicit control formulations in which the problem is expressed in terms of an implicit state variable such as tank level or pump station discharge. Such formulations suffer from the requirement of two-step optimization methodology in which the actual pump operating policies must be extracted from the solution of the implicit control problem. Attempts to circumvent this problem by use of explicit formulations in which pump run times are treated as the decision variables are limited due to the number of decision variables that can be
effectively considered. In the approach presented above the limitation is minimized by rank ordering different pump combinations and developing a single decision variable for each pump station for each control interval.

1.2.1.2 Simulated Annealing

The operation of water distribution systems affects the water quality in these systems. EPA regulations require that water quality be maintained at all points in the system including the point of delivery. Methods to optimize water system operations have been restricted to reducing costs related to pumping and costs related to sizing, construction, and/or maintenance of piping while meeting customer demands, pressure limits, and tank operation restrictions. There have been few attempts to optimize water system operations for both hydraulic and water quality performance and they have been restricted to simplified systems.

A new methodology that formulates the water distribution system problem as a discrete time optimal control problem was developed that linked the method of simulated annealing with EPANET for optimal operation of water distribution systems for both water quality and hydraulic performance. Most optimization techniques require the calculation of derivatives, response functions, or other methods that are limited to specific problems. Simulated annealing allows optimization for a variety of objective functions and can consider many modifications to operational conditions without reprogramming of the optimization procedure.

A water distribution system has two major concerns that govern its optimal operation, water quality and efficient pumping. Water quality can undergo significant changes as it travels through the water distribution system from the point of supply and/or
treatment to the point of delivery. For example, chlorine concentration will decrease with time in pipes and tanks through bulk decay and through reaction with the pipe wall. Also, chlorine mass leaves the system at demand nodes or can be added to the system at chlorine booster nodes. Mixing of water from different sources can also affect water quality (Clark et al., 1995).

Pumps are usually operated according to an operating policy, which includes the scheduling of pump operation, which can affect water quality if the system pumps draw water from sources with differing water quality. Pump operation will affect the turnover rate of storage tanks. Chlorine dosage may differ at each pump depending on the water source quality. For example, consider a pump that operates for 6 h divided into 1-h periods. The pump can either be on or off for any period. The number of combinations is $2^6 = 64$. A pump, which operates for 24 h divided into 1-h periods, has $2^{24} = 16,777,216$ combinations. The 6-h example can be solved by trial and error but the 24-h example is prohibitively large to solve by trial and error. For large combinatorial optimization problems, simulated annealing provides a manageable solution strategy. These considerations have generated a need for computer models and optimization techniques that consider water quality in the water distribution system as well as system hydraulics.

Simulated annealing is a combinatorial optimization method that uses the Metropolis algorithm to evaluate the acceptability of alternate arrangements and slowly converge to an optimum solution. The method does not require derivatives and has the flexibility to consider many different objective functions and constraints. Simulated annealing uses concepts from statistical thermodynamics and applies them to combinatorial optimization problems. Metropolis algorithm was developed to provide the
simulation of a system of atoms at a high temperature that slowly cools to its ground
energy state. If an atom is given a small random change there will be a change in system
energy $E$. If $E < 0$, the new configuration is accepted. If $E > 0$ or $E = 0$, then the decision to
change the system configuration is treated probabilistically. A random number evenly
distributed between 0 and 1 is chosen. If the number is smaller than the probability then
the new configuration is accepted; otherwise it is discarded and the old configuration is
used to generate the next arrangement.

The requirements for applying simulated annealing to an engineering problem are:

(1) A concise representation of the configuration of the decision variables,
(2) A scalar cost function,
(3) A procedure for generating rearrangements of the system,
(4) A control parameter and an annealing schedule, and
(5) A criterion for termination.

1.2.2 Water Distribution System Optimization

Ormsbee and Lansey (1994) reviewed methods used to optimize the operation of
water supply pumping systems to minimize operation costs. They developed multi-
objective genetic algorithms for pump scheduling in water supply systems. The pump
scheduling problem has been formulated as a cost optimizations problem which aims to
minimize marginal costs of supplying water, whilst keeping within physical and
operational constraints (e.g. maintain sufficient water within the system's reservoirs to
meet the required time varying consumer demands). To achieve the aforementioned aim
of pump scheduling, an algorithm has been devised to determine which of the pumps
available within the system, should be used during which interval of the optimization period.

Ostfeld and Shamir (1993) developed a water quality optimization model that optimized pumping costs and water quality using steady-state and dynamic conditions using GAMS/MINOS [General Algebraic Modeling System; Mathematical In-Core Nonlinear Optimization Systems]. The objective was to minimize total cost, which included the cost of water at the sources, of treatment, and of the energy to operate the system. The constraints included equations that describe the change in flow and quality over time throughout the system, the physical laws of flow and concentrations, and the requirements for level of service. The equations that describe concentrations in pipes are of a form that allows the flow direction to reverse during the iterative solution process.

Linear Programming approach in water resources system optimization has been thoroughly discussed by Jowitt and Germanopoulos (1992) in their research focused on optimal pump scheduling in water-supply networks There study presents a method based on linear programming for determining an optimal (minimum cost) schedule of pumping on a 24-hr basis. Both unit and maximum demand electricity charges are considered. Account is taken of the relative efficiencies of the available pumps, the structure of the electricity tariff, the consumer-demand profile, and the hydraulic characteristics and operational constraints of the network. The use of extended-period simulation of the network operation in determining the parameters of the linearized network equations and constraints and in studying the optimized network operation is described. The method was found to be robust and with low computation-time requirements, and is therefore suitable for real-time implementation.
Ormsbee (1991) suggested that on–off operation of pumps poses a difficulty for optimization techniques that require continuous functions. Computer-based water quality models exhibit the same difficulties. Most modeling methods include numerical methods rather than continuous functions. Linear superposition was used by Boccelli et al. (1998) for the optimal scheduling of booster disinfection stations by developing system-dependent discretized impulse response coefficients using EPANET and linking with the LP solver MINOS (Murtagh and Saunders, 1987).

Another approach known as Dynamic programming has been used for the optimal scheduling of pumps in a water distribution system by Coulbeck and Orr (1982) and Coulbeck et al. (1987). The method is sensitive to the number of reservoir states and the number of pumps. In general, the increase in number of discretizations and state variables increases the size of the problem dramatically, which is known as the curse of dimensionality. This has restricted the application of this method to small systems.

The nonlinear programming approach by Brion et al. (1991) evaluates gradients using finite differences. Optimum design and operation refers to the selection of pump type, capacity, and number of units as well as scheduling the operation of pumps that result in minimum design and operating cost for a given set of demand curves. The design criteria for such pumping stations are based fundamentally on some important and critical parameters, such as pump capacity, number of units, types of pumps, and civil works. The optimization process consists of three main steps: (1) determination of minimum yearly consumed energy; (2) minimization of the total cost for all sets of pumping stations; and (3) selection of the least-cost set among the feasible sets of pumping stations, recognizing a combination of the cited criteria.
Sakarya (1998) and Mays et al. (2000) considered a mathematical approach which solves nonlinear optimization problems by using the generalized reduced gradient method for scheduling pump operations to minimize energy and improve water quality. This approach considers three objective functions:

1. The minimization of the deviations of actual concentrations from a desired concentration,
2. Minimization of total pump duration times, and
3. The minimization of total energy while meeting water quality constraints.

Zessler and Shamir (1989) presented an approach for optimal operation of water supply system by progressive optimality, an iterative dynamic programming (DP) method. Given the forecasted demands for the coming 24 hr, the initial and final conditions of the reservoirs, the hydraulic properties of all system components, and the variable energy cost over the day, an optimal schedule of the pump operation is found. The algorithm cycles iteratively over the time steps and network subsystems, and converges to the optimum from any initial solution.

Nitivattananon et al. (1996) developed an optimization model to generate pump schedules in real-time operation for a complex water supply system by considering major difficulties posed by a complicated tariff; especially demand charges, discrete pump discharges, and other physical constraints. The model decomposes in space and time the system into several subsystems, and planning periods into operational periods. Progressive optimality is applied to solve a dynamic programming model. The pump discharges are discretized and arranged by heuristic methods in order to reduce the number of times pumps are switched on. In real-time operation, billing monthly demand
is estimated by a long-term model and then incorporated in a short-term model to obtain an optimal daily schedule in a relatively short computational time.

The requirements and basic components of a typical optimal control environment for water-supply pumping systems are presented and discussed by Ormsbee and Lansey (1994). The model components include hydraulic network models, demand forecast models, and optimal control models. Examined methodologies are classified on the basis of the type of system to which the methodology can be applied (single source-single tank or multiple source-multiple tank), the type of hydraulic model used (mass balance, regression, or hydraulic simulation), the type of demand model used (distributed or proportional), the type of optimization method used (linear programming, dynamic programming, or nonlinear programming), and the nature of the resulting control policy (implicit or explicit). The applicability of current technology to an existing water-supply pumping system is examined in light of existing technical limitations and operator acceptance issues.

Biscos et al. (2003) presented an approach for the operational optimization of potable water distribution networks. The maximization of the use of low-cost power (e.g. overnight pumping) and the maintenance of a target chlorine concentration at final delivery points were defined as important optimization objectives. The first objective is constrained by the maintenance of minimum emergency volumes in all reservoirs, while the second objective would include the minimization of chlorine dosage and re-dosage requirements. The combination of dynamic elements (e.g. reservoirs) and discrete elements (pumps, valves, routing) makes this a challenging predictive control and
constrained optimization problem, which is being solved by MINLP (Mixed Integer Non-linear Programming).

1.2.3 Reliability, Resilience and Vulnerability of Water Resources Systems

Loucks et al. (1982) have discussed the three criteria of reliability, resilience and vulnerability for evaluating the performance of water resources system. These measurements describe how likely a system is to fail (reliability), how quickly it recovers (resiliency) and how severe the consequences of the failure may be (vulnerability). These criteria can be used to assist in the evaluation and selection of effective design and operating policies for a wide variety of water resources projects. Later reservoirs operation reliability, resilience and vulnerability have been discussed by ReVelle et al. (1986).

1.2.4 Genetic Algorithm for Water Distribution Systems Optimization

There have been several attempts in recent years to develop optimal control algorithms to assist in the operation of complex water distribution systems. The various algorithms were oriented towards determining least-cost pump scheduling policies (proper on-off pump operation) and were based on the use of linear programming, nonlinear programming, dynamic programming, enumeration techniques, and general heuristics. However, the success of these procedures has been very limited and very few have actually been applied to real water distribution systems. Limited acceptance of optimal control models in engineering practice is partly because:

1. Such techniques are generally quite complex involving a considerable amount of mathematical sophistication (e.g., requiring extensive expertise in systems analysis and careful setting up and fine tuning of parameters).
2. They are generally highly dependent upon the number of pumps and storage tanks being considered along with the duration of the operating period.

3. They are generally subject to oversimplification of the network model and its components along with several simplifying assumptions to accommodate the nonlinear network hydraulics.

4. They tend to be extremely time-consuming resulting in added costs and inefficient use of the computer.

5. They may be easily trapped at local optima and may not lead to the global optimal solution. Another very important reason for their lack of acceptance and use was the unavailability of suitable and user-friendly pump optimization packages. As a result, most optimal control models developed to date have been mainly used as a research support tool.

Genetic algorithms are an adaptation procedure based on the mechanics of natural genetics and natural selection. They are designed to perform search procedures of an artificial system by emulating the evolution process (Darwin’s evolutionary principle) observed in nature and biological organisms. The evolution process is based on the preferential survival and reproduction of the fittest member of a population with direct inheritance of genetic information from parents to offspring and the occasional mutation of genes. The principal advantage of GAs is their inherent ability to intelligently explore the solution space from many different points simultaneously enabling higher probability for locating global optimum without having to analyze all possible solutions available and without requiring derivatives (or numerical approximations) or other auxiliary knowledge.
Zyl et al. (1995) presented an approach for application of genetic algorithms to pump scheduling for water supply. In their work a simple genetic algorithm has been applied to the scheduling of multiple pumping units in a water supply system with the objective of minimizing the overall cost of the pumping operation, taking advantage of storage capacity in the system and the availability of off-peak electricity tariffs. Again in 2004 Savic et al. have used genetic algorithm for optimal operation of water distribution system. According to their approach genetic algorithm (GA) optimization is well suited for optimizing the operation of water distribution systems, especially large and complex systems.

GAs have good initial convergence characteristics, but slow down considerably once the region of optimal solution has been identified. In their study the efficiency of GA operational optimization was improved through a hybrid method which combines the GA method with a hillclimber search strategy. Hillclimber strategies complement GAs by being efficient in finding a local optimum. Two hillclimber strategies, the Hooke and Jeeves and Fibonacci methods, were investigated. The hybrid method proved to be superior to the pure GA in finding a good solution quickly, both when applied to a test problem and to a large existing water distribution system.

Tu et al. (2003) further modified the use of genetic algorithm by developing a multi-commodity flow model to optimize water distribution and water quality in a regional water supply system. The optimization model was highly nonlinear and solved by a hybrid genetic algorithm (GA). Then a generalized reduced gradient (GRG) algorithm embedded in the GA was used to optimize the objective function for fitness evaluation. The proposed methodology was first tested and verified on a hypothetical
system and then applied to the regional water distribution system of the Metropolitan Water District of Southern California. The results obtained indicate that the proposed hybrid GA is a viable way of converting an undirected network to a directed network by separating the complicating variables, and that the resulting directed network model can be solved iteratively and efficiently by a gradient-based algorithm.

A new approach for reliability-based optimization of water distribution networks was presented by Simpson et al. (2004). The approach links a genetic algorithm (GA) as the optimization tool with the first-order reliability method (FORM) for estimating network capacity reliability. Network capacity reliability in this case study refers to the probability of meeting minimum allowable pressure constraints across the network under uncertain nodal demands and uncertain pipe roughness conditions. This research demonstrates the novel combination of a GA with FORM as an effective approach for reliability-based optimization of water distribution networks. Correlations between random variables are shown to significantly increase optimal network costs.

Zhang et al. (2005) have presented an approach for optimal operation of water supply systems with tanks based on genetic algorithm in which pump-station pressure head and initial tank water levels were treated as decision variables and the model of optimal allocation of water supply between pump-sources was developed. Genetic algorithm was introduced to deal with the model of optimal allocation of water supply. Methods for handling each constraint condition were put forward, and overcome the shortcoming such as premature convergence of genetic algorithm; a solving method was brought forward in which genetic algorithm was combined with simulated annealing technology and self-adaptive crossover and mutation probabilities were adopted.
1.2.5 Genetic Algorithm and Linear Programming for Water Resources Management Models

Genetic algorithm and linear programming hybrid approach for water management models have been precisely presented by Cai et al. (2001). Gradient-based nonlinear programming (NLP) methods can solve problems with smooth nonlinear objectives and constraints. However, in large and highly nonlinear models, these algorithms can fail to find feasible solutions, or converge to local solutions which are not global. Evolutionary search procedures in general and genetic algorithms (GAs) specifically, are less susceptible to the presence of local solutions. However, they often exhibit slow convergence, especially when there are many variables, and have problems finding feasible solutions in constrained problems with “narrow” feasible regions. For operating on such large nonlinear water resources models a combined approach involving GAs with linear programming has been adopted. The key idea is to identify a set of complicating variables in the model which, when fixed, render the problem linear in the remaining variables. The complicating variables are then varied by a GA and a linear problem is then solved to compute the optimal objective value for each set of fixed values suggested by GA.

1.2.6 Application of Genetic Algorithm in Groundwater Management Models

The design and management of groundwater systems is an important problem in both groundwater resource development for water supply and design of remedial actions for cleaning up contaminated aquifers. Many of these problems involve the solution of mathematical problems which have highly nonlinear and frequently discontinuous objective functions and constraint set. One of the feasible approaches is by combining
groundwater simulation models with genetic algorithm to help search for optimal groundwater system designs.

Groundwater management models are often characterized as non-convex, nonlinear programming problems. Such as there is no guarantee that a globally optimal solution will be found using traditional gradient based methods such as nonlinear programming. In the absence of an optimization tool to find the global optimum, GAs provide a good alternative method. In addition, GAs are capable of finding several near-optimal design alternatives that are different from each other and can be further analyzed or compared.

Nonlinear programming techniques have been used to solve groundwater management problems for the past decade. These methods employ gradient based algorithms to adjust decision variables so as to optimize the objective function of a management model. These algorithms require the computation of sensitivities of state variables, e.g. head or concentration, at certain location to decision variables, e.g. pumping rates, at other locations. Sensitivities can be obtained by either the adjoint sensitivity or perturbation methods. These sensitivities are difficult to program, in the case of adjoint sensitivity method, or computationally expensive to generate, in the case of perturbation methods, and in general, are not robust.

Further more, the cost functions of typical groundwater system components may be either discontinuous, e.g. well field capitol costs, or highly complicated, e.g., treatment process costs, making it difficult to calculate or estimate the derivatives of these functions with respect to the decision variables. Groundwater management models tend to highly nonlinear and non-convex mathematical programming problems especially in the case of
aquifer remediation design with mass transport constraints. As such there is no guarantee that a global optimum of a groundwater remediation design model will be found by nonlinear programming methods. Also these methods tend to be non-robust and require excessive computational resources.

McKinney and Lin (1994) have presented a comprehensive study in their research titled genetic algorithm solutions of groundwater management models. Various groundwater simulation models have been incorporated into a genetic algorithm to solve management problems, such as maximizing the pumping from an aquifer; minimum cost water supply development; and minimum cost aquifer remediation. The results have shown that genetic algorithm can effectively and efficiently be used to obtain globally optimal solutions to these groundwater management problems. The formulation of the method is straightforward and provides solutions which are as good as or better than those obtained by linear and nonlinear programming. Constraints are incorporated in the formulation and they do not require derivatives with respect to decision variables as in nonlinear programming. More complicated problems such as transient pumping and multiphase remediation have also been formulated and solved successfully. The computational time does increase with the complexity of the problem. To speed up the computation parallel computers have been used.

In 1998 Wang et al. presented an approach in which Genetic algorithms (GA) and simulated annealing (SA), two global search techniques, were coupled with MODFLOW, a commonly used groundwater flow simulation code, for optimal management of groundwater resources under general conditions. The coupled simulation-optimization models allow for multiple management periods in which optimal pumping rates vary with time to
reflect the changing flow conditions. The objective functions of the management models are of a very general nature, incorporating multiple cost terms such as the drilling cost, the installation cost, and the pumping cost. The models were first applied to two-dimensional maximum yield and minimum cost water supply problems with a single management period, and then to a multiple management period problem. The strengths and limitations of the GA and SA based models are evaluated by comparing the results with those obtained using linear programming, nonlinear programming, and differential dynamic programming.
Chapter 2: Simulation Models

2.1 Water Distribution System Model: EPANET

2.1.1 Introduction

EPANET is a computer program that performs extended period simulation of hydraulic and water quality behavior within pressurized pipe networks. A network can consist of pipes, nodes (pipe junctions), pumps, valves and storage tanks or reservoirs. EPANET tracks the flow of water in each pipe, the pressure at each node, the height of water in each tank, and the concentration of a substance throughout the network during a multi-time period simulation. In addition to substance concentrations, water age and source tracing can also be simulated. EPANET is designed to be a research tool for improving our understanding of the movement and fate of drinking water constituents within distribution systems. The water quality module of EPANET is equipped to model such phenomena as reactions within the bulk flow, reactions at the pipe wall, and mass transport between the bulk flow and pipe wall. Another feature of EPANET is its coordinated approach to modeling network hydraulics and water quality. The program can compute a simultaneous solution for both conditions together. Alternatively it can compute only network hydraulics and save these results to a file, or use a previously saved hydraulics file to drive a water quality simulation.

2.1.2 Hydraulic Modeling Capabilities of EPANET

Full-featured and accurate hydraulic modeling is a prerequisite for doing effective water quality modeling. EPANET contains a state-of-the-art hydraulic analysis engine that includes the following capabilities:

1. Places no limit on the size of the network that can be analyzed
2. Computes friction head-loss using the Hazen-Williams, Darcy-Weisbach, or Chezy-Manning formulas
3. Includes minor head losses for bends, fittings etc.
4. Models constant or variable speed pumps
5. Computes pumping energy and cost
6. Models various types of valves including shutoff, check, pressure regulating, and flow control valves
7. Allows storage tanks to have any shape (i.e., diameter can vary with height)
8. Considers multiple demand categories at nodes, each with its own pattern of time variation
9. Models pressure-dependent flow issuing from emitters (sprinkler heads) base system operation on both simple tank level or timer controls and on complex rule-based controls.

2.1.3 Chandler Water Resource System

Chandler is one of the fastest growing communities in the nation among cities with population over 150,000. Its history has seen a transformation from a small agricultural town at the turn of the 20th century to the high-tech oasis in the Silicon Desert of today. Chandler offers excellent quality of life amenities, superior schools, a rapidly expanding health care system, and a reputation as a global leader in technology. Location is a major factor in Chandler's prosperity, offering many advantages to existing business and industry and great opportunity for new business. Chandler has an expanding high technology base, with over 75 percent of its manufacturing employees in high
technology fields. The City continues to prosper from the availability of excellent regional and national transportation, economic climate, and quality of life.

The City of Chandler expects to have most residential areas built out in the next ten years, and the rest of the City may be built out within the next fifteen to twenty years. Therefore, the City required an accurate projection of water demands and sewer flows in order to plan capital improvements that will be needed over the next twenty years. Water resources together with infrastructure are essential elements in supporting the residential, commercial, and industrial activities of the City. The Water Resources portfolio and infrastructure, including potable and reclaimed water, provides residents of the City with an adequate supply of safe, dependable water to support the activities of the citizens and to allow for the continued, orderly, planned growth of the City.

2.1.4 Chandler Water Distribution Model

The Chandler city has created a hydraulic model for the water distribution system in the city. EPANET has been chosen for the modeling and simulation purposes. Figure 2.1 shows the Chandler water distribution system. The network consists of 11392 junctions with positive demands. The demand value varies with time and region. Each junction has a fixed elevation and base demand. Base demand is the average or nominal demand for water by the main category of consumer at the junction, as measured in flow unit. Most of the junctions have been assigned a fixed weekly pattern, which multiply with the base demand of the junction to decide the required value of the net demand for that time period. Many junctions serve as connectivity points i.e. they have zero demands. The maximum and minimum base demand values are around 58 and 0 GPM respectively. The minimum and maximum values of the elevations are 1155 and 1290
feet respectively. The elevation is some common point above reference and is used for calculation of pressure at the junction. It does not affect any other computed quantity.

**Figure 2.1:** Chandler water distribution system

There are 14258 pipes in the network. Each pipe has a unique label and a start node and an end node, which defines the direction of flow. Pipe diameters range from 6 to 60 inches depending upon the connection points. The roughness is unit less for Hazen-Williams or Chezy-Manning roughness and has unit of millifeet for Darcy-Weisbach roughness. A unit less loss coefficient has been assigned to each pipe for bends, fittings, joints etc. The network consists of around 123 pumps. Out of these, 23 are well pumps.
and the remaining are booster pumps. Wells pumps are the pumps which are directly connected to the reservoirs. Among these there are 13 pumps which are connected to the reservoirs on one side and to storage tanks on the other, where as the remaining 10 pumps directly connect the reservoirs to the pressurized network i.e. junctions. The remaining 100 pumps are the booster pumps which supply water from the storage tanks into the pressurized network. Each pump has a pump curve which describes the relationship between the head delivered by the pump and the flow though it. The speed setting of a pump is the factor which gets multiplied with the normal speed setting and produces a value, eg. 1.2 speed setting would result in a 20% higher speed of the pump. Each pump is also associated with a pattern values. These are actually equivalent to the speed setting which means that a 0 would represent no pumping. The patterns are weekly schedules with a tie period of one hour. The efficiency curve represents the pump’s wire to water efficiency as a function of flow rate. These are used for energy calculations. The average or nominal price of energy in monetary units i.e. per-kW-hr is 0.0599. Figure 2.2 shows water distribution network details for the Chandler city.

There are 24 tanks in the system out of which 13 tanks are directly connected to the well pumps one side and to booster pumps on the other. The remaining 11 tanks are directly connected to the booster pumps and valves, primarily used for storage and water supply. Each tank has a fixed elevation, minimum and maximum level. The water does not exceed the minimum and maximum limits of the tank level. The diameter of the tanks is in feet with a maximum value of 230 and minimum of 73. Similarly there is a minimum value of volume for the tanks below which the tanks do not operate. Tanks volume curves represents a relationship between the tank volume and its level. The type
of water mixing in the tank is assumed to be completely mixed type. The system has 32 reservoirs. Each reservoir has a hydraulic head which is the sum of the pressure head and the elevation head. In this system the reservoirs represents under groundwater head.

**Figure 2.2: Water distribution network details**

The system has three types of valves flow control valves (FCVs), pressure reducing valves (PRVs), and pressure sustaining valves (PSVs). The flow setting for the FCVs is around 3000 gpm and the pressure setting for the PSVs and PRVs is around 70 psi. Most of the FCVs have a diameter around 12 inches except two which have different flow settings; similarly the PRVs have a diameter around 16 inches and PSVs with 12 inches. The unit less minor loss coefficient that applies to the valves when they are fully open is 4.0 for all the types of valves. Various link and nodes use weekly, daily and
hourly patterns, which have been assigned to them. The system has around 90 patterns of different intervals. Figure 2.3 shows an EPANET weekly pattern “PATN1”.

<table>
<thead>
<tr>
<th>Pat ID</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PATN1</td>
<td>July 2005</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiplier</td>
<td>0.71</td>
<td>0.71</td>
<td>0.72</td>
<td>0.74</td>
<td>0.93</td>
<td>1.17</td>
<td>1.24</td>
<td>1.3</td>
</tr>
</tbody>
</table>

**Figure 2.3**: An EPANET weekly pattern “PATN1”

### 2.2 Groundwater Management Model: MODFLOW

#### 2.2.1 Introduction

MODFLOW is a computer program that numerically solves the three-dimensional groundwater flow equation for a porous medium by using a finite-difference method.

#### 2.2.2 Groundwater Flow Equation

The partial-differential equation of ground-water flow used in MODFLOW is

\[
\frac{\partial}{\partial x} \left( K_{xx} \frac{\partial h}{\partial x} \right) + \frac{\partial}{\partial y} \left( K_{yy} \frac{\partial h}{\partial y} \right) + \frac{\partial}{\partial z} \left( K_{zz} \frac{\partial h}{\partial z} \right) + W = S_s \frac{\partial h}{\partial t} \tag{2.1}
\]

Kxx, Kyy, and Kzz are values of hydraulic conductivity along the x, y, and z coordinate axes, which are assumed to be parallel to the major axes of hydraulic conductivity (L/T); h is the potentiometric head (L); W is a volumetric flux per unit volume representing sources and/or sinks of water, with W<0.0 for flow out of the ground-water system, and W>0.0 for flow in (T⁻¹); SS is the specific storage of the
porous material \((L^{-1})\); and \(t\) is time \((T)\). The equation 2.1, when combined with boundary and initial conditions, describes transient three-dimensional ground-water flow in a heterogeneous and anisotropic medium, provided that the principal axes of hydraulic conductivity are aligned with the coordinate directions. The Ground-Water Flow Process solves equation 2.1 using the finite-difference method.

Boundary conditions

1. Specified head boundaries (Dirichlet conditions): Hydraulic head is given for the boundary.

2. Specified flow boundaries (Neumann conditions): Flux (derivation of head) across the boundary is given (A no-flow boundary has a flux of zero).

3. Head-dependent flow boundaries (Cauchy or mixed conditions): Flux is dependent on the hydraulic head.

Packages included in MODFLOW for various uses

1. The Basic package (BAS)

2. The Block Centered Flow package (BCF)

Head dependent packages

3. River package (RIV)

4. Drainage package (DRN)

5. General Head Boundary package (GHB)

6. Evapotranspiration package (EVT)

Head-independent packages

7. Well package (WEL)

8. Recharge package (RCH)
Solver packages

9. Strongly Implicit Procedure package (SIP)
10. Successive Over Relaxation package (SOR)
11. Preconditioned Conjugate Gradient package (PCG)

The DRN, RIV, GHB, and EVT Packages are head dependent because the flow rate to or from the boundary is dependent on the head value within the adjacent model cell. The WEL and RCH Packages are considered head-independent packages because the user specifies the flow rate to or from the boundary.

2.2.3 Chandler Groundwater Model

Chandler is located in the east Salt River sub-basin of Phoenix. Studies of the hydrogeology of the east Salt River Valley have identified three major aquifer units: Upper Alluvial Unit (UAU), the Middle Alluvial Unit (MAU) and Lower Alluvial Unit (LAU) (Laney and Hahn, 1986). The definition of these three units is based on age and lithology of the sediments that comprise the Salt River alluvial basin. Beneath the Chandler water planning area all the three aquifer units are present and yield water to wells. However the Upper Alluvial Unit is unsaturated over much of the area to the east of the Chandler. Prior to 1970, water levels within the Chandler planning area were generally declining. However, since then water levels have steadily risen as the pumping, especially agricultural pumping, has declined.

2.2.4 Chandler MODFLOW Model

The Chandler planning area has been divided into 71 rows and 98 columns and it has three layers as already described in the previous section. The major units for calculation are days and feet for time and length. The total duration of simulation run is
one week and with one stress period and 7 time steps. The simulation is carried out in transient conditions. The IBOUND values for the three Chandler aquifer units are either 0s or 1s, i.e. either inactive or inactive. There are no constant head cells. All the three layers look quite identical with respect to the IBOUND distribution. In evapotranspiration the rate is calculated from the top grid layer only. The starting heads and hydraulic conductivities have been plotted. There are around 3500 wells (pumping and recharge) in the Chandler planning area. The wells pumps water from any of the three layers i.e. many of them have more than one screening lengths. The planning area also has around 125 drains. The groundwater model has been calibrated with the measured groundwater head data for the period between 1989 and 1998 (Southwest Ground-water Consultants, Inc. 2003, unpublished report).
Chapter 3: Genetic Algorithm

3.1 Introduction

There are myriad search and optimization techniques for optimization problems in the world. Researchers in engineering, economics, political science, psychology, linguistics, immunology, biology, and computer science need an efficient tool to tackle their optimization problems. It is difficult, however, to model realistic systems because the behavior of the systems is complex. In general, an optimization problem to be addressed has several objectives to be optimized. Thus, the complexity of the problem increases as the number of objectives increases because the objectives considered are often contradictory to one another. Such complex optimization problems have a lot of feasible solutions. However, only a few solutions among them are desirable.

In order to use an optimization technique for such complex optimization problems without difficulties, the technique should be robust. Goldberg defined robustness as “the balance between efficiency and efficacy necessary for survival in many different environment.” Then we can define two purposes in constructing an optimization technique as its efficacy and efficiency. Efficacy means whether the optimization technique can reach the optimum or not. The common purpose in constructing optimization techniques is this efficacy, that is, their convergence to the optimum of the problem. The other purpose, efficiency, means whether the technique can find a better solution under the constraints the problem has. The technique may not find the optimal solution of the problem due to the constraints, but it is important that better solutions are searched by the algorithm within the constraints. From this point of view, all search techniques are not robust because some search technique tends to find only the local
optimum due to its local scope, depends on existence of derivatives, or requires enormous computation time. Therefore Goldberg concluded that “the most important goal of optimization is improvement. Attainment of the optimum is much less important for complex systems.” As for complex systems, Zadeh (1988) also said “most realistic problems tend to be complex, and many complex problems are either algorithmically unsolvable or, if solvable in principle, are computationally infeasible.” Thus, robust algorithms which can find better solutions under a lot of constraints are required for optimizing complex systems.

The central theme of research on genetic algorithms (GAs) has been robustness. Genetic algorithms, first specified by John Holland in the early 1970’s are becoming an important tool for combinatorial optimization, function optimization, and machine learning. GAs are a kind of (i) stochastic search, (ii) multi-point search, (iii) direct search, and (iv) parallel search. These characteristic features of GAs contribute robustness of the algorithms. While it is easy to apply GAs to optimization problems, several researchers pointed out that the performance of GAs on some combinatorial optimization problems was a bit inferior to that of neighborhood search algorithms (e.g., local search, simulated annealing, and tabu search). Therefore hybridization of GAs with other heuristic methods is required for improving the performance of GAs. Genetic algorithms have been mainly applied to single-objective optimization problems. In order to handle multi-objective optimization problems, the objective functions should be combined into a scalar fitness function. But the characteristic features of GAs can be utilized for the search in the feasible region of multi-objective optimization problems. In general, optimization problems to be addressed have several objectives to be optimized. As the number of
objectives of the problem increases, the complexity of the problem becomes high because the objectives considered are often contradictory to one another. The researcher who tackles an optimization problem with multiple objectives needs a tool for optimizing their problem. Because objectives to be optimized in the problem are often contradictory to one another, the optimal solution of the problem is not obtained as a single solution. That is, a set of candidate solutions called non-dominated solutions is to be obtained for the problem. Since multiple solutions are to be obtained as candidate solutions of the problem, GAs as a kind of multi-point search potentially have an advantage for optimization problems with multiple objectives.

3.2 Genetic Algorithm for Single-Objective Optimization Problems

In this section, first we consider a simple genetic algorithm. In order to apply GAs to an optimization problem, each solution of the problem to be searched by GAs should be encoded as a finite-length string over some finite alphabet. We briefly describe the difference between the permutation coding and the binary coding. Next, genetic operators such as selection, crossover, mutation, and elitist strategy are described to construct GAs for optimization problems. These genetic operators should be carefully designed according to the property of the problems. The genetic operators for permutation strings are different from those for binary strings. Before applying GAs to optimization problems, several parameters such as population size, crossover probability, and mutation probability should be specified. After all the genetic operators and the parameters are specified for constructing GAs, we can apply GAs to the optimization problem.
3.2.1 Coding

In GAs, each solution of an optimization problem should be encoded as a finite-length string over some alphabet. The coding techniques can be categorized into the following two methods: a binary coding and a permutation coding. The binary representation is usually used for the coding of solutions. For example, the binary coding is often used for function optimization problems. In such problems, an input parameter vector $x$ on a constraint interval vector $[a, b]$ is encoded by the binary representation. The parameter vector $\hat{x}$ which optimizes a given function $f(x)$ is searched by GAs in the binary coding space. The other coding scheme, the permutation coding, is used for sequencing problems such as scheduling problems. For those problems, permutation strings of a set of numbers are more natural representation than binary strings.

![Figure 3.1: Genotype and Phenotype representation](image)

Figure 3.1 shows examples of strings by the binary coding and by the permutation coding. The string by the binary coding consists of “0” and “1”. The binary string treated in GAs is often decoded to the parameter value in integer, real number, and so on. The permutation string consists of numerals “1” to “$n$” where each numeral corresponds to a job in scheduling problems or to a city in traveling salesman problems, and $n$ is the total number of jobs or cities. Then jobs are processed according to their order in the
permutation when scheduling problems are considered, or cities are visited according to their order in the permutation in traveling salesman problems.

As shown in Figure 3.1, strings which consist of binary or numeral elements are called genotype, and solutions which are decoded from strings are called phenotype. GAs search over the genotype world and strings which are obtained by GAs are decoded into solutions in the phenotype world. That is, the users of GAs can get solutions of their optimization problems after the strings obtained by GAs are decoded into the solutions in the phenotype world.

### 3.2.2 Evaluation

Each of solutions which are decoded from the strings obtained by GAs is evaluated for optimization problems. GAs search a string with a better fitness value in the genotype world. In the case of function optimization problems, the function value $f(x)$ is calculated using a solution $x$ decoded from the corresponding binary string obtained by GAs. When the function value $f(x)$ is better, the string in the genotype world which corresponds to the solution $x$ gets a better fitness value. Then the function value of the solution $x$ is transformed to the fitness value in the genotype world. In the genotype world, it is easy for a string with a high fitness value to survive. For function optimization problems, if the function is to be maximized, the function value itself can be used directly for the fitness value. Otherwise, if the function is to be minimized, the fitness function should be defined as an increasing function by transforming the function in the phenotype world. For permutation problems, the same thing can be said. Scheduling problems have many evaluation functions such as the makespan, the total flowtime, the tardiness penalty, and so on. Traveling salesman problems also have evaluation functions such as
the total travel distance. Because permutations found by GAs are evaluated by the evaluation functions in the permutation problems, the function values can be transformed to the fitness values in the same way of function optimization problems. In this way, a fitness value is assigned to each string in the genotype world.

### 3.2.3 Selection

Selection is an operator to select two parent strings for generating new strings (i.e., offspring). In the selection, a string with a high fitness value has more chance to be selected as one of parents than a string with a low fitness value. In GAs, parent strings are selected by random choice. The parent strings, however, are not selected by a sheer random choice. The fitness value of each string is utilized for selecting parent strings. We describe two ways of selection schemes which are often employed: the roulette wheel selection scheme and the rank-based selection scheme.

One way is the roulette wheel selection. The roulette wheel selection scheme is often used as a selection operator. Let \( N_{\text{pop}} \) be the number of strings in each population in GAs, that is, \( N_{\text{pop}} \) is the population size. We denote \( N_{\text{pop}} \) strings in the current generation by \( \Psi = \{x_1, x_2, \ldots, x_{N_{\text{pop}}} \} \). Each solution \( x_i \) is selected as a parent string according to the selection probability \( P_s(x_i) \). In the roulette wheel selection scheme, the selection probability \( P_s(x_i) \) is defined as follows:

\[
P_s(x_i) = \frac{f(x_i)}{\sum_{j=1}^{N_{\text{pop}}} f(x_j)}, \quad \text{for } i = 1, 2, \ldots, N_{\text{pop}}
\]

Where \( f(\cdot) \) is the fitness value of the solution \( x \).

The other way is the rank-based selection scheme. In this selection scheme, the population is sorted according to the fitness value. Then, the string which has the best
fitness value in $\Psi$ is ranked as the first string, and the string which has the second best fitness value is ranked as the second string, and so on. The number of selected strings is decided according to their own rank. For example, Table 3.1 shows that the number of strings selected as parent strings is defined as a percentage of $N_{\text{pop}}$.

Table 3.1: The number of selected strings in the rank-based selection scheme

<table>
<thead>
<tr>
<th>Rank</th>
<th>% of $N_{\text{pop}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20%</td>
</tr>
<tr>
<td>2</td>
<td>10%</td>
</tr>
<tr>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>N_{\text{pop}} - 1</td>
<td>0.1%</td>
</tr>
<tr>
<td>N_{\text{pop}}</td>
<td>0%</td>
</tr>
</tbody>
</table>

3.2.4 Crossover

Crossover is an operator to generate new strings (i.e., offspring) from parent strings. Various crossover operators have been proposed for GAs. The crossover operators for permutation strings are different from those for binary strings because permutation problems usually have a requirement that each element of a string should appear only once in the string.

3.2.5 Crossover for Binary Strings

The standard one-point crossover is typical operator for binary strings. The operator is applied to selected parent strings as follows: a crossover point is randomly selected between two adjacent elements. Two new strings are generated by swapping all elements in the head part of the strings. Figure 3.2 shows an example of the one-point crossover operator. In this Figure, a crossover point is selected between the sixth position and the seventh position in the strings. All the elements from the first position to the sixth position are swapped. In this way, two new offspring are generated.

3.2.6 Mutation

Mutation is an operator to change elements in a string which is generated by a crossover operator. Such a mutation operator can be viewed as a transition from a current
solution to its neighborhood solution in local search algorithms. When we apply mutation operators to strings in GAs, we should be careful as in applying crossover operators.

![Crossover point](image)

**Parent 1**: 0 0 1 0 0 1 0 0 1  
**Parent 2**: 1 0 0 1 0 0 1 0 0 0  
**Offspring 1**: 1 0 0 1 0 0 0 0 1  
**Offspring 2**: 0 0 1 0 0 1 1 0 0  

*Figure 3.2*: The standard one-point crossover for binary strings

### 3.2.7 Mutation for Binary Strings

A mutation operator for binary strings is quite simple. An element (i.e., “0” or “1”) is changed to the other element by the mutation operator. An example of this mutation is shown in Figure 3.3 where “0” in the second position is changed to “1”, and “1” in the fourth position and “1” in the seventh position are changed to “0”.

![Mutation](image)

*Figure 3.3*: A mutation for binary strings

### 3.2.8 Elitist Strategy

The elitist strategy is a strategy that the best string in the population should be kept in the next population as it is. That is, the best string is not affected by genetic operators such as crossover and mutation operators. The best string has a lot of chance to be selected as parent strings. It is also kept in the next population as it is in the current population. This strategy is often used for in GAs.

### 3.2.9 The Algorithm

We considered some operations such as coding, evaluation, selection, crossover, mutation, and elitist strategy to construct GAs for optimization problems. We can
construct GAs by employing the above operations. The outline of GAs can be written as follows:

**Step 0 (Initialization):** Randomly generate an initial population of $N_{pop}$ strings where $N_{pop}$ is the population size.

**Step 1 (Evaluation):** Decode strings to solutions in the phenotype world. Next calculate the value of the objective function for each solution. Then transform the value of the objective function for each solution to the value of the fitness function for each string in the genotype world.

**Step 2 (Selection):** Select a pre-specified number of pairs of strings from the current population according to the selection probability discussed above.

**Step 3 (Crossover):** Apply the pre-specified crossover operator to each of the selected pairs in Step 2 to generate $N_{pop}$ strings with the pre-specified crossover probability $P_c$.

**Step 4 (Mutation):** Apply the pre-specified mutation operator to each of the generated strings with the pre-specified mutation probability $P_m$.

**Step 5 (Elitist strategy):** Randomly remove a string from the current population and add the best string in the previous population to the current one.

**Step 6 (Termination test):** If a pre-specified stopping condition is satisfied, stop this algorithm.

Otherwise, return to **Step 1**. The outline of GAs is illustrated in Figure 3.4

### 3.3 Genetic Algorithm for Multi-Objective Optimization Problems

Multi-objective genetic algorithms usually try to find all the non-dominated solutions of an optimization problem with multiple objectives. Let us consider the following multi-objective optimization problem with $n$ objectives:
Maximize \( f_1(x), f_2(x), \ldots, f_n(x) \)

Where \( x \) is a vector to be determined, and \( f_1(.), f_2(.), \ldots, f_n(.) \) are \( n \) objective functions to be maximized. If a feasible solution is not dominated by any other feasible solutions of the multi-objective optimization problem, that solution is said to be a non-dominated solution. When the following inequalities hold between two solutions \( x \) and \( y \), it is said that the solution \( x \) is dominated by the solution \( y \):

\[
fi(x) \leq fi(y) \quad \text{and} \quad fj(x) < fj(y)
\]

Examples of non-dominated solutions are shown in Figure 3.5 where dominated solutions and non-dominated solutions are depicted by open circles and closed circles in a two-dimensional objective space, respectively. The two-dimensional objective space in Figure 3.5 corresponds to the following two-objective optimization problem:

Maximize \( f_1(x) \) and \( f_2(x) \)

Figure 3.4: Outline of GAs for single objective optimization problem
Figure 3.5: Non-inferior solutions (closed circles) and inferior solutions (open circles). The Arrows represent the search directions.

As is shown in Figure 3.5, multi-objective optimization problems usually have several non-dominated solutions. The aim of our multi-objective algorithms is not to determine a single final solution but to find all the non-dominated solutions of the multi-objective optimization problem. Since it is difficult to choose a single solution for a multi-objective optimization problem without iterative interaction with the decision maker, one general approach is to show the set of non-dominated solutions to the decision maker. Then one of the non-dominated solutions can be chosen depending on the preference of the decision maker.

The vector evaluated genetic algorithm is one of “population-based approaches” because its selection to form $n$ subpopulations is implemented according to one of the $n$ objectives separately. We show the search directions of these approaches in Figure 3.5 for the case of the two-objective optimization problem. As we can expect from Figure 3.5, these approaches can easily find the solutions A and D, but it is not easy to find the solutions B and C. In order to find all the non-dominated solutions by GAs, the variety of individuals (i.e., solutions) should be kept in each generation. The solution to this is in their selection scheme, two candidates for selection are picked randomly from the current population. A comparison set which consists of a predefined number of individuals was
also selected from the current population. Then each of the candidates is compared against each individual in the comparison set using the inequalities mentioned above (when all objectives are to be maximized). If one candidate is dominated by the comparison set but the other is not dominated, the latter is selected for the crossover operator. If neither or both are dominated by the comparison set, a fitness sharing technique is adopted.

3.3.1 Evaluation

When we apply GAs to the \( n \)-objective optimization problem, we have to evaluate the values of \( n \) objective functions for each solution. Using these values of \( n \) objective functions, the fitness value of each string should be defined. GAs search a string with a better fitness value in the genotype world as in single-objective optimization problems. A way to transform the values of objective functions to the fitness value of each string in the genotype world is to combine the \( n \) objective functions into a scalar function as follows:

\[
f(x) = w_1^* f_1(x) + w_2^* f_2(x) + \ldots + w_n^* f_n(x)
\]

Where \( f(x) \) is the fitness function of \( x \), and \( w_1, \ldots, w_n \) are non-negative weights for the \( n \) objectives. These weights satisfy the following relations:

\[ w_i \geq 0 \text{ for } i = 1, 2, \ldots, n, \]

If we use constant weight values for the two-objective optimization problem, the search direction by GAs is fixed as shown in Figure 3.6. The search direction in Figure 3.6 corresponds to the weight vector \( \mathbf{w} = (w_1, w_2) = (0.5, 0.5) \) in the two-dimensional objective space. When the search direction is fixed, it is not easy to obtain a variety of non-dominated solutions. In the case of Figure 3.6, GAs with the constant weight vector
\( w = (w_1, w_2) = (0.5, 0.5) \) may easily find the solutions B and C, but it is very difficult to find the solutions A and D. From the above discussions, we can see that neither the constant weight value approach nor the choice of one objective is appropriate for finding all the non-dominated solutions of the multi-objective optimization problem. This is because various search directions are required to find a variety of non-dominated solutions.

**Figure 3.6**: The search direction determined by the constant weight vector

### 3.3.2 Selection

When a pair of parent strings are to be selected from a current population \( \Psi \) for generating an offspring by a crossover operator, first \( n \) weight values \( (w_1, w_2, \ldots, w_n) \) are randomly specified. Then the fitness value of each solution \( x \) in the current population \( \Psi \) is calculated as the weighted sum of the \( n \) objectives by the optimization equation 3.3. The selection probability \( P_s(x_i) \) of each string \( x \) based on the linear scaling is defined by the roulette wheel selection as follows:

\[
P_s(x_i) = \frac{f(x_i) - f_{\min}(\Psi)}{\sum_{j=1}^{N_{\text{pop}}} \{f(x_j) - f_{\min}(\Psi)\}}, \quad \text{for } i = 1, 2, \ldots, N_{\text{pop}}
\]  

3.3
Where \( f_{\text{min}}(\psi) \) is the minimum fitness value (i.e., the worst fitness value) in the current population \( \psi \). According to this selection probability, a pair of parent strings is selected from the current population \( \psi \).

3.3.3 Elitist Strategy

During execution of GAs for multi-objective optimization, two sets of solutions are stored: a current population and a tentative set of non-dominated solutions. After evaluating all the strings in the current population, the tentative set of non-dominated solutions is updated by the current population. That is, if a string in the current population is not dominated by any other strings in the current population and the tentative set of non-dominated solutions, this string is added to the tentative set. Then all solutions dominated by the added one are eliminated from the tentative set. In this manner, the tentative set of non-dominated solutions is updated at every generation in GAs for multi-objective optimization problems. From the tentative set of non-dominated solutions, a few solutions are randomly selected and added to the current population. The randomly selected non-dominated solutions may be viewed as a kind of elite solutions because they are added to the current population with no genetic operations.

3.3.4 Multi-Objective Genetic Algorithm (MOGA)

We considered some modified operations such as evaluation, selection, and elitist strategy in the previous sections in order to construct a genetic algorithm for multi-objective optimization problems. We can construct a multi-objective genetic algorithm (MOGA) by employing those operations for multi-objective optimization. The outline of the MOGA can be written as follows:
Step 0 (Initialization): Randomly generate an initial population of $N_{\text{pop}}$ strings where $N_{\text{pop}}$ is the population size.

Step 1 (Evaluation): Decode strings to solutions in the phenotype world. Next calculate the values of the $n$ objectives for each solution. Then update the tentative set of non-dominated solutions.

Step 2 (Selection): Repeat the following procedure to select parent strings to generate $N_{\text{pop}}$ strings.

(i) Specify the weight values in the fitness function

(ii) According to the selection probability in select a pair of parent strings.

Step 3 (Crossover): Apply a crossover operator to each of the selected pairs in Step 2.

Step 4 (Mutation): Apply a mutation operator to each of the generated strings with mutation probability $P_m$.

Step 5 (Elitist strategy): Randomly remove $N_{\text{elite}}$ solutions from the generated $N_{\text{pop}}$ solutions, and add $N_{\text{elite}}$ solutions that are randomly selected from the tentative set of non-dominated solutions.

Step 6 (Termination test): If a pre-specified stopping condition is satisfied, stop this algorithm.

Otherwise, return to Step 1. Figure 3.7 shows the outline of MOGA.
Figure 3.7: The outline of MOGA
Chapter 4: Development of Chandler Production Well Management Model

4.1 The Governing Equations

A pressurized water distribution system is a network, which must considers both continuity equation at nodes and energy equation in arcs. The nodes include junction nodes, demand nodes, supply nodes, valves, and tanks. Junction nodes and valves have no storage capacity; however, tanks are able to store water over time. The arcs include pipes and pumps. The continuity equation at each node is (Ahuja et al. 1993)

\[ \sum_{j \in N} q_{ij,t} - \sum_{j \in N} q_{ji,t} = b_{i,t} \quad \text{for } i \in N, \text{ and } t = 1, 2, \ldots, T \]

where \( A \) is the arc set of the system; \( N \) is the node set of the system; \( q_{ij,t} \) is the total flow rate in arc \((i, j)\) from node \( i \) to node \( j \) during time \( t \); \( b_{i,t} \) is the supply/demand node \( i \); and \( T \) is the number of time steps. \( b_{i,t} \) is zero for junction nodes and valves. \( b_{i,t} > 0 \) represents a supply node. \( b_{i,t} < 0 \) represents a demand node. \( b_{i,t} \) also represents the storage variation in tanks. The energy equation for each arc connecting nodes \( i \) and \( j \) is (Rossman 2000; Goldman and Mays 2005)

\[ h_{i,t} - h_{j,t} = f(q_{ij,t}) \quad \text{for } (i, j) \in A, \text{ and } t = 1, 2, \ldots, T \]

where \( h_{i,t} \) is the hydraulic head at node \( i \) and time \( t \). \( f(q_{ij,t}) \) is the headloss function in the arc \((i, j)\). The headloss can be a negative value to represent the head gain from pumps. The number of unknowns is \((A + N)T\) and the total number of equations 4.1 and 4.2 is also \((A + N)T\).

The transient groundwater flow equation in a three-dimensional heterogeneous and anisotropic aquifer is (Harbaugh et al. 2000)
\[ \frac{\partial}{\partial x} \left( K_{xx} \frac{\partial \phi}{\partial x} \right) + \frac{\partial}{\partial y} \left( K_{yy} \frac{\partial \phi}{\partial y} \right) + \frac{\partial}{\partial z} \left( K_{zz} \frac{\partial \phi}{\partial z} \right) + Q = S_s \frac{\partial \phi}{\partial t} \quad 4.3 \]

where \( \phi \) is the groundwater head; \( K_{xx}, K_{yy}, \) and \( K_{zz} \) are hydraulic conductivity along principle directions \( x, y, \) and \( z, \) respectively; \( Q \) is the pumping/recharge term; and \( S_s \) is the specific storage.

The linkage between the pressurized water distribution system and the aquifer system is through the well pumps, where the groundwater is pumped into either free-surface tanks or the pressurized junction nodes. Once the pumping rates are determined by the hydraulics in equation 4.2, the pumping rates in groundwater model in equation 4.3 are distributed according to the transmissivity-weighted pumping rates at the vertical pumping (sink) nodes for multi-layer wells

\[ Q_{i,\ell} = \frac{T_{i,\ell}}{\sum_k T_{i,k}} q_{i,\ell} \quad i \in P \in N \quad 4.4 \]

where \( P \) is pumping well set, a subset of node \( N; \) \( Q_{i,\ell} \) is the pumping rate at well \( i, \) layer \( \ell, \) time \( t; \) and \( T_{i,\ell} \) is the transmissivity value at well \( i, \) layer \( \ell. \) Therefore, \( q_{i,\ell} = \sum_\ell Q_{i,\ell}. \) For the simplicity purposes, the pumping rate is denoted as \( Q_{i,\ell} = q_{i,\ell} \) in the following objective functions.

The groundwater head at the well pump location will not be affected by the pressurized water distribution system. However, the well pump efficiency and flow rate are governed by the groundwater head and the pressurized water distribution system through equation 4.2. In other words, the aquifer is considered as a boundary node to the distribution system, which serves as a dynamic water source.
4.2 Conjunctive Management Model

4.2.1 Decision Variable: Well Pump Operation Patterns

Once the above simulation models are successfully linked, the complexity of the conjunctive management model depends on decision-makers’ objectives in order to better the integrated water system performance. The important decision variables include booster pump operations which directly deliver tank water into the distribution system, surface water treatment plant (SWTP) operations which control major treated surface water release to the demand nodes, boundary source node operations which deliver trans-boundary surface water into the water distribution system through federal and/or state water projects, and well pump operations which directly impact on the groundwater system and the pressurized distribution system. Although it is possible to incorporate all kinds of decision variables in the conjunctive management model, the current stage focuses on well pump operations optimization because well pumps directly control amount of groundwater into the distribution system.

The well pump operations optimization problem is mainly a scheduling optimization problem, which optimizes the well pump on/off schedule. Although there is a great amount of studies on pumping rate optimization in groundwater management, actual pumping rates are hardly to be controlled due to hydraulic constraints. Instead, we can control the on/off schedule such that the pumping rates are indirectly controlled close to the desired pumping efficiency level.

In this study, each well pump has an individual 24-hour operation pattern (daily operation pattern) with 1-hour time interval. The same operation pattern repeats daily for one week. Therefore, each operation pattern has 24 decision variables, a sequence of
binaries (1s and 0s), to determine the on and off status for each hour. The actual pumping rate is $Q'_{i,t} = Q_{i,t} z_{i,t}$, where $z_{i,t} = \{0, 1\}$ is a binary variable which controls the pumping status of well $i$ at time $t$. For example, if $z_{i,t} = 1$, the well pump $i$ is on and the pumping rate is $Q_{i,t}$ at time $t$. If $z_{i,t} = 0$, the well pump $i$ is off and $Q'_{i,t} = 0$ at time $t$.

The reasons of considering periodic daily operation patterns for a week are two-fold. First, the periodic daily operation patterns make well pumps easier to be operated because the daily water demand patterns are typically repeated for next day. Second, the number of decision variables will be very large, and unreasonable well pump operations may be resulted if the operation allows to be changed hourly for the whole week in the optimization model. Moreover, the periodic daily operation patterns will help the ending tank water level be close to its initial tank water level after 24-hour operations. The periodic water variation in tanks will reflect a proper system operation and ease system management. This study proposes to develop a conjunctive management model that has the capability of providing the 24-hour real-time well pump operations while forecasting the groundwater system and water distribution system responses for 7 days. The proposed management objectives are the followings.

4.2.2 Management Objectives

4.2.2.1 Objective 1: Energy Cost Minimization

The energy (kW-h) consumed by a well pump $i$ at time $t$ is calculated by

$$E_{i,t} = \frac{\gamma_w Q'_{i,t} H_{i,t}}{\eta_{i,t}}$$

where $\gamma_w$ is the water specific weight; $H_{i,t}$ is the lifted head by the pump $i$ at time $t$; and $\eta_{i,t}$ is the pump efficiency for pump $i$ at time $t$, which is determined by the efficiency
The unit energy cost can vary according to electricity consumption during the peak hours or off-peak hours. The overall energy cost is

\[
\min_{z_{i,t} \in \{0,1\}} \sum_i \sum_t \frac{\gamma_n O_{i,t}^t H_{i,t}}{\eta_{i,t}} p_{i,t}
\]

where \( p_{i,t} \) [$/kW-h] is the unit energy cost for pump \( i \) at time \( t \).

### 4.2.2.2 Objective 2: Daily Tank Level Deviation Minimization

One of the purposes of developing the management model is to optimize the well pump operations such that the ending tank water level is close to its initial water level after 24-hour operations. Again, the periodic water variation in the tanks over a week will help a proper system operation and management because the daily water demand pattern is typically periodic. The importance of keeping similar ending storage level to the beginning storage level has been highly recommended in the multi-reservoir system optimization (REFs). The tank levels in the pressurized water distribution system play very a crucial role in the pump scheduling optimization problem (Chase and Ormsbee 1993; Lansey and Awumah 1994). The tank level deviation minimization objective function is

\[
\min_{z_{i,t} \in \{0,1\}} \sum_k \sum_n \left[ H_{k,24n}^{\text{Tank}} - H_{k,24(n-1)}^{\text{Tank}} \right]^2
\]

where \( H_{k,24n}^{\text{Tank}} \) is the tank level at tank \( k \) at the end of day \( n \); and \( H_{k,24(n-1)}^{\text{Tank}} \) is the tank level at the beginning of day \( n \).

### 4.2.2.3 Objective 3: Weekly Drawdown Minimization

The weekly drawdowns at the pumping sites are calculated to evaluate the depletion of the aquifer system due to groundwater withdrawal, which has been one of the major objectives in the conjunctive use of groundwater and surface water
management (Huang and Mayer 1997, Miles and Lence 1996, Richards et al. 1993, and McPhee and Yeh 2004.) Negative drawdowns can be resulted if the replenishment rate is higher than the pumping rate. Only positive drawdowns are considered in the objective function

$$\min_{z_{ij} \in [0,1]} \sum_i \max \left[ 0, \phi_{i,T} - \phi_{i,0} \right]$$  \hspace{1cm} 4.8

where $\phi_{i,T}$ is the groundwater head at pumping site $i$, last time step $T$; and $\phi_{i,0}$ is the initial groundwater head at pumping site $i$.

**4.2.2.4 Objective 4: Tank Water Residence Time Minimization**

The tank water residence time (TR) is defined as the duration of water residing in the tank. Farmani et al. 2006 has discussed the approach of conflicting objectives in water distribution system optimization, with residence time as one of the objectives. Resident time is an indicator to water quality in the tank; and reducing residence time of water in tanks is important for maintaining good water quality in the distribution system.

In this study, the residence time is calculated in one hour interval by dividing the volume of the tank water $S_{i,t}$ by the net outflow rate $b_{i,t}$ in equation 4.1:

$$TR_{i,t} = \min \left[ 1, \frac{S_{i,t}}{b_{i,t}} \right]$$

If $TR_{i,t} > 1$ hour, the residence time is taken as one hour; otherwise the residence time is $TR_{i,t}$. Moreover, if there is no outflow rate or the net outflow rate is negative, the residence time is taken as one hour. The objective for the total tank water residence time minimization is

$$\min_{z_{ij} \in [0,1]} \sum_i \sum_t \min \left[ 1, \frac{S_{i,t}}{b_{i,t}} \right]$$  \hspace{1cm} 4.9
4.2.2.5 Objective 5: System Reliability Maximization

Opposite to the residence time minimization, we also consider the system reliability to avoid releasing tank water too fast. Conceptually, more water in the tanks with longer time increases the distribution system reliability. The residence time $TR_{i,t}$ less than one hour is undesired because the net outflow rate at that hour completely drains the tank and may lead to lack of water to the system in the later time. Increasing total number of $TR_{i,t} = 1$ over the planning horizon will increase the system reliability. This study defines the system reliability is percentage of the total residence time with $TR_{i,t} = 1$ with respect to the planning horizon. The objective function is

$$
\min_{z_{i,t} \in \{0,1\}} \frac{1}{N_t \times T} \sum_{i} \sum_{t} \text{INT} \left[ TR_{i,t} \right]
$$

where $N_t$ is the number of tanks; and $\text{INT} \left[ TR_{i,t} \right]$ gives the integer value 0 if $TR_{i,t} < 0$ and 1 if $TR_{i,t} = 1$.

4.2.2.6 Objective 6: Pressure Violation Minimization

Pressure violation is the result of the water pressure deficit or excess at the junction nodes. Pressure violation has a very close relationship to the well pump operations and tank water levels in the distribution system such that pressure violation minimization contradict the energy cost minimization (Objective 1). Delivering sufficient pressure to the water distribution with respect to optimal operation is important (Lansey and Basnet 1991, Broad et al. 2005, Miles and Lence 1996, Pezeshk and Helweg 1996, and Cheung et al. 2003). From the design point of view, the water distribution system is usually divided into several pressure zones according to the water source locations and the system topography. Using one single designed pressure level across the entire
distribution system is not efficient. For example, the Chandler distribution is considering two pressure zones; and two different minimum designed pressure levels are specified.

Considering the pressure violation as a hard constraint will restrict the flexibility on well pump operations and will complicate the optimization problem. In the real operation, pressure violation is acceptable at a certain level. For the planning purpose, the desired pressure level is often intentionally high in order to investigate the system capacity. Therefore, this study considers the total pressure violation as a penalty function in the objective function as the following,

\[
\min_{z_{i,j} \in [0,1]} \sum_{j} r_{j} \sum_{i} \max_{t} (0, P_{\text{design}} - P_{i,j})
\]

where \( P_{\text{design}} \) is the designed pressure head for zone \( j \); \( P_{i,j} \) is the pressure head at junction node \( i \), time \( t \), in zone \( j \); and \( r_{j} \) is penalty term for zone \( j \). Pressure violation is not considered if the pressure head is larger than the designed pressure head.

### 4.2.3 Multi-Objective Model

The well pump management model considers above objectives and penalties and renders a multi-objective function using the weighting method

\[
\min_{z_{i,j} \in [0,1]} \begin{align*}
& w_{1} \sum_{i} \sum_{t} \frac{\gamma_{w} Q_{i,j} H_{i,t}}{\eta_{i,t}} p_{i,t} + w_{2} \sum_{k} \sum_{n} [H_{\text{Tank}}^{k,24} - H_{\text{Tank}}^{k,24(n-1)}]^{2} + w_{3} \sum_{i} \max(0, \phi_{i,T} - \phi_{0}) \\
& + w_{4} \sum_{i} \sum_{t} \min \left[ 1, \frac{S_{i,t}}{b_{i,t}} \right] - w_{5} \frac{1}{N_{i} \times T} \sum_{i} \sum_{t} \text{INT}[TR_{i,t}] \\
& + \sum_{j} r_{j} \sum_{i} \min_{t} (0, P_{\text{design}}^{j} - P_{i,j}^{j})
\end{align*}
\]

where \( w_{i}, i = 1, 2, \ldots, 5 \) are the weights for individual the objectives functions.
4.3 Solution Strategy for the Chandler Case Study

The multi-objective model is a binary integer nonlinear programming problem (BINLP), which involves groundwater modeling and hydraulic pipe network modeling. The difficulty of solving equation 4.12 depends on the number of binary integer variables and complexity of the system. One common approach is to relax the binary variable to the real variable such that BINLP becomes NLP to be solved by the gradient-based method. However, the relaxation of the binary variables generally alters the true optimal locations and NLP converges to a local optimum. Although the cutting plane algorithm and branch-and-bound method have been proven to be very efficient in the integer linear programming (ILP), the BINLP has to be linearized as a successive integer linear problem to be solved iteratively by ILP method. Linearization of equation 4.12 may be very difficult for the Chandler problem because Jacobian matrices have to be evaluated from the groundwater and water distribution models. Moreover, the enumeration method only works for small amount of binary variables. Local search method may be feasible for a fairly amount of binary variables; however, the solution is easily to be trapped at the local optimum. Other than blindly determine what optimization algorithm(s) to be used, we need to understand the Chandler system in order to find an approximate solution method.

The City of Chandler has been developing a detailed pressurized water distribution model using EPANET (Rossman 2000), which solves equations 4.1-4.2, as a tool to plan for the City’s buildout in the next 15 years. The detailed distribution system is shown in Figure 2.1, where 11,392 junctions and 14,258 pipes are currently considered in the system. The distribution system incorporates the major water supply sources for the
City from Salt River Project (SRP), Roosevelt Water Conservation District (RWCD), Roosevelt Dam New Conservation Storage, Central Arizona Project (CAP), Reclamation facilities, and Safe Yield Pumping. The project waters are delivered into the system through the trans-boundary reservoirs. The surface water treatment plant (SWTP) is the major surface water supplier to the system. The safe yield pumping is conducted through the 23 well pumps inside the City, which are connected to the groundwater model. 13 pumping wells directly pump groundwater into the storage tanks and 10 pumping wells deliver groundwater into the pressurized system. There are 24 storage tanks in the model, which includes aforementioned 13 tanks receiving water from the well pumps and 11 tanks distributed in the network without being associated with the well pumps. Other than the well pumps, there are 100 booster pumps connected to the storage tanks to deliver water into the pressurized system. The booster pump operations are given and will not be considered as the decision variables in the management model. Two pressure zones in the distribution system are considered. Zone 2 covers the southeastern region (a new developing area) with a designed pressure head $P_{\text{design}}^{\text{Zone2}} = 435.8$ meters (1430 feet) and Zone 1 covers rest of the City with a designed pressure head $P_{\text{design}}^{\text{Zone1}} = 419.1$ meters (1375 feet) (Chandler City Master Plan, unpublished report).

The Chandler is located on the Salt River alluvial basin, which contains three main unconfined aquifer units. Most of the Chandler wells withdraw groundwater model from the Middle Alluvial Unit and Lower Alluvial Unit. The Upper Alluvial Unit is mostly unsaturated over much of the area to the east of Chandler (Laney and Hahn, 1986). MODFLOW (Harbaugh et al. 2000), which solves equation 4.3, was employed by the City to predict the groundwater resource and drawdown over a large area much larger
than the City. The groundwater model considers the pumping wells inside and outside the Chandler. The aquifer system is divided into three layers according to three main aquifer units. Each layer is discretized into 71 rows and 98 columns in the model. The City of Chandler is bounded from 14th to 57th rows and from 21th to 70th columns. The model considers the groundwater budget through the drainage, evaporation, areal recharge, injection wells and pumping wells. The groundwater model has been calibrated with the measured groundwater head data for the period between 1989 and 1998 (Southwest Ground-water Consultants, Inc. 2003, unpublished report).

This study focuses on management model evaluation during the week 7/10/2005–7/16/2005 since we have the measured data during this period to compare the actual operations and energy cost against those using optimized well pump operations under different scenarios. Therefore, the total number of time steps is $24 \times 7 = 168\). The total number of the binary variables is $24 \times 23 = 552\) according the 23 well pumps, which renders the total possible solutions to be around $1.4742 \times 10^{166}\). Although there may exist many potential optimization approaches to solve the formulated BINLP problem method, feasible algorithms are limited. The large amount of binary variables and the form of the objective functions make the BINLP intractable by either the relaxation or linearization methods. The local search method is expected to be computationally demanding and gives the local optimal solution.

### 4.4 Genetic Algorithm

From our experience, a genetic algorithm (GA) would be a straightforward solution method for solving the BINLP due to the intrinsic binary chromosome in GA. GA is one of the evolutionary algorithms, which has a potential to search for global or
near-global optimal solution while only analyzing a small fraction of possible solutions.

GA is a derivative-free algorithm and only requires the objective function evaluation. GA evolves the solutions from multiple starting points and uses probabilistic transition rules to reduce the possibility of being trapped in a local optimum. In past decade, GA has become one of important optimization techniques to solve the water resources system optimization problem (van Zyl et al. 2004, Cai et al. 2001, Mckinney and Lin 1994, Sharif and Wardlaw 2000, Dandy et al. 1996, and Savic et al. 1999).

We adopt a GA solver (Carroll, 1996) to our problem. A binary chromosome with 552 bits is designed to represent a feasible solution for the well pump operations. Each bit has either 0 or 1 value. The GA searches for the optimal solution using a set of processes and operators analogous to bio-evolution processes (selection, crossover, mutation, reproduction, and replacement) to evolve a population of chromosomes such that the best fitness among chromosomes is maximized generation by generation. The fitness is given as the negative value of the objective function for the minimization purpose. With the Chandler infrastructure, the model objective in equation 4.12 becomes

\[
- \text{fitness} = w_1 \sum_{i=1}^{23} \sum_{t=1}^{168} \frac{\gamma^i Q_{i,t} H_{i,t}}{\eta_{i,t}} p_{i,t} + w_2 \sum_{k=1}^{24} \sum_{n=1}^{7} \left[ H_{k,24n}^{\text{Tank}} - H_{k,24(n-1)}^{\text{Tank}} \right]^2 + w_3 \sum_{i=1}^{23} \max \left[ 0, \phi_{i,168} - \phi_{i,0} \right] \\
+ w_4 \sum_{i=1}^{24} \sum_{t=1}^{168} \min \left[ 1, \frac{S_{i,t}}{b_{i,t}} \right] - w_5 \frac{1}{24 \times 168} \sum_{i=1}^{24} \sum_{t=1}^{168} \text{INT} \left[ TR_{i,t} \right] \\
+ r_1 \sum_{i=1}^{168} \max \left( 0, 0.419.1 - P_{i,1}^{\text{Zone1}} \right) + r_2 \sum_{i=1}^{168} \max \left( 0, 0.435.8 - P_{i,2}^{\text{Zone2}} \right)
\]

4.13

The GA links with EPANET and MODFLOW as a simulation-optimization model to obtain the optimal 23 well pump daily operation patterns.
4.5 Linking Chandler EPANET and MODFLOW Models

The linkage between the Chandler EPANET and MODFLOW models is made through the 23 well pumps. We assume that the groundwater head changes are insignificant within a day. Therefore, the groundwater levels in EPANET are updated weekly using MODFLOW results. The weekly stress period is considered in the groundwater model. The MODFLOW well package is updated using the EPANET pumping rates. The data transfer between the EPANET and MODFLOW for the transient simulation is the following:

Step 1: Specify well pump operation patterns.
Step 2: Run EPANET for 168 hours with the given well pump operation patterns.
Step 3: Record the calculated well pumping rates, energy consumed by the well pumps, and tank levels from EPANET.
Step 4: Calculate the weekly pumping rates from EPANET.
Step 5: Update the MODFLOW well package with the pumping rates in Step 4 and run the MODFLOW for 7 days. The evapotranspiration, drainage, and surface recharge remain the same within the year of 2005.
Step 6: Calculate the drawdown for the week.

Once obtained through Steps 1-6 in GA, the optimal operation patterns can be employed as the first 24-hour real-time operations. The 7-day forecast in the water distribution system and aquifer system can also be obtained. For the next 24-hour real-time operations, updating the EPANET groundwater reservoir levels is recommended in Step 7: Use the MODFLOW groundwater heads to update the EPANET groundwater reservoir levels for the next day. Repeat Steps 1-6.
4.6 The User Interface

The Chandler production well management model’s user interface is based on Microsoft EXCEL. The calculations within the codes and the results produced in between and finally are transported into the EXCEL sheet for further processing, so as to make them presentable and distinguishable. The following data are exported into the EXCEL work sheets.

1. The optimal and initial energy values for all the 23 pumps for one week i.e. an array of $23 \times 168$ along with the optimal and initial total energy values for the whole simulation.

2. The optimal pump scheduling patterns array ($23 \times 24$).

3. The optimal and initial pumping rates array ($23 \times 168$).

4. The optimal and initial values of reliability and residence time.

5. The optimal and initial values of tank level difference minimization function as discussed previously in the report.

6. The optimal and initial values of the pressure difference minimization function for zone one and zone two.

7. The optimal and initial values of the tank levels for all the 13 tanks for one week i.e. an array of $13 \times 168$.

8. The optimal drawdown distribution ($71 \times 98$ array) for three layers.

9. The optimal and initial values of total drawdown i.e. summation for all the three layers.

10. The maximum and minimum pressure values in both the zones along with the time of their occurrences and the respective junction nodes where these occur.
11. The optimal and initial head distribution for three layers.

The only values imported in the main FORTRAN code from EXCEL are the values of the weighting coefficients in the main objective function equation. These values are input by the user in the main EXCEL worksheet.

The Microsoft EXCEL main work sheet for the Chandler production well management model is shown in Figure 4.1. Some major post and pre-simulation results are printed on the main sheet. On the top of the sheet the date and time when the simulation is started are shown. Along with this the actual time i.e. for the week July 10 to July 16, 2005, for which the simulation is being carried out is also given. Therefore the number of days analyzed are 7 as printed on the sheet. Below this is the main objective function equation. It is a minimization problem, so we try to minimize this equation which has all the 7 different objectives.

The only user input in the simulation is the weighting coefficients in the equation; these are input from the red cells under the heading weights. In the same row on right, the definition for each objective is given along with the nature of optimization they are going through. The simulation is run for 100 GA generations using 5 chromosomes. The final objective values (i.e. after 100 GA generations) for the 7 different objectives are printed out in the same row, following are the values obtained after multiplying these with their respective weighting coefficients (red cells). Then we sum all these values to get the total value of the objective function after optimization. Following these values in the same row the objective values at the beginning of the simulation are printed put, then these are multiplied with their respective weighting coefficients and are summed up to get the total value of the objective function before optimization. The last column in the same row
gives us the percent decrease or percent increase in the different objectives. For example, the energy which is a minimization function has decreased by about 30% in 100 GA generations. On the lower left side we calculate the pressure violations. The magnitudes of the maximum and minimum pressures in the two pressure zone are calculated, along with the time of their occurrence. The junctions where these maximum and minimum pressures occur during the simulation are also printed out. Below in the same column we calculate the maximum percent violation for the maximum and minimum pressures in both the zone with respect to the thresholds. For zone 1 and 2 the respective thresholds are 1375 and 1430 feet. The lower middle part gives us the final objective value of the best chromosome after 100 GA generations. The total time taken for the simulation is also calculated. The curve on the lower right gives us variations of the energy and the pumping rates for the total simulation. The energy is represented in the right side axis with two curves one representing the energy at the beginning and other at the end of optimization. Similarly the pumping rates are given on the left side of the axis with curves for before and after the optimization.

The EXCEL sheet will have the following curves along with their respective arrays

1. Optimal pattern curves.
2. Optimal, initial and original tank level, energy and pumping rate curves.
3. Optimal, original and starting head curves.
4. Optimal drawdown curves.

Figures 4.2 and 4.3 represent the linking of the simulation model and the main optimization framework under genetic algorithm, respectively.
The main objective function

$$\min \sum_{i=0}^{n} \left( W_1 \sum_{j=1}^{m} (F_{ij} - R_{ij})^2 + W_2 \sum_{j=1}^{m} (F_{ij} - R_{ij})^2 + W_3 \sum_{j=1}^{m} \left( H_{ij} - \left( E_{ij} - \frac{H_{ij}}{2} \right) \right)^2 + W_4 R_{BL} + W_5 \sum_{j=1}^{m} \left( h_{ij}^{(n)} - h_{ij}^{(n-1)} \right)^2 + W_6 \sum_{j=1}^{m} (p_j^2 - p_j^2) + W_7 \sum_{j=1}^{m} (p_j^2 - p_j^2) \right)$$

Table of results:

<table>
<thead>
<tr>
<th>Input Values in red cells</th>
<th>Post Simulation Results</th>
<th>Pre-Simulation Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>W1 540960.9</td>
<td>Total Energy (Min)</td>
<td>1.7384891607504E+05</td>
</tr>
<tr>
<td>W2 196.0</td>
<td>Residence Time (Min)</td>
<td>1.834160738802E+03</td>
</tr>
<tr>
<td>W3 500.0</td>
<td>Daily Load (Max)</td>
<td>1.584920579059E+02</td>
</tr>
<tr>
<td>W4 16.0</td>
<td>Reliability (Max)</td>
<td>7.498764085586E+01</td>
</tr>
<tr>
<td>W5 35.8</td>
<td>Total Breakdown (Min)</td>
<td>2.612391902028E+01</td>
</tr>
<tr>
<td>W6 500</td>
<td>Zone 1 Pressure Diff (Min)</td>
<td>8.217080608375E+01</td>
</tr>
<tr>
<td>W7 500</td>
<td>Zone 2 Pressure Diff (Min)</td>
<td>9.736827073919E+01</td>
</tr>
</tbody>
</table>

Pressure calculations and violations:

<table>
<thead>
<tr>
<th>Zone 1</th>
<th>Zone 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min. pres. (m)</td>
<td>969.72</td>
</tr>
<tr>
<td>Min. pres. Node ID</td>
<td>594</td>
</tr>
<tr>
<td>Time</td>
<td>201600.00</td>
</tr>
<tr>
<td>Max. pres. (m)</td>
<td>1764.26</td>
</tr>
<tr>
<td>Max. pres. Node ID</td>
<td>699</td>
</tr>
<tr>
<td>Time</td>
<td>500000.00</td>
</tr>
</tbody>
</table>

Figure 4.1: The main EXCEL worksheet
Run EPANET for 168 hours

Retrieve and calculate the following values from the standalone program during the simulation:
- Retrieve pumping data for all the well pumps (i,t)
- Retrieve tank level for all 13 tanks (i,t)
- Retrieve energy consumed (i,t)
- Retrieve pressure at each junction for zone 1 and zone 2 (i,t)

Calculate the tank level difference function
Calculate the pressure difference function
Calculate residence time
Calculate reliability

Calculate the average pumping rate in cubic feet per day

Calculate the pumping rates based on the transmissivity of the corresponding wells, since most of the wells pump water from both 2nd and 3rd layers

Substitute the pumping rates in the MODFLOW well package at the corresponding well locations

Run MODFLOW for one week (7/10/2005 – 7/16/2005)

Retrieve the final head distribution for all the three layers.

Calculate the summation of the drawdown values for all the pumping locations

Transfer all the objective function values into the main optimization program

Figure 4.2: Linking EPANET and MODFLOW models
The main optimization program starts with;
MAXGEN as the maximum number of generations to be carried out
Set i=1

POPSIZE is the maximum number of chromosomes in one generation. Generate POPSIZE
number of random 23 X 24 matrix with 0’s and 1’s (Pump Patterns); set j=1

The randomly generated POPSIZE number of chromosomes are assigned
j values i.e. the order in which they enter the simulation

Substitute the generated patterns (for specified j) into the EPANET input file

Substitute the reservoir head values from the
original MODFLOW head file at 7/10/2005 into EPANET input file

Run the linked simulation model
(As described in the previous flowchart)

Calculate the fitness for the chromosome i

Increment j by 1

If j < POPSIZE
Yes
No

If i < MAXGEN
Yes

Take two chromosomes with larger fitness value
and generate POPSIZE number of chromosomes
/crossover

Increment i by 1

Select the maximum fitness chromosome from the final generation and
Run the linked simulation model
(As described in the previous flowchart)

Transfer all the calculated data into Excel using VB application

Figure 4.3: The main optimization framework under genetic algorithm
4.7 Parallel Genetic Algorithm

Given a set of weight and penalty values, we employed the micro GA (Carroll, 1996) with 5 chromosomes over 200 generations to minimize the objective function in order to obtain the optimal pumping patterns. The computation time for running one groundwater model is 3 seconds and for running one water distribution model is 6.5 minutes under the DOS command lines on a PC with 3.2 GHz Intel Pentium V processor and 2 GB RAM. It took 4.5 days for running the models $5 \times 200 = 1000$ times in GA. Obviously, the management model will lose its goal of predicting the 24-hour real-time operations using a single computer. Parallelization of the current genetic algorithm solver and the use of a cluster of computers have to be implemented to obtain the solution much less than 24 hours computation time.

Parallelizing genetic algorithm is easy to be implemented by distributing one chromosome to one computer such that groundwater model and water distribution model with different well operation patterns can be run simultaneously at different computers. The MPI (Message Passing Interface) library is added to the GA solver (Carroll, 1996). The parallel codes were run in a cluster, called MANGAL, which has 64 Intel Pentium IV 1.66GHz processors. Each processor has 1GB RAM. Currently, we only use 12 processors to demonstrate the efficiency of the parallel genetic algorithm to our problem. Therefore, instead of using micro GA, we use 12 chromosomes (because of 12 processors) in one population and consider 75 generations. It took 8 hours to finish $12 \times 75 = 900$ model runs in the parallel GA. Using more processors is possible, but the use of 12 processors is able to achieve the goal that the optimized operation patterns are
obtained much less 24 hours computation time. Figure 4.4 shows a comparison between single and parallel processor computation speed.

**Figure 4.4:** Comparison between Single and Parallel Processor Computation Speed
Chapter 5: Results and Conclusion

5.1 Results

Before conducting the optimization, we first run the groundwater and water distribution models using the documented well pump operations during the week 7/10/2005–7/16/2005 as the original model to develop the baseline against the optimized model and two extreme cases. The two extreme cases are all well pumps off (pattern values are 0) and all well pumps on (pattern values are 1), respectively. Through the extreme case analysis as show in Table 5.1, we are able to acknowledge the upper and lower limits for all the objective function values.

**Table 5.1: Original and Extreme Cases**

<table>
<thead>
<tr>
<th></th>
<th>Original Model</th>
<th>“All-on” Model</th>
<th>“All-off” Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Energy (kW-h)</td>
<td>456003.00</td>
<td>512363.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Daily Tank Level Deviation (m)</td>
<td>113.32</td>
<td>64.52</td>
<td>20.32</td>
</tr>
<tr>
<td>Total Drawdown (m)</td>
<td>2.75</td>
<td>3.58</td>
<td>0.23</td>
</tr>
<tr>
<td>Residence Time (h)</td>
<td>2134.65</td>
<td>2174.54</td>
<td>216.32</td>
</tr>
<tr>
<td>System Reliability (%)</td>
<td>96.33</td>
<td>99.65</td>
<td>8.90</td>
</tr>
<tr>
<td>Zone 1 Pressure Violation (m)</td>
<td>23073.75</td>
<td>12624.32</td>
<td>30011.16</td>
</tr>
<tr>
<td>Zone 2 Pressure Violation (m)</td>
<td>5070.11</td>
<td>1823.09</td>
<td>8431.19</td>
</tr>
</tbody>
</table>

The small energy difference between the original model and “all-on” model indicates that the original well pumps are operated most of the time, which gives the system reliability more than 96%. The high system reliability is the original intention of operating the Chandler water distribution system without considering daily periodic operation patterns. The original model gives small variation of hourly energy output from 23 well pumps from 2,650 kW-h to 2,920 kW-h over a week. The total energy consumed by the well pumps is 456,003 kW-h. The minimum and maximum hourly pumping rates are $1.89 \times 10^5$ m$^3$/day and $2.32 \times 10^5$ m$^3$/day, respectively, in the original model.
original well pump operations lead to high tank levels such that the daily tank level deviation is small. The daily tank level deviation is also small in the extreme cases because most of the tanks are either full or empty within one day.

The maximum drawdown, 0.4 meters, occurs at the “West Peco Road” well site in the Middle Alluvial Unit. Most of the drawdowns are in the northeastern part of the city. The total drawdowns are 2.75 meters by summing all the positive drawdowns at 23 well pump locations according Objective 3. As shown in Table 5.1, the zone 1 and zone 2 pressure violations take places even though all well pumps are tuned on. The original model give the pressure violations on zone 1 and zone 2 are 23,074 meters and 5,070 meters respectively. The total residence time in the original model is 2,135 hours, which is 97.8% of total residence time in one week.

In summary, the original operation patterns can be improved to avoid high water levels in the tanks and reduce the unnecessary high system reliability.

With the multiple objectives, the first task is to determine the reasonable weight and penalty values according to the priority and sensitivity of each objective to the system. Sensitivities of objectives to the system will depend on the objective function values. Through many tests, we found that energy cost and pressures violation are the most sensitive to the system because these two objectives give vary large numbers. We consider the energy cost minimization (objective 1) and pressure violation minimization (objective 6) are the first priorities. The second priorities are the tank water residence time minimization and daily tank level deviation minimization. The system reliability is less important under current consideration. Moreover, the pressure violation minimization will indirectly increase the system reliability. The groundwater drawdown
is important, but is not sensitive to the system. Moreover, the energy cost minimization will have a considerable impact on the drawdown minimization. Therefore, this study does not focus on the drawdown minimization. A systematic search was performed by varying one weight at a time and keeping other weights constant. We found a set of reasonable weight and penalty values listed in Table 5.2 as Case 1, which gives the maximum decreases 62% with respect to the total weighted objective value using the original well pump operations. Again, Table 5.2 couldn’t be achieved without the parallel GA.

Table 5.2: Optimized Result, Case 1

<table>
<thead>
<tr>
<th></th>
<th>Weights</th>
<th>Optimized</th>
<th>W*Optimized</th>
<th>% Dec./Inc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Energy (kW-h)</td>
<td>500000</td>
<td>173840.11</td>
<td>86920054688</td>
<td>61.88</td>
</tr>
<tr>
<td>Daily Tank Level Deviation (m)</td>
<td>500</td>
<td>350.35</td>
<td>175176.58</td>
<td>-209.18</td>
</tr>
<tr>
<td>Total Drawdown (m)</td>
<td>25</td>
<td>2.06</td>
<td>51.57</td>
<td>24.97</td>
</tr>
<tr>
<td>Residence Time (h)</td>
<td>100</td>
<td>1834.16</td>
<td>183416.08</td>
<td>14.08</td>
</tr>
<tr>
<td>System Reliability (%)</td>
<td>10</td>
<td>74.86</td>
<td>748.63</td>
<td>-22.28</td>
</tr>
<tr>
<td>Zone 1 Pressure Violation (m)</td>
<td>500</td>
<td>18952.41</td>
<td>9476202.83</td>
<td>17.86</td>
</tr>
<tr>
<td>Zone 2 Pressure Violation (m)</td>
<td>5000</td>
<td>2967.31</td>
<td>14836547.79</td>
<td>41.47</td>
</tr>
<tr>
<td>Total Weighted Objective Value</td>
<td></td>
<td></td>
<td>86944726831</td>
<td>61.86</td>
</tr>
</tbody>
</table>

The total energy obtained in Case 1 is 173,840 kW-h, which is substantially less than that in the original model. The minimum and maximum hourly energy consumed by the well pumps are 600 kW-h and 2,400 kW-h respectively over a week as shown in Figure 5.1. Unlike the original energy variation, the optimized energy values follow a daily periodic energy consumption pattern with peak value occurring at 6:00 pm each day. The amount of total optimized pumping rates from all the 23 wells is 770,268 m³/week, which decreases 49% of the total pumping rates in the original model. These considerable amounts of reductions in the total energy and pumping rates result in the decrease in the system reliability from 96% in the original model to 75% in Case 1. As
mentioned earlier we do not concentrate on maintaining unnecessarily high system reliability; therefore, a 75% reliable system is acceptable as long as we are able to reduce the drawdown and tank residence time.

![Figure 5.1: Original, Extreme, and Case 1 Energy values](image)

The tank levels have been considerably lowered with respect to the original high tank levels as shown in Figure 5.2. The total daily tank level deviation has increased from 113.32 m to 350.35 m due to not keeping tank full or empty as in the original model and extreme cases. Most of the tanks follow a daily periodic pattern, which avoids developments of high pressures at nodes in the vicinity of the tanks. However, some tank levels are not periodic due to non-periodic or non-constant booster bump operations to the tanks.
The reduction in pumping rates has a direct impact on the total drawdown, which has decreased from 2.75 meter in the original model to 2.06 meters in Case 1 as shown in Figure 5.3. A maximum drawdown of 0.2 meter occurs at the “West Peco Road” well site in the Middle Alluvial Unit. Most of the drawdown in Case 1 and original model occurs in the northeastern part of the City. Although this study does not focus on minimizing drawdown, the total drawdown values have reduced by 25%. Moreover, the reduction in pumping rates also contributes to the decrease in the tank water residence time from 2,135 hours (97.8%) in the original model to 1,834 hours (84%) in Case 1, which may indirectly improve the water quality in the system. The pressure violations also have
decreased to 18,952 meters in zone 1 and 2,967 meter in zone 2, showing a reduction of 18% and 42% for zone 1 and 2 to the original model, respectively.

Figure 5.3: Original and Case 1 Drawdown Distribution for Middle Alluvial Unit
The well production management model has a number of conflicting objectives, which needs thorough analyses to obtain compromising solutions according to decision makers. However, in this study we focus on the compromising solutions by changing the weights between the energy cost and zone 1 pressure violation because these two objectives are our top priorities and they are highly sensitive to the system performance. In Case 2, the energy objective weight is decreased from 500,000 to 100,000 and penalty for zone 1 pressure violation is increased from 500 to 1,000. In Case 3, the energy objective weight is further decreased to 50,000 and penalty for zone 1 pressure violation is further increased to 10,000.

As shown in Figure 5.4, the total energy values in Cases 2 and 3 are 224,103 kW-h and 354,265 kW-h, showing a decrease of 51% and 22%, respectively, with respect to the original model. Due to decrease in the energy weight the energy values in Case 2 and 3 are higher than the Case 1 values by 50,263 kW-h and 180,425 kW-h, respectively. Also, the minimum and maximum hourly energy consumptions for Cases 2 are 700 kW-h and 2,500 kW-h, and for Case 3 are 1,650 kW-h and 3,800 kW-h, respectively, over a week. The pumping rates also increase in Cases 2 and 3 as shown in Figure 5.5. The total pumping rate in Cases 2 and 3 are 884,264 m$^3$/week and 969,634 m$^3$/week, showing 41% and 35% decrease, respectively, with respect to the original pumping rate.

Higher pumping rates result in an increase in the system reliability from 75% in Case 1 to 78% and 85% in Case 2 and 3, respectively. Also, due to increased pumping rate the tank water levels have been increased; and therefore, shown in Table 5.3 the total daily tank level deviation has decreased. The deviations are slightly lower than that in Case 1, but are still larger than the original model.
Figure 5.4: Daily Energy Consumption for Original, Extreme, and Cases 1, 2, and 3

Figure 5.5: Daily Pumping Rate for Original, Extreme, and Cases 1, 2, and 3
The maximum drawdown values for Cases 2 and 3 are still recorded at the “West Peco Road” well site as 0.25 meters and 0.30 meters respectively, at, in the Middle Alluvial Unit. The drawdown distributions are shown in Figure 5.6. The total drawdowns have increased in Cases 2 and 3 in comparison to Case 1. The total drawdown values are 2.20 meters and 2.54 meters, showing a decrease of 19.95% and 7.49% in Cases 2 and 3, respectively, with respect to the original. Similar to Case 1 and original, most of the drawdown occurs at the northeastern part of the city.

Table 5.3: Optimized Results for Case 2 and Case 3.

<table>
<thead>
<tr>
<th></th>
<th>Case 2</th>
<th></th>
<th>Case 3</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Weights</td>
<td>Optimized</td>
<td>% Dec./Inc.</td>
<td>Weights</td>
</tr>
<tr>
<td>Total Energy (kW-h)</td>
<td>100000</td>
<td>224103.35</td>
<td>50.85</td>
<td>50000</td>
</tr>
<tr>
<td>Daily Tank Level Deviation (m)</td>
<td>500</td>
<td>333.95</td>
<td>-194.70</td>
<td>500</td>
</tr>
<tr>
<td>Total Drawdown (m)</td>
<td>25</td>
<td>2.20</td>
<td>19.95</td>
<td>25</td>
</tr>
<tr>
<td>Residence Time (h)</td>
<td>100</td>
<td>1898.54</td>
<td>11.06</td>
<td>100</td>
</tr>
<tr>
<td>System Reliability (%)</td>
<td>10</td>
<td>78.35</td>
<td>-18.66</td>
<td>10</td>
</tr>
<tr>
<td>Zone 1 Pressure Violation (m)</td>
<td>1000</td>
<td>17881.19</td>
<td>22.50</td>
<td>10000</td>
</tr>
<tr>
<td>Zone 2 Pressure Violation (m)</td>
<td>5000</td>
<td>2698.13</td>
<td>46.78</td>
<td>5000</td>
</tr>
</tbody>
</table>

The tank water residence time is increased to 87% (1,899 hours) in Case 2 and 90% (1,965 hours) in Case 3. By increasing the penalty values, the zone 1 pressure violation shows a decrease of 22.5% and 32.77%, respectively, with respect to original model. These values show a considerable pressure violation reduction with respect to Case 1. The increased pumping rates also reduce zone 2 pressure violation to 2698.13
meter (46.78%) and 2609.8 meters (48.53%), with respect to original model. The complete optimization results for all the Cases are shown in Table 5.4.

Figure 5.6: Case 2 and Case 3 Drawdown Distribution for Middle Alluvial Unit
Figure 5.7 shows the Pareto curve, the trade-off between the energy minimization and zone 1 pressure violation minimization. Although the detailed results are not shown in this study, Case 4 and Case 5 were conducted to expand the range of the Pareto curve. Each point is the optimal solution obtained by the parallel GA and represents the non-inferior solution. Ideally, these points should lie exactly on the Pareto curve. However, the obtained optimal solution may not be the global optimal solution due to the optimization method and the problem complexity. For example, a number of parameters, e.g., crossover and mutation probabilities, in GA will affect the optimal solution. Find the optimal GA parameters are possible with large number of trial-and-error, which is not considered in this study. Nevertheless, the Pareto curve implies the global or near-global optimal solutions.

Figure 5.7: The Pareto Curve
Table 5.4: Optimization Results

<table>
<thead>
<tr>
<th>Case 1</th>
<th>Weights</th>
<th>Original</th>
<th>W*Original</th>
<th>Optimized</th>
<th>W*optimized</th>
<th>% Dec./Inc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Energy</td>
<td>500000</td>
<td>456003.00</td>
<td>2.28002E+11</td>
<td>173840.11</td>
<td>86920054688</td>
<td>61.88</td>
</tr>
<tr>
<td>Residence Time</td>
<td>100</td>
<td>2134.65</td>
<td>213465.00</td>
<td>1834.16</td>
<td>183416.08</td>
<td>14.08</td>
</tr>
<tr>
<td>Daily Tank Level Diff.</td>
<td>500</td>
<td>113.32</td>
<td>56659.27</td>
<td>350.35</td>
<td>175176.58</td>
<td>-209.18</td>
</tr>
<tr>
<td>Reliability</td>
<td>10</td>
<td>96.33</td>
<td>963.25</td>
<td>74.86</td>
<td>748.63</td>
<td>-22.28</td>
</tr>
<tr>
<td>Total Drawdown</td>
<td>25</td>
<td>2.75</td>
<td>68.74</td>
<td>2.06</td>
<td>51.57</td>
<td>24.97</td>
</tr>
<tr>
<td>Zone 1 Pres. Violation</td>
<td>500</td>
<td>23073.75</td>
<td>11536875.26</td>
<td>18952.41</td>
<td>9476202.83</td>
<td>17.86</td>
</tr>
<tr>
<td>Zone 2 Pres. Violation</td>
<td>5000</td>
<td>5070.11</td>
<td>25350555.33</td>
<td>2967.31</td>
<td>14836547.79</td>
<td>41.47</td>
</tr>
<tr>
<td>Total (Obj. Value):</td>
<td>2.28039E+11</td>
<td>86944726831</td>
<td>61.87</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Case 2</th>
<th>Weights</th>
<th>Original</th>
<th>W*Original</th>
<th>Optimized</th>
<th>W*optimized</th>
<th>% Dec./Inc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Energy</td>
<td>100000</td>
<td>456003.00</td>
<td>45600300000</td>
<td>224103.35</td>
<td>22410350000</td>
<td>50.85</td>
</tr>
<tr>
<td>Residence Time</td>
<td>100</td>
<td>2134.65</td>
<td>213465.00</td>
<td>1898.54</td>
<td>189854.00</td>
<td>11.06</td>
</tr>
<tr>
<td>Daily Tank Level Diff.</td>
<td>500</td>
<td>113.32</td>
<td>56659.27</td>
<td>333.95</td>
<td>166977.06</td>
<td>-209.18</td>
</tr>
<tr>
<td>Reliability</td>
<td>10</td>
<td>96.33</td>
<td>963.25</td>
<td>78.35</td>
<td>783.50</td>
<td>-18.66</td>
</tr>
<tr>
<td>Total Drawdown</td>
<td>25</td>
<td>2.75</td>
<td>68.74</td>
<td>2.20</td>
<td>55.03</td>
<td>19.95</td>
</tr>
<tr>
<td>Zone 1 Pres. Violation</td>
<td>1000</td>
<td>23073.75</td>
<td>23073750.53</td>
<td>17881.19</td>
<td>17881866.49</td>
<td>22.50</td>
</tr>
<tr>
<td>Zone 2 Pres. Violation</td>
<td>5000</td>
<td>5070.11</td>
<td>25350555.33</td>
<td>2698.13</td>
<td>13490661.36</td>
<td>46.78</td>
</tr>
<tr>
<td>Total (Objective Value):</td>
<td>45648995462</td>
<td>22442064517</td>
<td>50.83</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Case 3</th>
<th>Weights</th>
<th>Original</th>
<th>W*Original</th>
<th>Optimized</th>
<th>W*optimized</th>
<th>% Dec./Inc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Energy</td>
<td>50000</td>
<td>456003.00</td>
<td>22800150000</td>
<td>354265.33</td>
<td>17713266250</td>
<td>22.31</td>
</tr>
<tr>
<td>Residence Time</td>
<td>100</td>
<td>2134.65</td>
<td>213465.00</td>
<td>1965.32</td>
<td>196532.00</td>
<td>7.93</td>
</tr>
<tr>
<td>Daily Tank Level Diff.</td>
<td>500</td>
<td>113.32</td>
<td>56659.27</td>
<td>308.25</td>
<td>154125.17</td>
<td>-172.02</td>
</tr>
<tr>
<td>Reliability</td>
<td>10</td>
<td>96.33</td>
<td>963.25</td>
<td>84.57</td>
<td>845.68</td>
<td>-12.21</td>
</tr>
<tr>
<td>Total Drawdown</td>
<td>25</td>
<td>2.75</td>
<td>68.74</td>
<td>2.54</td>
<td>63.59</td>
<td>7.49</td>
</tr>
<tr>
<td>Zone 1 Pres. Violation</td>
<td>10000</td>
<td>23073.75</td>
<td>230737505.25</td>
<td>15512.91</td>
<td>155129081.66</td>
<td>32.77</td>
</tr>
<tr>
<td>Zone 2 Pres. Violation</td>
<td>5000</td>
<td>5070.11</td>
<td>25350555.33</td>
<td>2603.89</td>
<td>13048977.20</td>
<td>48.53</td>
</tr>
<tr>
<td>Total (Objective Value):</td>
<td>2.31E+10</td>
<td>1.79E+10</td>
<td>22.73</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Case 4</th>
<th>Weights</th>
<th>Original</th>
<th>W*Original</th>
<th>Optimized</th>
<th>W*optimized</th>
<th>% Dec./Inc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Energy</td>
<td>1.00E+4</td>
<td>4.56E+05</td>
<td>4.56E+09</td>
<td>4.15E+05</td>
<td>4.15E+09</td>
<td>8.99</td>
</tr>
<tr>
<td>Residence Time</td>
<td>100</td>
<td>2134.65</td>
<td>213465.00</td>
<td>1998.55</td>
<td>199855.00</td>
<td>6.38</td>
</tr>
<tr>
<td>Daily Tank Level Diff.</td>
<td>500</td>
<td>113.32</td>
<td>56659.27</td>
<td>352.12</td>
<td>176062.18</td>
<td>-210.74</td>
</tr>
<tr>
<td>Reliability</td>
<td>10</td>
<td>96.33</td>
<td>963.25</td>
<td>87.65</td>
<td>267.16</td>
<td>-9.01</td>
</tr>
<tr>
<td>Total Drawdown</td>
<td>25</td>
<td>2.75</td>
<td>68.74</td>
<td>2.75</td>
<td>68.66</td>
<td>0.11</td>
</tr>
<tr>
<td>Zone 1 Pres. Violation</td>
<td>50000</td>
<td>23073.75</td>
<td>230737505.25</td>
<td>11617.22</td>
<td>5.81E+08</td>
<td>49.65</td>
</tr>
<tr>
<td>Zone 2 Pres. Violation</td>
<td>5000</td>
<td>5070.11</td>
<td>25350555.33</td>
<td>2306.92</td>
<td>11534607.36</td>
<td>45.50</td>
</tr>
<tr>
<td>Total (Objective Value):</td>
<td>5.78E+09</td>
<td>4.78E+09</td>
<td>17.30</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Case 5</th>
<th>Weights</th>
<th>Original</th>
<th>W*Original</th>
<th>Optimized</th>
<th>W*optimized</th>
<th>% Dec./Inc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Energy</td>
<td>1.00E+6</td>
<td>4.56E+05</td>
<td>4.56E+11</td>
<td>173651.98</td>
<td>1.7E+11</td>
<td>61.89</td>
</tr>
<tr>
<td>Residence Time</td>
<td>100.00</td>
<td>2134.65</td>
<td>213465.00</td>
<td>1830.08</td>
<td>183107.67</td>
<td>14.22</td>
</tr>
<tr>
<td>Daily Tank Level Diff.</td>
<td>500.00</td>
<td>113.32</td>
<td>56659.27</td>
<td>352.12</td>
<td>178062.18</td>
<td>-210.74</td>
</tr>
<tr>
<td>Reliability</td>
<td>10.00</td>
<td>96.33</td>
<td>963.25</td>
<td>74.02</td>
<td>740.23</td>
<td>-9.01</td>
</tr>
<tr>
<td>Total Drawdown</td>
<td>25.00</td>
<td>2.75</td>
<td>68.74</td>
<td>2.07</td>
<td>68.66</td>
<td>0.11</td>
</tr>
<tr>
<td>Zone 1 Pres. Violation</td>
<td>100.00</td>
<td>23073.75</td>
<td>230737505.05</td>
<td>19210.12</td>
<td>19210117.75</td>
<td>16.74</td>
</tr>
<tr>
<td>Zone 2 Pres. Violation</td>
<td>5000.00</td>
<td>5070.11</td>
<td>25350555.33</td>
<td>3003.89</td>
<td>15019427.79</td>
<td>40.75</td>
</tr>
<tr>
<td>Total (Objective Value):</td>
<td>4.57E+11</td>
<td>1.74E+11</td>
<td>61.82</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5.2 Conclusions

The well pump energy consumption has decreased by about 62% in Case 1 as compared to the original value which allows us to represent a saving of around $16000 for the water distribution system in a week. Including the booster pumps in the scheduling optimization would further enhance the efficiency of the system. Therefore,
the best performance of the periodic operations should consider the pump-tank-booster joint operations in order to optimize the tank water levels. The presence of conflicting objectives has been clearly identified. With increasing energy consumption the pressure violations decrease due to which residence time increases and so the reliability increases. Therefore we try to set up a compromising solution so that it able to tackle all the conflicts without introducing any potential problem in the water distribution system.

Genetic algorithm has proven to be very effective for our operational optimization problem. Over other non-linear solvers like MINOS, CONOPT, GA was fast and efficient. We managed to produce satisfactory and consistent results after optimization. Optimizing a water distribution system using GA under a parallel processing environment exposes us to a vast set of solutions and the study of their inter-relationship. Parallel processing helps us recognize the best weighting coefficients for the optimization equation. Without parallel processing we could not have produced a Pareto curve.

The drawdown minimization does not reveal the significance in Chandler production well management model due to the short-term (7 days) pumping activities. If a certain maximum level of drawdown is allowed, the drawdown maximization can be investigated in the future study for withdrawing more groundwater from the aquifer without compromising the environmental detriment. The Chandler production well management model is able to be expanded to consider the conjunctive use of the surface water and groundwater. The current Chandler production well management model considers that the surface water is a fixed input to the system. If more groundwater water is allowed to be withdrawn, Chandler production well management model can be expanded to investigate how much surface water can be reduced.
References


Vita

Vineet Katiyar was born on March 10th, 1983, in Uttar Pradesh, India, to Mrs. Vimla Katiyar and Mr. Vinod Kumar Katiyar. He graduated from Indian Institute of Technology Roorkee, with a Bachelor of Technology in Civil Engineering degree in the year 2005. He came to the United States to join the graduate school at Louisiana State University Agricultural and Mechanical College in the Department of Civil Engineering in fall 2005. He defended his master’s thesis on December 4th, 2006. He will be awarded the degree of Master of Science in Civil Engineering at the May 2007 commencement.