APPLICATION AND CONCEPTUAL DEVELOPMENT OF GENETIC ALGORITHMS FOR OPTIMIZATION IN THE WATER INDUSTRY

SJ van Vuuren · PG van Rooyen · JE van Zyl · M van Dijk

WRC Report No. 1388/1/05

Water Research Commission

APPLICATION AND CONCEPTUAL DEVELOPMENT OF GENETIC ALGORITHMS FOR OPTIMIZATION IN THE WATER INDUSTRY

Report to the
Water Research Commission

by

SJ van Vuuren, PG van Rooyen, JE van Zyl and M van Dijk

UNIVERSITY OF PRETORIA
Department of Civil & Biosystems Engineering

WRC Report No. 1388/1/05
ISBN NO. 1-77005-318-2 MARCH 2005

Disclaimer

This report emanates from a project financed by the Water Research Commission (WRC) and is approved for publication. Approval does not signify that the contents necessarily reflect the views and policies of the WRC or the members of the project steering committee, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

Printed by Silowa Printers: 012 804 1164

Executive Summary

Background and motivation

In the previous study, WRC Project K1144 – "POTENTIAL APPLICATION OF GENETIC ALGORITHMS IN THE WATER INDUSTRY", the application of Genetic Algorithms was reviewed. Findings from that study indicated that:

- GAs are not used to its full potential in the optimisation in the water industry in South Africa.
- Potential applications of the technique within the South African context are:
  - Hydrology and water resources assessment.
  - Network optimisation.
  - Optimisation of rehabilitation, extension and upgrading of distribution networks during the planning and design phase.
  - Operation and maintenance scheduling
- Little formal teaching on GAs is included in the curriculum of civil engineering.
- Feedback from Rand Water reflected the need for the development of software utility programs that can be used in practice and stimulate the further exploitation of this technique.
- The pipeline diameter optimisation program that had been developed under this study was well accepted in practice.

Objectives of this study

This study evaluated the application of genetic algorithms in the optimisation of different components of water supply projects and conceptually developed the procedures for the implementation thereof.

Based on the available literature study, as well as the feedback from water supply authorities, the need for the application of GAs as an optimisation technique in the water industry was defined. The potential applications of GAs in the water industry in South Africa are:

- Hydrology and water resources assessment.
- Network optimisation.
- Optimisation of rehabilitation, extension and upgrading of distribution networks during the planning and design phase.
- Operation and maintenance scheduling

This study objective was to provide the **conceptual development of procedures to implement GAs** as an optimisation technique for water resources assessment and network optimisation.

Methodology

The focus of this study was to conceptualise the optimisation problems that were identified and to conceptually develop the required procedures for the implementation of GAs in these areas in a follow-up study.

Literature on optimisation techniques was reviewed, problem variables and interdependence were identified and the GA operators for these problems were conceptually described.

In the case of the network optimisation the GA procedures were coded and software was developed and tested for network optimisation. The GA Water Utility programs, GANEO, can be used for the assessment and optimisation networks (Rehabilitation, replacement, placement of parallel pipes or the assessment of new networks).

Results

The use of GA has been conceptually developed for water resources assessment where the WRYM (Water Resources Yield Model) and WRPM (Water Resources Planning Model) are used. Conceptual models for applying the GA procedure to network optimization and operational optimization problems have also been developed.

A GA front-end interface (see **Figure i**) for the public domain software, EPANET, has been developed that optimizes water distribution systems. The results of the developed optimization procedure were tested against benchmark problems that have been reviewed by numerous International Researchers. The results are extremely promising and the optimization process's computational time was reduced (**Figure ii**).

Figure i: Water distribution network optimization software (GANEO)

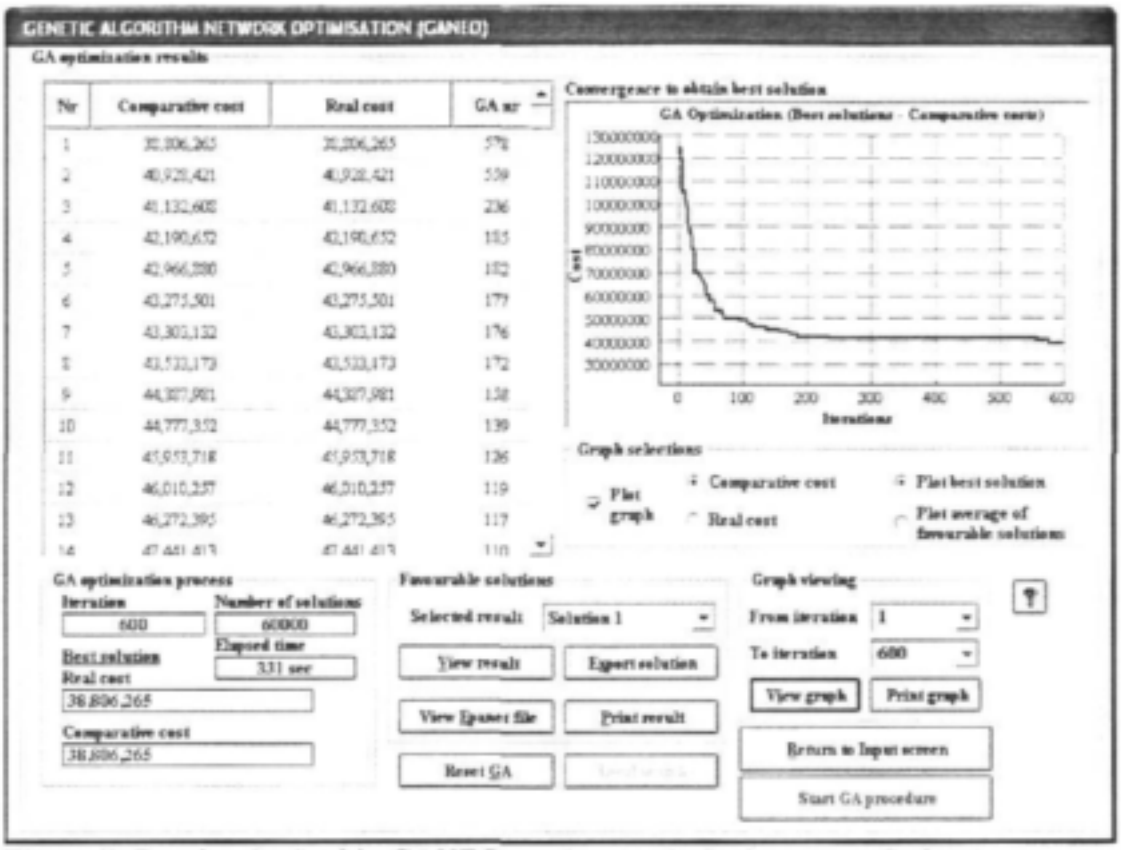


Figure ii: Results obtained in GANEO (various optimized system solutions)

This inclusion of the GA procedure for the network optimisation provides the opportunity for the optimisation of new networks, refurbishment of existing networks or the evaluation of the extension to existing networks.

International benchmark problems were analysed with the utility, in all cases providing the same or an improved optimal solutions in a much shorter computational time.

The GANEO software is capable to solve the following objectives for any complex network:

- Find a solution that meets the minimum pressure requirements at every node,
- Find a solution that meets the maximum velocity criteria set for each pipe, and
- Minimize the capital cost

Recommendation

In this study the conceptual procedures for GAs has been developed for:

- Pump optimisation and operational scheduling,
- Water Resources assessment and
- Water distribution networks.

These developments created the opportunity of conceptualising the problems and now require the extension of the study to develop the utilities to implement these concepts and create interfaces with existing software.

It is recommended that the following aspects should be attended to:

- Conceptual development of a genetic algorithm multi-objective optimisation model for the optimisation of water distribution systems specifically for rural developments, combining design and operational objective functions, including Life Cycle Costing analysis of networks.

- Utilization of the developed GA Water Utility Programs in the design of a new water distribution system in South Africa to illustrate/demonstrate its applicability and benefits.

- Promote the inclusion of GA techniques in the graduate program.

- The GA software should be used to enhance the understanding of GA's and transfer knowledge to the water sector.

- Development of a GA application for the optimisation of the operating rules between reservoirs.

- Review and testing of parameter values and alternative operators to improve the performance of the GA procedure (mutation percentage, population size, crossover methods etc.).

- Develop a generic inter reservoir optimisation model.

The software program can be downloaded from either the WRC or the University of Pretoria web sites (links below).

WRC - http://www.wrc.org.za
University of Pretoria - http://www.up.ac.za/academic/civil/divisions/water/ga.html

ACKNOWLEDGEMENTS

The research presented in this report was undertaken as a project funded by the Water Research Commission, whose support is greatly acknowledged. Based on the findings of the first study, conceptual procedures have been developed as part of this study for a number of potential areas where GAs may be used.

These developments sourced on the leadership of the Chairman, Mr Jay Bhagwan and the contribution of the Steering Committee.

Mr JN Bhagwan - Water Research Commission
Mr R Dube - Water Research Commission
Mr B Valentyn - Eastern Cape Technikon
Dr JG Ndiritu - University of Witwatersrand
Prof GGS Pegram - University of Natal
Mr DDP Trebicki - Rand Water

The input from the project team, their commitment and support is acknowledged. They are: Mr P G van Rooyen, Prof JE Van Zyl, and Mr M van Dijk.

A number of students have conducted specific developments as part of their formal studies, which contributed to this project. They are:
Mr M van der Merwe (UP) and
Mr J Borthwick (RAU)

S J van Vuuren (Project Leader)

APPLICATION AND CONCEPTUAL DEVELOPMENT OF GENETIC ALGORITHMS FOR OPTIMIZATION IN THE WATER INDUSTRY

Table of Contents

LIST OF TABLES

LIST OF FIGURES

WRC Project K5/1338

Application and Conceptual Development of Genetic Algorithms for Optimization in the Water Industry

1. INTRODUCTION

1.1 Background

Genetic algorithms are recognised as a powerful optimisation technique with potential use in the water environment. In a previous research study, the potential application in the Water Industry in South Africa has been indicated (**Potential Application of Genetic Algorithms in the Water Industry, WRC Report 1144/1/01**).

Previous studies on the subject identified the following potential applications of GAs:

- WRSM2000 model calibration,
- Water resources assessment
- Network analyses and scheduling (based on EPANET – Free Network analysis software).

During the Inauguration Meeting, the Project the Steering Committee requested that the available software should be briefly reviewed to determine the accessibility of potential users to the products. This was added as a focus of the project. Furthermore, reference was made that the WRC intend to fund a Research Project that were to update the WRSM 2000 models and to include the automated calibration for which the use of GA's have been identified in a previous study (WRC Report 1144/1/01). This resulted that the aspect related to the calibration optimisation of the WRSM 2000, was excluded from this study.

The level of detail of this study was to conceptually develop procedures for the incorporation of GA's to optimise a number of water related systems.

1.2 Objectives of the Research

The objective of this study was to extend the understanding and review the implementation of the GA-procedures in a number of priority water related fields in South Africa and to transfer the knowledge to potential users.

The objective was to be achieved through:

- Conceptualisation of procedures to use GAs in the assessment of the yield of surface water resources and optimise networks and distribution systems and
- To present a Course that will stimulate the use of GAs.
- *In* addition the research team included further work on the development of utility software (Freeware) that could be used to optimise water distribution networks. This was an addition to the task list.

1.3 Methodology

The focus of this study was to conceptualise the optimisation problems that were identified and to conceptually develop the required procedures for the implementation of GAs in these areas in a follow-up study.

Literature was reviewed of optimisation techniques, the problem variables and interdependence were identified and the GA operators for these problems were conceptually described.

In the case of the network optimisation the GA procedures were coded and software was developed and tested for network optimisation. The GA utility program, GANEW, can be used for the assessment and optimisation networks (Rehabilitation, Replacement, Placement of parallel pipes or the assessment of new networks).

1.4 Research Outputs (Deliverables)

The following research products emanated from the study and are captured in this report:

- Technology transfer
- The extension of the pipeline optimisation program (GAPOP)
- Conceptual development of procedures to use GAs for:
  - water resources assessment and
  - network analyses and pump scheduling
- Development of the GANEW Software utility program that is capable of solving numerous optimisation objectives of water distribution networks.

1.5 Structure of the Report

This report contains the developments and findings of this research conducted to evaluate the application and reflect the conceptual development of the application of genetic algorithms in the water industry. The Report has been structured to provide an overview of the previous work, a description of the working of GA's and the conceptual development of the application of GA's in a number of focus areas. These aspects are covered in the following chapters:

Chapter 1:	Introduction (This Chapter) – Provides an introduction to the Study
Chapter 2:	An Overview of the findings of previous studies
Chapter 3:	An Overview of the workings of GA's – This chapter reflects the working of the GA procedure and shows the different GA types.
Chapter 4:	Available GA Software Tools – The availability and development of software programs are discussed
Chapter 5:	Development of a Conceptual Model for the Optimisation of Operating Rules for Water Resources – This chapter reflects the variables that should be included in the optimisation of the yield from surface water resources
Chapter 6:	Conceptual Development of Genetic Algorithms for Pipe Network Optimisation – The aspects that should be considered and included for the optimisation is discussed.
Chapter 7:	Conceptual development of a Method for Operational Optimisation – The various focus areas in the optimisation of operational control are discussed in this chapter.
Chapter 8:	Extensions of the GA Utility Programs
Chapter 9:	Conclusion
Chapter 10:	Recommendations
Chapter 11:	References
Appendix A:	Contains a step-by-step hand calculation of optimising a water distribution network utilizing a genetic algorithm

2. AN OVERVIEW OF THE FINDINGS FROM PREVIOUS STUDIES

2.1 Introduction

In a previous study (Van Vuuren et al., 2001) the potential application of GA's in South Africa was reviewed, against the background of an ever expanding computational capacity which far exceeds the capacity of the "instructor" to define options to be evaluated when optimisation, as well as the South African Government's objective to provide "water for all" in the near future.

The variables that could influence the selection and hence the final cost of water systems include:

- The high variance in rainfall and runoff,
- the availability of alternative water supply,
- the demand pattern variability,
- the operational ability of the system,
- the maintenance requirements,
- the running cost-especially power cost and
- the affordability and willingness to pay for services.

The determination of the optimal selection of system components requires techniques that can be employed to assist the decision-maker in finding the appropriate solution within the environment of all the possible solutions (Solution space).

2.2 The potential use of GA's

Genetic Algorithms (GAs) have been developed (Holland, 1975) to assist in searching through complex solution spaces for the optimum solution. GAs have been applied as search techniques for various engineering problems such as, structural design optimisation, water distribution network evaluation, pump scheduling, hydrological runoff predictions and resource utilisation, this technique has however not been generally used in South Africa.

The application of genetic algorithms in the optimisation of different components of water supply projects include:

- pump selection and scheduling
- optimal pipeline diameter selection
- valves and surge alleviating devices selection
- management of water supply projects
- alterations and extensions required in the upgrading of the capacity of water supply infrastructure.

The outset of the previous study (Van Vuuren et al. 2001) was, that should the benefit of GAs be proven, further research would be conducted to develop the application of genetic algorithms in the water industry, which should also be used as educational tools.

2.3 Findings of previous study

During the previous study (van Vuuren et al., 2001), the procedures of Genetic Algorithms were discussed and the Potential Application of Genetic Algorithms in the Water Industry was evaluated. It was concluded that:

- GAs are not used to its full potential in the optimisation processes in the water industry of South Africa.
- Potential applications of the technique within the South African context are:
  - Hydrology and water resources assessment.
  - Network optimisation.
  - Optimisation of rehabilitation, extension and upgrading of distribution networks during the planning and design phase.
  - Operation and maintenance scheduling
- A minimum amount of teaching on GAs has been included in the curriculum of civil engineering.
- Feedback from Rand Water reflected the need for the development of software utility programs that can be used in practice and stimulate the further exploration of this technique.
- The pipeline diameter optimisation program (GAPOP) that has been developed under this study is well accepted in practice and could be extended.

2.4 Recommendation emanating from previous study

Based on the findings, it was recommended that:

- A course be presented on the use of GAs in the water industry. World renowned leaders in this field must be invited to participate in such a course in South Africa
- Procedures for the implementation of the technique in:
  - Network optimisation.
  - Water resources assessment and
  - Operational scheduling.

  needs to be defined.

- The software that has been developed should be improved and used to transfer knowledge and educate students in the use of GAs.

It was also recommended that the following aspects should be addressed and researched:

- Opportunities for the inclusion of this technique in the graduate program should be promoted.
- Procedures for the implementation of the technique in:
  - Network optimisation,
  - Water resources assessment and
  - Operational scheduling,

  needed to be conceptually defined and developed for implementation.
- The software that was developed (GAPOP), should be extended and distributed through the WRC with the objective to enhance the use and transfer of knowledge regarding the value of GAs in the water sector.
- Application of GA to optimise the operating rules between reservoirs in the WRYM.
- Conceptual development of the procedures for the coding and "fitness" functions for Water Distribution Networks.

It has also been indicated that research in subsequent phases should be considered and that the following aspects needed to be addressed:

The potential applications of GAs in the Water Resource field to improve and automate the optimisation of operating rules of water resource systems should be developed. An automated optimisation tool would reduce the time required by the model user and has the potential benefit of reducing pumping costs, which could amount to millions of rands in savings. Optimising the sizes of proposed infrastructure components in combination with optimising the operating rules could also have significant benefits.

It is therefore proposed that further research and development be undertaken to apply GAs in the optimisation of the operating rules and sizing of proposed water resource infrastructure. Although the ultimate aim of such an optimisation tool would be to simulate the large integrated systems, it is proposed that the research and development be carried out in incremental steps to first prove the viability of the development.

3. OVERVIEW OF THE WORKING OF GA's

3.1 Introduction

According to Michalewicz (1994), Genetic Algorithms (GA's) can basically be described as artificial evolution search methods based on the theories of natural selection and mechanisms of population genetics. GA's emulate nature's optimization technique of evolution, based on:

- survival and reproduction of the fittest members of the population
- the maintenance of a population with diverse members
- the inheritance of genetic information from parents
- and the occasional mutation of genes

Genetic algorithms are best suited to solving combinatorial optimization problems that cannot be solved practicably using more conventional methods, as discussed in the previous section.

GA optimization is a powerful approach to identifying a near optimal solution (Goldberg and Kuo, 1987). However, for a GA optimization to run efficiently, careful parameter tuning and configuration is needed. Inappropriate penalty levels will distort the results away from the end user's perception of optimum.

Too high or low a rate of genetic interchange ('crossover' and 'mutation') results in degeneration to a random search or stagnation, causing a failure to converge on the global optimum. Because of the complexity involved in setting up a GA to operate effectively, this form of optimization is not well suited to trivial problems, or problems for which the number of possible solutions is small.

The benefits of GA's stem from their ability to converge rapidly on an optimal or near-optimal solution, having analyzed only a small fraction of the number of possible solutions available. For large problems such as those typically associated with networks, the exhaustive analysis of all options is unlikely to ever be feasible.

In a recent water distribution network expansion and reinforcement problem analyzed in the Centre for Water Systems, University of Exeter, using GA's, there were approximately 300 pipe links, each of which could adopt any of 14 diameters (including the option of not employing a pipe). This presents a problem size of 14^{300} or 6.89×10^{343}. It would take an estimated 4.6 billion years to analyze each possible solution. In this instance, the GA optimization search was able to converge on the lowest cost solutions by carrying out several hundred thousand evaluations.

3.2 Concepts and Methodology

A GA evolves optimal solutions by sampling from all the possible solutions. The best of these solutions are then combined, using the genetic operators of crossover and mutation, to form new solutions. This process continues until some termination condition is fulfilled. A flow diagram of the basic GA methodology is given in **Figure 3.1**.

Figure 3.1: The basic genetic algorithm methodology

3.3 Terminology

Much of the terminology used in GAs is borrowed from natural science. A *population* is made up of a number of *organisms* (individual solutions), each of which is defined by its *chromosomes* (set of decision variables). Chromosomes are made of units called *genes* (individual decision variables) arranged in linear succession where every gene controls one or several characteristics. The position of a gene is called its *locus* and the actual value of the gene its *allele*.

Each organism represents a potential solution to the optimisation problem. Successively-created populations are called *generations* and these generations are created by selecting organisms from the current population to be *parents*, from which *children* are produced to form the next generation.

Encoding. Encoding refers to the way that chromosomes are represented in computer code. This can be done in a number of ways. The simplest approach uses binary encoding where each chromosome consists of a string of binary digits (*binary encoding*), i.e. a sequence of ones and zeros. Binary encoding is relatively simple to apply and the vast majority of initial experimental GAs used this method (Goldberg, 1989). However, binary encoding may not be the best technique in all circumstances. Other representations such as *integer encoding* and *real number encoding* have been successfully applied to various problems (Davis, 1991). In these representations, each chromosome consist of a string of integer or real numbers representing the variable values, for example pipe diameter, pump status or pipe roughness.

For binary coding, a gene normally consists of a string of bits. For non-binary coding, a gene normally corresponds to one variable regardless of the number of possible values required for a particular gene.

Population. The complete set of organisms at any given stage of the optimisation represents a single generation. Typically, the population size will vary between 50 and 150 organisms (Robertson and Willet, 1995). An initial population of organisms is generated randomly at the beginning of an optimisation run. The two main ways in which subsequent generations are produced are:

- Creating sufficient children to replace the original population, a procedure called *generational replacement*; or

- Creating sufficient children to replace only a few (normally the worst) children, a procedure called *steady-state replacement*.

Fitness Function. For each organism in the population, a fitness value is calculated from the objective function. The fitter a chromosome is the more likely it is to survive and then to pass parts of itself to future generations. In nature, some organisms will survive merely because they are lucky, but on average the offspring will tend to derive from the fitter parents, thus increasing the average fitness of the population.

Genetic drift, is the random change in allele frequency from one generation to the next due to chance fluctuations (as opposed to due to selection pressure).

Objective value, is similar to the fitness, and in many cases the value of the fitness is simply assigned as the objective value. The objective value is the number, which either maximized or minimized by the GA. In a minimization problem the smallest objective value will reflect the most optimal solution.

To incorporate soft constraints into a genetic algorithm, *penalty functions* are used. A cost or penalty is calculated, depending on the extent to which the constraints are violated. This cost is then added to the objective function. Several types of penalty functions such as linear, quadratic and cubic functions can be used.

The efficiency of a GA will be determined in large part by the extent to which the fitness function accurately characterizes the function to be optimised. Therefore, it is often necessary to map the underlying natural objective function to a fitness function form. The objective function value may be used unchanged or may be subjected to windowing, scaling or normalization procedures in order to improve the GA performance. Some of the common remapping methods are (Davis, 1991):

- *Windowing* using the value of the least-fit member of the population. Each member of the population is given a fitness equal to the amount by which its value exceeds the value of the worst member. A variation is to assign a minimum value to prevent the possibility that the worst member has no chance at all of reproduction.

- *Linear normalization* (A, B) where A is a constant value and B the rate of decrement. The chromosomes are sorted by decreasing evaluation where the fittest chromosome is given fitness equal to A.

3.4 Types of GA's

When considering genetic algorithms there are three main general types that need to be discussed (Zebulum, Pacheco and Vellasco, 2002):

- *Generational GA*: is the most classic type of GA, and involves replacing the whole GA population at each generation

- *Steady State GA*: in this type one member of the population is changed at a time. To perform selection a member of the population is chosen according to its fitness. It is copied and the copy mutated. A second member of the population is selected which is replaced by the mutated string. In crossover two members of the population are chosen, a single or more children can be created which would then replace members of the population. There are different replacement techniques:

  - Replace worst
  - Randomly chosen member
  - Using negative fitness

  There are two types of operation for a GA, steady-state and generational. The major difference between steady-state and generational GA's is that for each P members of the population generated in the generational GA there are 2P selections. Consequently the selection strength and genetic drift for a steady-state GA is twice that of the generational GA. The steady-state GA, therefore, appears twice as fast although it can lose out in the long term because it does not explore the landscape as well as the generational GA.

- *Generation Gap*: this is a cross between the generational and steady state types. A new population is generated and added to the old population before the selection procedure occurs. The genetic drift caused by the generation gap can then be calculated for the step, the larger the generation gap the smaller the genetic drift.

The purpose of such approaches is to scale the evaluations so that a super-fit organism does not dominate the population and exclude lesser organisms, thus causing premature convergence, and accentuating the differences in fitness values that are very similar.

Various types of GAs have been developed to improve their efficiency on specific types of problems. Some of the more common types of GAs are listed below:

- *Hybrid GA.* GAs are often used in combination with a problem-specific or local search procedure, especially in commercial applications (Goldberg and Voessner, 2000). The goal of using a problem specific search method is to improve the efficiency of the GA, either in terms of the time required to find a good solution, or the quality of the solution found. Early hybrid GAs were introduced by Smith (1985) and Grefenstette et al (1985), and are commonly used today in serious GA applications (Goldberg and Voessner, 2000).

  Hybrid GA searches generally work by employing the GA first to search for the global optimum, or more probably, the attraction basin of the global optimum. The local search method is then employed to find the optimum itself (Van Zyl, Savic and Walters, 2003b).

- *Multi-objective GA (MOGA).* In a multi-objective GA, optimisation is not done for a single objective (e.g. cost), but for multiple, often-conflicting objectives (e.g. cost and reliability). In multi-objective GAs, there is no single optimal solution. The interaction among different objectives gives rise to a set of compromised solutions, largely known as the trade-off, non-dominated, non-inferior or Pareto-optimal solutions (Savic, 2002).

- *Co-operative co-evolutionary genetic algorithm (CCGA).* CCGAs do not utilise a single population, but rather a number of sub-populations in the GA process. Each sub-population in the CCGA represents a variable or a part of the problem that needs to be optimised. A combination of an individual from each sub-population leads to a complete solution to the problem where the fitness value of the complete solution is calculated in the usual way.

- *Fuzzy GA.* In fuzzy GAs, fuzzy logic and genetic algorithms are integrated to solve a problem.

3.5 Operators and Parameters

The different operators and parameters typically used in genetic algorithms are:

- Parent selection/generation
- Crossover
- Mutation and
- Replacement

These operators are discussed below.

3.5.1 Parent Selection

In parent selection, organisms are selected from the population for the purpose of producing offspring. The GA has to give a higher probability of reproduction to the fitter members of the population, i.e. to members representing a better solution. Various parent selection techniques exist. The most common ones are briefly discussed below:

- *Uniform random.* Parents are randomly chosen from the population.

- *Roulette wheel.* Each organism is assigned a probability of selection according to the ratio of its fitness over the total finesses of the population.

- *Stochastic remainder selection without replacement (SRSWR).* Works similarly to roulette wheel selection. The ratio of each organism's fitness to the average fitness of the population is calculated. For each organism where this ratio is above one, a number of clones equal to the whole number of the ratio are placed in the selection pool. The remaining organisms to make the size of the chosen set equal to that of the population are chosen from these remainder values using roulette wheel selection.

- *Tournament.* Involves the random selection of a designated number of organisms from the population, who compete in a tournament to determine the organisms selected for reproduction. Tournament selection in its simplest form consists of a binary tournament involving the direct competition of two organisms for survival by the comparison of their fitness function, with the fitter of the two surviving.

- *Rank biased.* Uses the ranking of an organism to bias the selection. The probability of an organism being selected is higher for higher-ranking organisms.

3.5.2 Crossover

Crossover is the component that makes GAs different from other evolutionary algorithms. The function of crossover is to produce child organisms that share the properties of their parents. Various crossover techniques can be applied, including:

- *One or multiple point crossover.* A point or points on the chromosome are randomly selected, and sections of the chromosome on different sides of the points are exchanged to create child organisms. Three-point crossover is schematically shown in **Figure 3.2**.

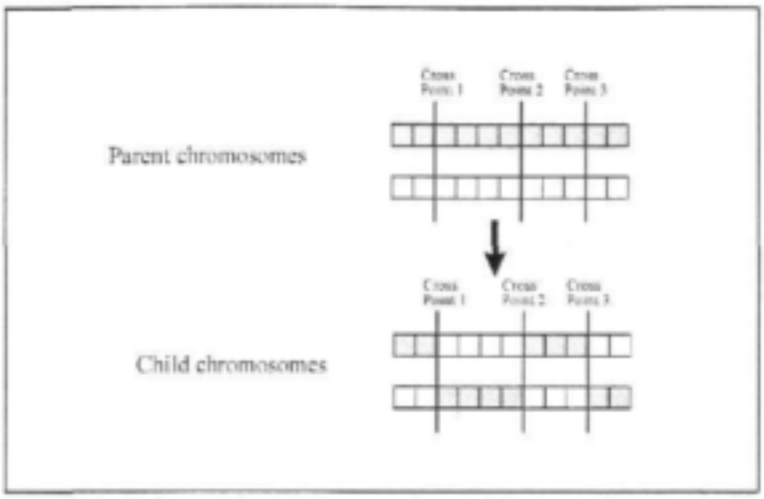


Figure 3.2: Schematic representation of three point crossover

- *Uniform random.* Uniform crossover differs from multi-crossover in that each bit is considered as a section. A decision string, known as a mask, is randomly generated deciding whether the bit should be placed into the first child or the second child.

- *Arithmetic.* This is a real type coding crossover operator. The child's genes are calculated as randomly weighted sums of the respective genes in the parents.

3.5.3 Mutation

Mutation is the final operator in the genetic process and is performed after the crossover operator. Mutation makes a random change to a gene value or replaces it with a randomly generated value. It operates on single chromosomes and its main benefit lies in reintroducing diversity into a population whose members have become very similar. Mutation can be performed on a bit by bit for binary encoding (or on a gene by gene basis for non-binary encoding), where each bit is altered with a small probability, usually between 0.001 and 0.1.

Alternatively, a chromosome is selected for mutation using a high probability (usually between 0.80 and 1) and a random gene on the chromosome altered. If the probability of mutation is set too high, the GA degenerates into a random search process.

3.5.4 Replacement

In a steady-state GA, a strategy has to be selected to replace organisms in the population with newly generated child organisms. It is not always the best strategy to replace the weakest member of the population as this can reduce the population diversity. Some of the replacement strategies are:

- *Weakest.* The member in the population with the lowest fitness value is replaced, if its fitness is lower than that of the child organism.

- *First weaker*. The first member found in the population with a lower fitness than the child organism is replaced.

- *Random*. A random member of the population is replaced.

3.5.5 Termination Conditions

Following the selection, crossover and mutation operators and introduction of the new child organisms into the population, the process is repeated until an appropriate termination condition is met. The simplest technique is to use a fixed number of generations or iterations, but often it is not known how many generations might be appropriate for a particular problem. Alternative termination conditions include when:

- *Complete convergence* has occurred, when the variability in the population's fitness values is below a certain value, or all the chromosomes are identical to each other (steady-state GA only) or;

- *No improvement* in the fitness value of the best chromosome has occurred in some fixed number of the previous generations.

Some experimentation is normally required to establish appropriate convergence criteria. Besides convergence criteria, premature convergence is another issue. It occurs generally early in the process when 'relatively good' organisms are dominating the rest of the population. By dominating all the other chromosomes, those 'relatively good' ones are involved in almost every selection and recombination and cause all organisms to become similar to each other. The GA will thus be trapped in a local optimum. There are several ways to tackle premature convergence, such as fitness scaling, increasing the population size, adding non-random chromosomes in the initial population and increasing the mutation rate (South, Wetherill and Tham, 1993).

3.6 Advantages and Limitations of Genetic Algorithms

3.6.1 Genetic Algorithm Advantages

The benefits of GAs stem from their ability to converge rapidly on an optimal or near-optimal solution, although the process has analysed only a tiny fraction of the number of possible solutions. For large problems such as those typically associated with water systems, the exhaustive analysis of all options is unlikely to ever be feasible.

Other advantages compared with conventional optimisation methods are as follows (Goldberg, 1989):

- They are able to perform both efficient exploration and exploitation of the search space since they search the space from a population of points, not just from a single point.

- They require only the evaluation of the objective function. No matrix inversion or partial derivative calculations are required. They also work with a coding of the parameter set, not the parameters themselves.

- They can work with both discrete and continuous decision variables. Discrete decision variables are more commonly used when solving engineering design problems. Nominal pipe diameters, for instance, always have discrete values.

- They generate alternative solutions that can be used later and compared if additional criteria are added.

- They use probabilistic transition rules to guide their search in the solution space. This reduces the probability of being trapped in a local optimum.

- Therefore, genetic algorithms can be applied to large and complex problems that are nonlinear and have multiple local optima.

3.6.2 Genetic Algorithm Limitations

GA optimisation is a powerful approach with a proven ability to identify near-optimal solutions. However, the correct operation of a GA optimisation depends on careful setting up and parameter tuning, requiring appropriate skills and experience. Inappropriate parameter values and penalty levels will distort the results away from the optimum.

Too high a rate of genetic exchange causes the process to degenerate into a random search while too low a rate of genetic exchange causes convergence on a local optimum. Because of the complexity involved in setting up a GA to operate effectively, this form of optimisation is not well suited to 'trivial' problems, i.e. those for which the number of possible solutions is small.

Other disadvantages compared with conventional optimisation methods are as follows:

- CPU time is normally much larger compared to other optimisation techniques. This will be less significant in the future since faster computers are continually being developed.

- GAs do not guarantee that the global optimum will be found, although they generally provide good solutions. To overcome this, several runs with different random seeds are usually performed. In fact, since the genetic algorithm approach is a stochastic search technique, the solutions obtained from different random seed values may be different.

- Since GAs are unconstrained optimisation techniques, penalty functions are used. In fact, the use of such functions may affect the efficiency of the exploration and exploitation of the search space.

4. AN OVERVIEW OF GA USE IN THE WATER INDUSTRY

4.1 Available Software Tools

A number of free simulation and optimization software tools are available for use in the optimization model. In this section, a number of these tools are discussed to show how they can be incorporated into the proposed model.

4.1.1 Open Source GA Libraries

Genetic Algorithms is an established optimization technique with applications in a large number of fields. For this reason, various public-domain software libraries are available on the Internet.

One of the best known GA code libraries is *GAlib*, developed by Mathew Wall at MIT (2003). GAlib contains a set of C++ genetic algorithm objects. The library includes tools for using genetic algorithms to do optimization in any C++ program using any representation and genetic operators. The documentation includes an extensive overview of how to implement a genetic algorithm as well as examples illustrating customisations to the GAlib classes.

GAlib contains four types of genetic algorithms. The first is the standard 'simple genetic algorithm' described by Goldberg (1989). This algorithm uses non-overlapping populations and optional elitism. Each generation the algorithm creates an entirely new population of individuals. The second is a 'steady-state genetic algorithm' that uses overlapping populations. In this variation, you can specify how much of the population should be replaced in each generation. The third variation is the 'incremental genetic algorithm', in which each generation consists of only one or two children. The incremental genetic algorithms allow custom replacement methods to define how the new generation should be integrated into the population. So, for example, a newly generated child could replace its parent, replace a random individual in the population, or replace an individual that is most like it. The fourth type is the 'deme' genetic algorithm. This algorithm evolves multiple populations in parallel using a steady-state algorithm. Each generation the algorithm migrates some of the individuals from each population to one of the other populations.

A good multi-objective GA library is *MOMHLib++* (Multiple Objective MetaHeuristics Library in C++), developed by Jaszkiewicz and Ziembinski (2004). MOMHLIB++ is a library of C++ classes that implements a number of multiple objective metaheuristics. It has been developed on the basis of former implementations of Pareto simulated annealing (PSA) and multiple objective genetic local search (MOGLS).

The MOMHLIB++ code library includes the following methods:

- Pareto simulated annealing (PSA)
- Serafini's multiple objective simulated annealing (SMOSA)

- Ulungu's et al. multiple objective simulated annealing (MOSA)
- Pareto memetic algorithm
- multiple objective genetic local search (MOGLS)
- Ishibuchi's and Murata's multiple objective genetic local search (IMMOGLS)
- multiple objective multiple start local search (MOMSLS)
- non-dominated sorting genetic algorithm (NSGA) and controlled NSGA II
- Strength Pareto Evolutionary Algorithm

4.1.2 Epanet

The Epanet hydraulic modelling package was developed by the United States Environmental Protection Agency to help water utilities to better understand the movement and transformations undergone by water in water distribution systems (Rossman, 2000).

Epanet can perform both snapshot (steady-state) and extended-period hydraulic analyses of incompressible flow in pipe networks. It can handle various hydraulic components including reservoirs, tanks, pipes, pumps and control valves. Epanet can also model water quality by simulating the behaviour of chemical constituents in the water distribution system with time. The water quality capabilities of Epanet were extended to allow water age and source tracing analyses to be performed.

The Epanet hydraulic and water quality engine has become the standard for the modelling of water distribution systems and is used by several commercial software packages. Epanet also has its own user interface and is available as a stand-alone software package.

A very useful addition to Epanet was the release of the Epanet Programmers Toolkit. The Epanet Programmers Toolkit is a dynamic link library of functions that allows researchers and developers to customize Epanet's computational engine for their own specific needs. The Epanet Programmers Toolkit can be used to query and set element properties, run hydraulic and water quality simulations and obtain simulation results.

The Epanet Programmers Toolkit is used extensively in research on water distribution modelling as is evident from the numerous papers referring to it. The Toolkit is written in ANSI compatible C using an efficient procedural programming style.

4.1.3 Ooten

Procedural programming (as used in Epanet) is a style of computer programming that focuses on actions, which are programmed into functions and then into units. A disadvantage of procedural programming is that the code can be difficult to write and maintain, especially when programs become large.

A programming style that addresses this problem and that has gained much popularity is object-oriented programming. Object-oriented programming has a number of advantages over the more traditional programming styles, such as ease of developing, checking, expanding, sharing and maintaining programming code. Water distribution systems are particularly suitable for object-oriented description since they consist of different components (e.g. pipes, junctions and valves) that can be modelled efficiently using objects.

A need was identified for a toolkit that can make the Epanet source code functions available in an object-oriented way. As a result, the RAU Water Research Group developed an Object-Oriented Toolkit for Epanet (Ooten) (Van Zyl, Borthwick and Hardy, 2003a) as an object-oriented alternative to the Epanet Programmers Toolkit as part of another WRC Project K5/1389: "Integration of Software Packages for the Probabilistic Analysis of Complex Water Supply Systems". The aim of this project was to continue the WRC-supported drive to establish probabilistic analysis of water distribution systems as a powerful adjunct to the usual, well-established deterministic analysis used by water engineers to design and improve water supply systems. The main product of this project was the development of a stochastic analysis tool, called Mocasim II, by integrating the stochastic analysis methodology with the public-domain software Epanet.

Object-Oriented Toolkit for Epanet (Ooten) provides many of the Epanet source code functions as methods of relevant objects. Ooten acts as a shell of the Epanet source code in the sense that very little data are duplicated. The project data remain in the Epanet source code, and Ooten calls the relevant Epanet source code functions to make the data and functions available in an object-oriented way.

All the Epanet Programmers Toolkit functions are incorporated into Ooten, but the functionality of Ooten has also been extended to include methods that are not available in the Epanet Programmers Toolkit. The main additions are the capabilities of Ooten to handle pressure-dependent demands and curves.

A comprehensive help program is provided with Ooten, giving extensive help on the capabilities and use of Ooten. Ooten is available as a free download from http://general.rau.ac.za/civil/wrg/ooten.htm.

4.1.4 Mocasim II

Mocasim II is a software package combining stochastic analysis with hydraulic analysis. It allows various types of analysis to be performed, including capacity versus reliability analysis for service reservoirs. It is based on a methodology that uses Monte Carlo analysis.

Mocasim II was developed under WRC Project K5/1389: "Integration of Software Packages for the Probabilistic Analysis of Complex Water Supply Systems". It is available as a free download from http://general.rau.ac.za/civil/wrg/mocasim.htm.

The stochastic analysis technique was developed by Nel (1993) and improved by Van Zyl and Haarhoff (2002). The initial method used a simple mass balance model for calculating the flows in simple linear distribution networks and was used to quantify the relationship between the capacity and reliability of service reservoirs.

The method was developed further, by incorporating the Epanet hydraulic engine in the stochastic analysis method. The Epanet hydraulic engine had to be modified to handle pressure-dependent demands and isolated network sections, for example when simulating pipe failures. The new method allows stochastic analysis to be applied to large and complex water distribution systems.

Typical stochastic analysis runs have long simulation run times and would not be suitable for inclusion in an optimization routine. However, various techniques may be applied to speed up the stochastic analysis process, or combine it with a GA optimization run to overcome this problem.

4.2 Water Distribution Systems

Genetic algorithms have been incorporated into a number of commercial software packages, mainly for network calibration and design. Most of software packages implementing GAs are not commonly used in South Africa and are often very expensive to buy and maintain. These software packages are thus not necessarily a solution for South African conditions, but do strengthen the point that GAs are the optimisation technique of choice for water distribution systems.

The software packages that have already incorporated GAs are listed below and are compared in **Table 4.1**.

H2ONet. This model is developed by MW Soft in the USA and runs in an Autocad for Windows environment. Its link to Autocad enables it to have good graphical capabilities. It also has the ability to do automated fire flow modelling and energy saving analyses. The H2ONET Calibrator uses GAs to determine the optimal pipe roughness coefficients to match field observations and best reflect what is actually occurring in your system. Their H2ONET Designer uses GAs to determine cost-effective rehabilitation, replacement, strengthening, and expansion options to reliably supply projected demands at adequate levels of service.

KYPipe. Developed at the University of Kentucky, this product claims to be the most widely used hydraulic simulation model in the world. KYPipe2000 does implicit optimisation based on a genetic algorithm approach to optimally adjust pipe roughnesses, valve settings, tank levels, demand distribution, and other data to provide a calibrated model. The program minimizes the difference between observed field data and model predictions considering all test data simultaneously to provide the best calibration possible.

OptiDesigner. Optiwater has a product called OptiDesigner, which is a Windows based software product for the optimal design of water distribution systems using genetic algorithms. The program uses the Epanet hydraulic simulator.

Watercad. Haestad Method's water distribution modelling package developed for the Windows environment. It links automatically with Autocad, Microstation and Intergraph to allow for transfer of data, drawings, etc. Watercad's Darwin Calibrator is an integrated calibration environment coupled with a powerful genetic algorithm optimisation engine with the hydraulic network simulator. It allows hydraulic modellers to quickly and easily calibrates steady state, multi-steady state and extended period hydraulic models. Their Darwin Designer uses GAs to evaluate a range of options for creatively implementing sound and economical design, rehabilitation, and enhancement solutions.

Table 4.1: Comparison of genetic algorithms in existing network models

Software network models	Genetic Algorithms	Website
H20Net	Calibration Design Pump scheduling	www.mwsoft.mw.com
Optiwater	Design Calibration available soon	www.optiwater.com
KYPipe	Calibration	www.kypipe.com
WaterCad	Calibration Design	www.haestad.com

5. CONCEPTUAL DEVELOPMENT OF GA'S FOR DETERMINING THE OPTIMAL OPERATING RULE FOR WATER RESOURCE SYSTEMS

5.1 Background

In the literature survey that was carried out as part of a previous study (Van Vuuren, *et al*, 2001), it was clear that Genetic Algorithms could be applied for the optimisation of the water resource system operating rules. All the examples found in the literature, however, is based on water resource modelling techniques and analysis procedures that are different from those used in South Africa. The challenge therefore lies in developing a genetic algorithm that is applicable to the currently applied analysis technology and uses the existing water resources models as point of departure.

The operating costs (in particular pumping costs) of the various water resource systems in South Africa amounts to hundreds of millions of Rand per year and much time is spend by water resource planners to implement cost saving operating rules that do not jeopardise the assurance of supply over the long term.

The method of optimisation that is currently carried out is through a process of scenario simulation, where different rules are analysed and evaluated using network simulation models. This process requires a high level of human intervention in defining the scenarios and processing the results in an iterative cycle. Typically the selection of the scenario of rules is bases on the experience of the annalist and the knowledge of the particular system being optimised. Although this process has provided acceptable results in the past, the question that always lingers is whether or not a more optimum-operating rule exists that has not been considered by the annalist. The descriptions given in the report assumes that the reader has a basic knowledge of the Water Resources Yield Model (WRYM) used by the Department of Water Affairs and Forestry. The reader can refer to the "*Water Resources Yield Model User's Guide*" DWAF (1999), describing the data structure and functionality of the model and the "*Water Resources Yield Model Procedural Manual*" DWAF (2005), which describes the processes that is applied when building and using the WRYM. The Procedural Manual describes the model elements (reservoirs, nodes and channels) and the related analysis procedures and the User Guide is structured around and focuses on the input files and the data formats.

5.2 Purpose and Layout of Chapter

The purpose of this chapter is to describe, conceptually, a Genetic Algorithm for the optimisation of water resource system using the existing network simulation models and analysis techniques that are applied in South Africa.

Section 5.3 provides guidelines that were identified to direct the design of the conceptual model and gives the reader an understanding of framework within which the conceptual model was developed.

Section 5.4 describes how the operating rules of water resource systems are defined in the Water Resources Yield Model (WRYM) and **Section 5.5** presents the proposed structure of the GA and the connections with the WRYM. In **Section 5.6** the proposed analysis procedure is presented showing how the GA will interact with WRYM.

5.3 Development Guidelines and Framework

This section describes the guidelines that were applied in the design of the conceptual GA model and serve as a means of informing the reader of the thought processes and background that led the authors to arrive at the particular model description.

5.3.1 Pointers from the current modeling environment to be considered in the development of Gas

Introducing a new procedure (the GA) into the water resource analysis environment, where a large pool of knowledge is founded on present technologies, requires careful planning. The aim is to implement the GA so that a high degree of the current knowledge is maintained and future analysis requirements are recognised.

To this end, relevant guidelines are listed below:

- The GA should be developed for the existing water resource models that are used in practice for the management of most of the countries water resources. These models are the Water Resources Yield Model (WRYM) and the Water Resources Planning Model (WRPM). The advantage of using the existing models is that the existing pool of knowledge that reside in water resource analysts, which are currently applying these models, can be extended further by adding optimisation functionality.

- Cognisance should be taken of the analysis techniques that are used in practice in South Africa. Currently an important parameter that quantifies the supply capability of a water resource system is the Firm Yield that can be abstracted. Two types of Firm Yield values are usually determined:

  - the Historical Firm Yield, which is based on analyses of the historical sequence and
  - the Stochastic Firm Yield, which also defines the reliability at which the yield is available (DWAF, 1987).

- The aim with the optimisation of an operating rule for a system is, in most cases, to maximise the firm yield in order to "stretch" the resource as far as possible. Minimising the operating costs, in particular pumping costs, is normally a secondary objective and usually opposes the objective of maximising the firm yield.

- In order to build confidence in the application of GA technology it is deemed essential to introduce (develop and verify) the GA technology in phases. The intention is to prove the value of a GA by using basic yield analysis and relatively simple systems as a first step. The GA can then thereafter be extended to include probabilistic projection analyse of complex systems, involving changing system characteristics where demands are growing, new reservoirs and transfer conduits are implemented over a planning horizon.

- It is proposed that the GA first be developed to only optimise the operating rules of a system, and as a further phase, be extended to include the optimisation of the capacities (sizes) of water resources infrastructure such as conveyance capacity of water conduits and the storage capacity of reservoirs.

- Although the time it will take the GA to compute a solution is an important factor to consider, the focus should initially be on ensuring the viability of the technique and the validity of the results. Once the integrity of the GA has been established and proven, resource can be employed to reduce the computational requirements.

- Within the context of the above guideline, the applied water resource analysis techniques allow for different types of analysis which required different computational requirements. For example, historical analysis, where only one hydrological sequence is analysed is computationally significantly less intensive compared to stochastic analyses where up to a thousand sequences are analyses in some cases. The conceptual design of the GA should allow for both types of analysis techniques.

- The basic design of the GA should be to maintain the flexibility and generic structure that is provided by the network models. This will involve making the definition of the GA as generic as possible so that it can be applied to most of the water resource systems and defined through model input data files.

- The requirement of the National Water Act (Act 36 of 1998) to implement Licensing will change the emphasis of water resource analysis in South Africa from yield determining assessments of large systems to system analysis dealing with multiple users that are dispersed over the catchments of the different river systems. Water resource analysis for Licensing will broadly focus on two aspects, (a) accounting for the interaction users have on one another in the system and (b) determining how the reconciliation of water balances will be achieved in a fair and equitable manner. The GA design will have to take these requirements into account.

5.3.2 Pointers from the literature – focus on the GA

Information from past research, as presented in the literature, indicates that the formulation of a GA consists of four main elements,

- the encoding scheme,
- the GA operators,
- the process (water resource system) that is optimised, and
- the objective function as translated into the GA fitness function.

These processes are described briefly in the following sections.

5.3.2.1 Encoding scheme and initialisation process

Encoding involves the process where the "genes" of the GA is converted into the input parameters of the process to be optimised. In the case of water resources the parameters are the variables defining the operating rules and the process to optimise, in the water resource system. An important component of this encoding process is to ensure that the constraints or limitations on the parameter are adhered to when assigning their values.

From the literature it has been noted that some form of random process is used to generate the initial set of genes (Goldberg, 1989; Oliveira and Loucks, 1997 and Ndiritu, 2003). Thereafter the genes are converted into parameter values and the feasibility of the values verified. If a parameter set is infeasible, that gene is rejected and the process is repeated until a full set of initial members (also called the Initial Population) is produced.

From the literature it was pointed out that there are advantages to define an encoding scheme that groups decision variables together, as apposed to dealing with each parameter individually (Oliveira and Loucks, 1997). In doing this, one may also be able to capture some of the constraints and limitation presented above.

A further important requirement specified in the literature is that the selection of the initial population should be such that the parameter values must be uniformly distributed over the search space (Oliveira and Loucks, 1997).

5.3.2.2 GA operators

The GA operators have been discussed in Chapter 3 and the specific use of them within the context of water resources analyses are reflected below.

Fitness function

The fitness function consists of two components (a) the objective function representing the yield, reliability and pumping costs, and (b) the fitness scheme which involves the calculation of a numerical value for each gene that reflect how good a gene has performed relative to the others.

The objective function can include multi objectives and typically a weighting factor is used to represent the importance of different objectives.

The fitness values of each gene are used to select which genes become parents for the determination of the next generation of children genes.

Pairing
Pairing is the process where pairs of genes are selected based on their fitness (the genes with high fitness values have a higher probability of being selected) for the next generation. The applicable selection process to water resources yield modeling, is referred to as "Tournament Selection".

Crossover
Working with real-values chromosomes, as opposed to binary chromosomes, is a clear-cut way for the design of the GA. Given that real-value chromosomes will be use in the proposed development, two means of crossover were suggested by Oliveira and Loucks (1997), namely, uniform and quadratic. Uniform crossover basically involves the copying of parameter values from one of the parents to a child (next generation). The selection of which parent's parameter values are copied is based on a binary random variable. For the quadratic crossover a weighting system is applied where a new values is calculate for the child chromosomes based on both the parent's parameter values. The weights are such that the parent with the best fitness has the largest influence (contributes the most) on the child's parameter value.

As was the case with the initial population, the feasibility of the children's chromosomes or genes has to be tested. Adjustments are then made to the parameters of the infeasible cases where needed.

Mutation
This is the process where, usually a small number, of child chromosomes are changed "arbitrarily" in order to create some maverick chromosome to assess the fitness of a diverse set of cases in the search space. Mutated chromosomes also have to be tested for feasibility or the method of mutation should be adapted to ensure that the feasibility is preserved after mutation.

5.4 Variables Defining Water Resource System Operating Rules

5.4.1 Overview

There are four basic variables defining the operating rule of a water resource system in the Water Resources Yield Model (WRYM) as listed below:

- Number of operating zones in a reservoir.
- Bottom elevation level of reservoir zones.
- Penalty value of the water in reservoir zones.
- Channel penalties.

These are described in detail in the following sections.

5.4.2 Number of operating zones in a reservoir

The number of zones in any reservoir can be as many as twenty in the current WRYM. Although the WRYM data structure is such that the number of zones in all reservoirs is the same, the bottom zone level (Section 5.4.3) of individual reservoirs can be used to eliminate zones by making the storage in them zero. Since the number of zones does not directly influence the behavior of a water resource system it is proposed to make this variable a fixed size, of say ten zones. It might also be considered to optimize the penalty values themselves obtaining the best operational sequencing of the system.

5.4.3 Bottom elevation level of reservoir zones

The difference between the bottom elevations of two zones, that are "stacked" on top of each other, defines (indirectly through the elevation-volume characteristics) the volume of storage that is contained in a zone. It is important to note that, with the WRYM, different zone levels can be defined for each of the 12 months of the year. This flexibility allows for the design of operating rules that accommodate seasonal variability. However, in most of the systems that are modelled in South Africa, constant zone levels are used throughout the year.

5.4.5 Penalty value of the water in reservoir zones

WRYM currently provides for the definition of one penalty value per reservoir zone for all the months of the year. This implies that the sequence of drawdown among the zones in different reservoirs is the same throughout a particular scenario analysis. It would be possible to change the input data structure of WRYM to accommodate different zone penalties for the 12 months of the year. However, it is proposed to only consider such an enhancement once the GA has been proven viable with the current structure.

There are certain constraints that have to be adhered to when assigning values to the zone penalties. In most cases, where the "Rule Curve" (see WRYM Users Guide for definition) is set to be equal to the full supply level of the reservoir, the penalty values that are assigned to the zones must be in ascending order, following the stacked sequence of the zones from top to bottom.

The application of the reservoir zone penalties to implement an operating rule in a water resource system, is best explained by an example as shown in **Figures 5.1** and **5.2**. **Figure 5.1** and **Figure 5.2** illustrates the schematic layout of the Usutu River System with the main components being the three reservoirs Westoe, Jericho and Morgenstond, indicated by the triangles numbered 21, 22 and 23 respectively. Each of these reservoirs has an upstream "Dummy dam" which simulates the combined affect of small farm dams in the respective catchments. Node 24 represents Churchill Diversion Weir, which diverts water from an adjacent tributary into Westoe Dam through channel 37.

Channel 32 conveys water under gravity from Westoe Dam to Jericho Dam from where it is pumped to the users as indicated by channel 36, which is the main yield channel of the system. Water is pumped from Morgenstond Dam through channel 34 in support of Jericho Dam.

The symbol "$\mathcal{S}$" represents the runoff from the respective catchments that is, for example, labelled as "I114" for Churchill Dam. In cases where the runoff from a catchment is split, to enter the network system at more than one point, an associated fraction (see the values given in brackets "(0.15)" in **Figure 5.1**) is use to indicate the portion of the runoff that is being routed through the particular node or reservoir.

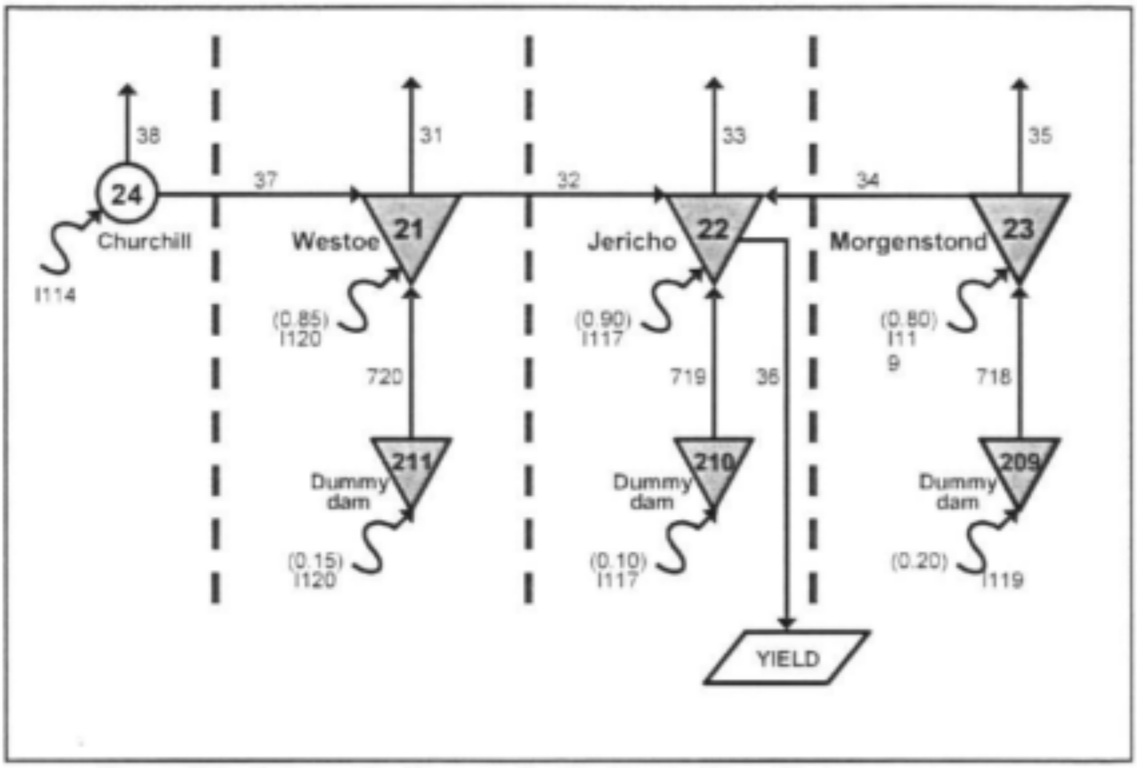
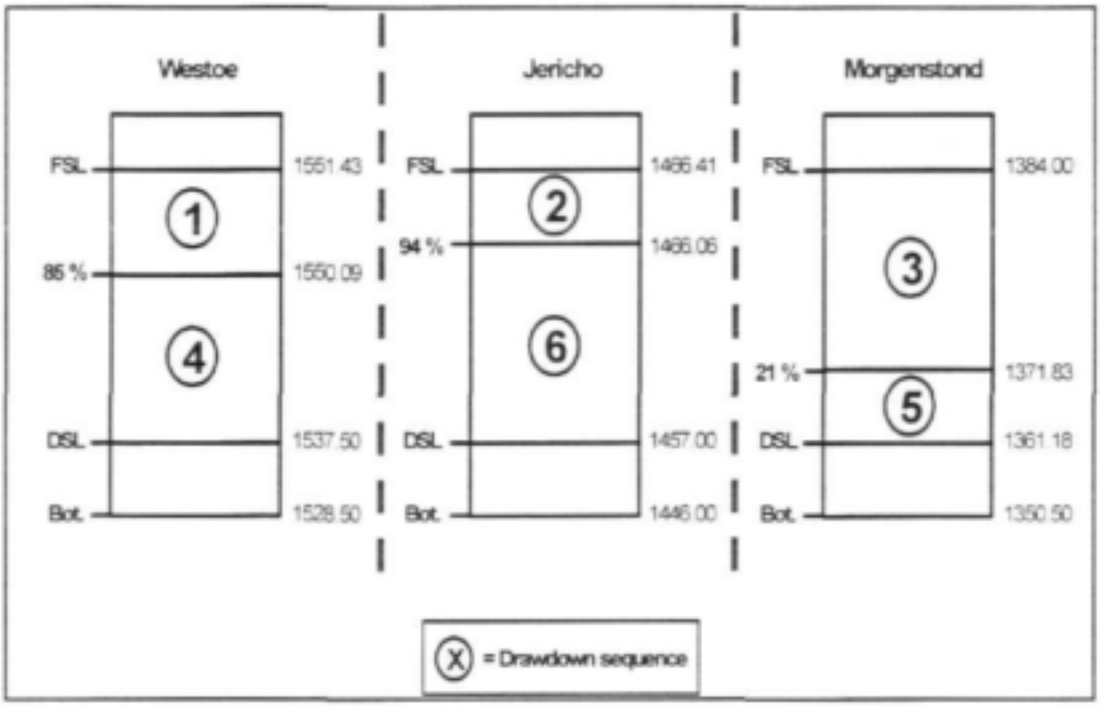


Figure 5.1: Schematic diagram for the Usutu River System

Figure 5.2 gives a typical reservoir zone definition of the three main reservoirs in the Usutu River System. The bottom levels of each zone are indicated in metres above sea level (at the right hand side each of the reservoir diagrams) and the circled numbers presented in each of the zones, represent the sequence in which the zones will be utilised during drawdown periods.

To achieve the indicated draw down sequence of the zones in the three dams any set of penalty values can be defined as long as the relative magnitude of the penalties complies with the required sequence. It should be noted that the penalties assigned to the channels linking the reservoirs also impacts on the operation of the system and the definition of the channel elements are presented in the following section.



Notes: FSL = Full Supply Level, DSL = Dead Storage Level, Bot. = Bottom of reservoir
The zone levels are given in metres above sea level.

Figure 5.2: Reservoir zones and drawdown sequence

It is important to note that the penalty values assigned to each reservoir zone have two functions. Firstly it represents the "unit constraint" of drawing water from the reservoir zone and secondly the, same penalty value serve as a "unit gain" of storing water in the particular zone. This second functionality is one mechanism of how the network model facilitates transferring water from one reservoir to another. The model also has a second mechanism for transfer between reservoirs as defined by channel **Type (b)**, described in **Section 5.4.5** below.

The intention is that the optimisation process will involve searching for the zone levels and zone penalties, as two of the three types of parameter, to maximise the fitness function. The third parameter will be the penalty values of channels, which are presented in the following section

5.4.5 Channel penalties

A channel is the modelling element that represents a conduit (links) between reservoirs or nodes. The penalty structure of channels can be divided into two basic types and the distinction between the types lies in the number and direction of the arcs that makes up a channel.

For this discussion, **Type (a)** is a channel that is only defined by one arc and the (flow) direction of the arc is the same as that of the channel. In this case the penalty value represents a "unit constraint" that has to be overcome before water will flow in the channel.

For flow to occur in the channel water must be drawn or pull through at a "unit gain" that is greater than the "unit constraint" of the channel. **Figure 5.3** below shows graphically the configuration of a **Type (a)** channel and indicates the notation that is used to describe a channel.

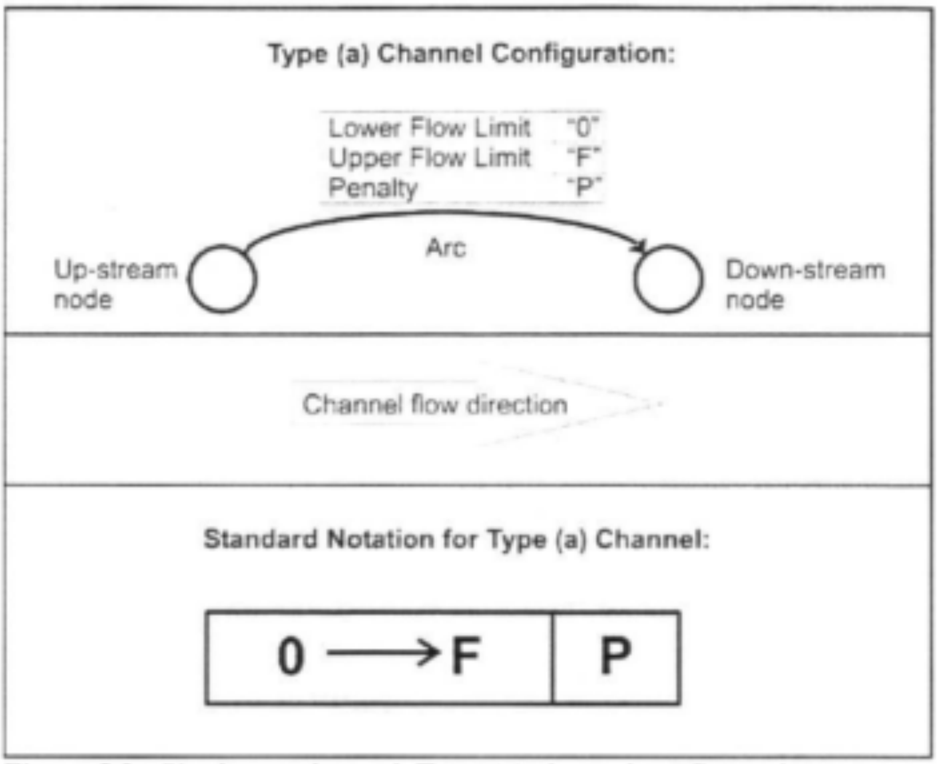


Figure 5.3: Single arc channel, Type (a) channel configuration

Type (b), see **Figure 5.4**, is a channel that consists of two arcs, one arc in the same direction as the channel (also referred to the forward arc) and the second arc, in the opposite direction of the channel or backwards arc.

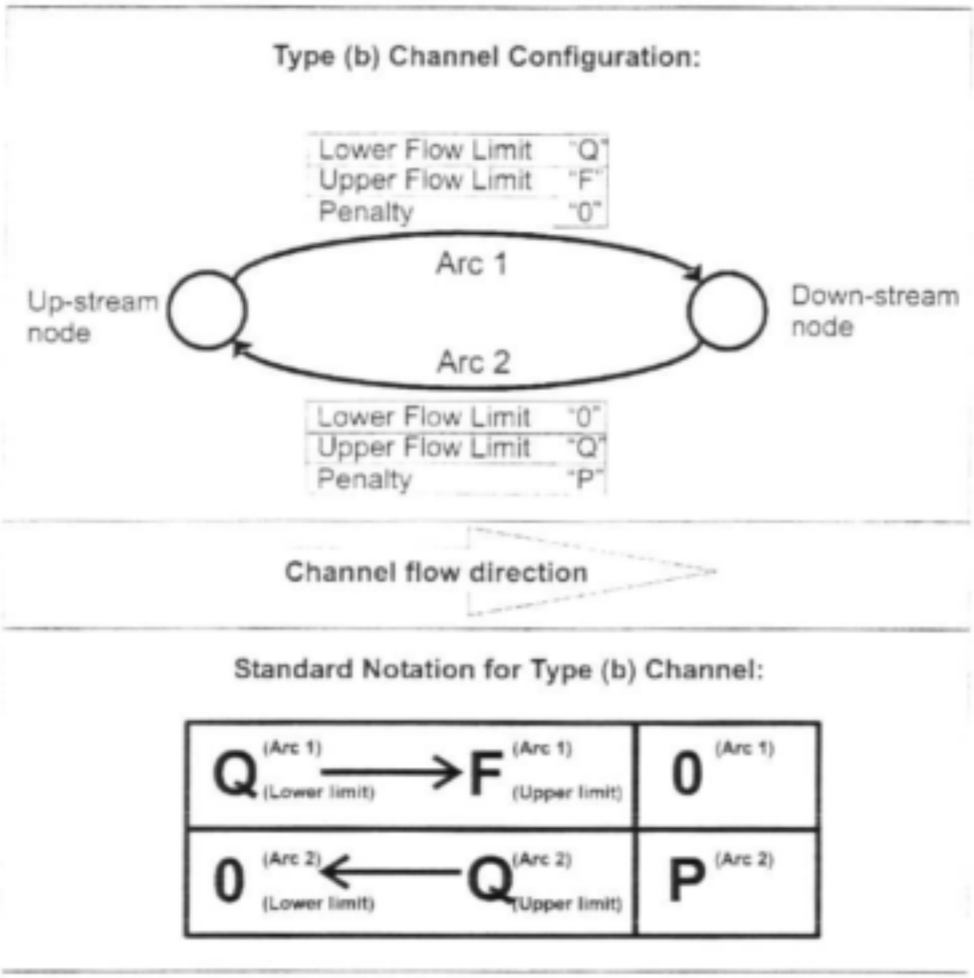


Figure 5.4: Two arc channel, Type (b) channel configuration

The operation of a **Type (b)** channel is as follows: The minimum flow defined for Arc 1, the forward arc, will always be "Q" and can be as high as "F", which is the upper limit of the arc. The purpose of the second arc, with the flow direction opposite to the channel's flow direction, is to allow the "relaxation" of flow in the channel up to a volume of "Q". Relaxation flow through the second arc occurs at a penalty "P". Therefore, if the penalty "P" is higher that the penalty assigned to the water in a reservoir zone, for example, "Q" flow will be drawn (pulled) from the reservoir and no relaxation is needed (no flow is routed through the relaxation arc or Arc 2).

In summary the function of a **Type (a)** channel in a water resource system is to serve as a conduit through which water can be drawn by means of either a reservoir zone penalty value (unit gain) or a **Type (b)** channel, in other words a **Type (a)** channel cannot "pull" water through itself. However, a **Type (b)** channel is capable to drawing water through itself with a "pulling force" that is equal to the penalty value "P" that is assigned to Arc 2 as indicated in **Figure 5.3**.

During the operating rule optimisation process the channel penalty values are changed in relation to the penalties assigned to the reservoir zones to implement a particular operating rule. In general the type of channel (**Type (a)** or **Type (b)**) would remain the same for all the scenarios with the only changes being to the penalty values of the channels. In the example of the Usutu River System, the penalty values of channels 32 and 34 would be the only possible candidates that would form part of the optimisation process.

5.5 GA definition

The processes that make up the GA and their interactions with one another are best illustrated by the flow diagram shown in **Figure 5.5**. The indicated processes and the linkages between them were compiled using similar algorithmic flow diagrams and GA descriptions found in the literature.

5.5.1 Description of GA processes

The steps presented below are shown on the schematic diagram, **Figure 5.5**.

5.5.2 Step 1: Initial gene pool generation

The first process of the GA involves generating initial sets of genes that defines the operating rule values for the intended water resource system elements using a random selection method that results in parameter values that are uniformly distributed over the search space. In order to be able to duplicate results across different computer systems it is required to implement a "software" random number generator (similar to the subroutines that are currently applied in the WRYM) as apposed to a computer dependant random number generator. These random number generators make use of seed values that will be changed during the evaluation and testing phase to illustrate that convergence is achievable with different initial gene populations. Once it is shown that convergence is achieved, the seed values will be fixed in the production version of the model.

In order to only allow penalty values that are not in conflict with the penalties of the water resource elements that are excluded from the optimisation process, it is necessary to limit the possible selection of penalty values for optimisation within a pre-defined range. This range of penalty values will be fixed for a particular system configuration, however, the range has to be defined in relation to the penalty values assigned to the water resource elements that are not optimised. Both the penalty value search range and the penalties for the water resource elements that are not optimised have to be defined by the model user.

5.5.3 Step 2: Test the feasibility of the gene sets

This procedure evaluates each gene set to identify if they represent feasible penalty definitions by complying with the following constraints:

- Penalty values should be in ascending order from the top zone to the bottom zone for each reservoir.
- *Penalty values for* **Type (b)** *channels have to be such that the penalty* for Arc 2 must always be greater than the penalty for Arc 1.

5.5.4 Step 3: Calculate the fitness of the gene sets

This step involves three basic processes; first each gene set is converted into penalty values that define an operating rule scenario. The water resource system is then analysed iteratively, with the operating rule scenario in place, to determining the firm yield of the system. (The iterations referred to here are different runs with the same inflow sequence, historical or stochastic, where the target abstraction of the yield channel is changed between the runs to find the Firm Yield value.) Other simulations results such as the flow through pumping channels will be extracted from the simulation output (this will be for the run where the firm yield is imposed as the target draft) and summarised in a form required by the fitness function. Additional simulation result, such as the volume of spillage from the system, may also be used as a component of the fitness function. The model will be developed in such a way that the user can select which channel flows should be included in the fitness function and allowances will be made to assign weights to each of the selected channel flows and the firm yield result.

5.5.5 Step 4: Tournament selection

This procedure uses the result of the fitness function of a particular gene set, relative to the sum of the fitness function results of all the gene sets, to select which of the parent gene sets are to form pairs for crossover. The selection procedure is carried out by a roulette-wheel parent-selection scheme (Goldberg, 1989), which assures a higher probability of selecting the parents that has the best fitness function. The operation of the roulette-wheel procedure requires the generation of random numbers, and for the same reasons as presented in *Step 1*, a software random number generator will be used for this purpose.

Oliveira and Loucks (1997) performed tests on two selection schemes, window and ranking (Davis, 1991), and found that the ranking scheme was preferred. It is therefore proposed to also apply the ranking scheme in the GA model.

In most of the GA applications presented in the literature an additional feature is applied in the selection of the child genes call **elitism**. This involves using the genes of the best solution, unchanged, as one of the child genes. This ensures that the best solution of one generation is kept unchanged to compete with the next new generation of children genes.

5.5.6 Step 5: Crossover operation

Crossover is the GA operator that creates a new generation of gene sets, two genes from each parent pair. In the classic strain representation of a gene (strain is a string of zeros and ones), crossover involves exchanging (swapping) parts of the two parent genes to form the new children genes. Given that real value chromosomes will be used for the GA the procedure would be somewhat different. In the context of the selection of the reservoir zone and channel penalty values, it is proposed to use either uniform or quadratic methods of crossover to produce the new generation as suggested by Oliveira, et al. (1997).

In order to maintain the constraints the suggestion of Olivera, et.al (1994) and others are that parameters should be grouped rather than working with each parameter individually. In the context of the WRYM this would mean that the penalty values for the zones in a reservoir and for the arcs in a channel would respectively be treated as groups. Another group would be the zone levels of each reservoir. The uniform crossover method would then select one of the groups of parameters of a parent as the parameters for the child. The selection process to pick which parent's parameter group is transferred to the child will be based on the results of a random generator that produces either a one or a zero value. The selection is then as follows, if the random generator's result is one, the parameter group from the first of the two parents is used and vice versa if the random generator's value is zero. The generation and selection process is repeated for all the parameter groups that are part of the optimisation process.

5.5.7 Step 6: Mutation

Mutation is the process where maverick genes are generated to ensure the search space is explored randomly to avoid the optimisation process only achieving local optimum solutions. From the literature it is show that the percentage genes of each generation on which mutation should be applied ranges from 0.05 to 0.20.

In order to maintain the feasibility of the operating rules after mutation it will be required to only allow mutated parameter values, penalties and zone levels, to change within defined limits. A zone level and a zone penalty value could only be changed to a value that lies between the values of the zones above and below. The value will be calculated by using a uniform distribution selection method.

5.5.8 Step 7: Termination

The final step is to check whether an optimum answer has been found. This will be based on a criteria that will indicate by how the rate of improvement from the one iteration to the next has changed and if the rate is large the process will loop back to Step 3 or stop.

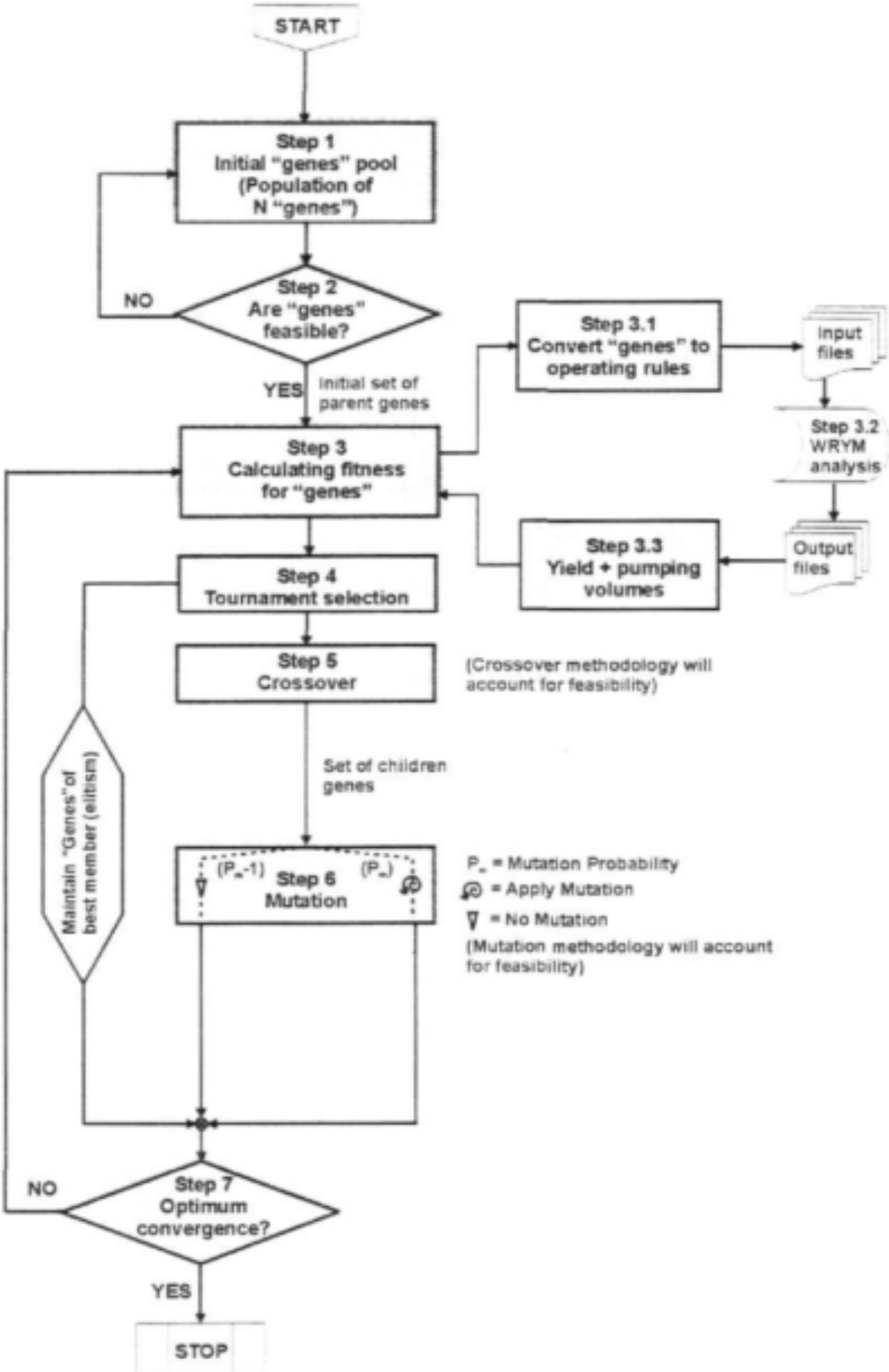


Figure 5.5: Flow diagram of Genetic Algorithm

5.6 Proposed Analysis Procedure

Within the context of the existing structure of the WRYM, **Figure 5.6** presented below gives a schematic presentation of how the GA can be implemented.

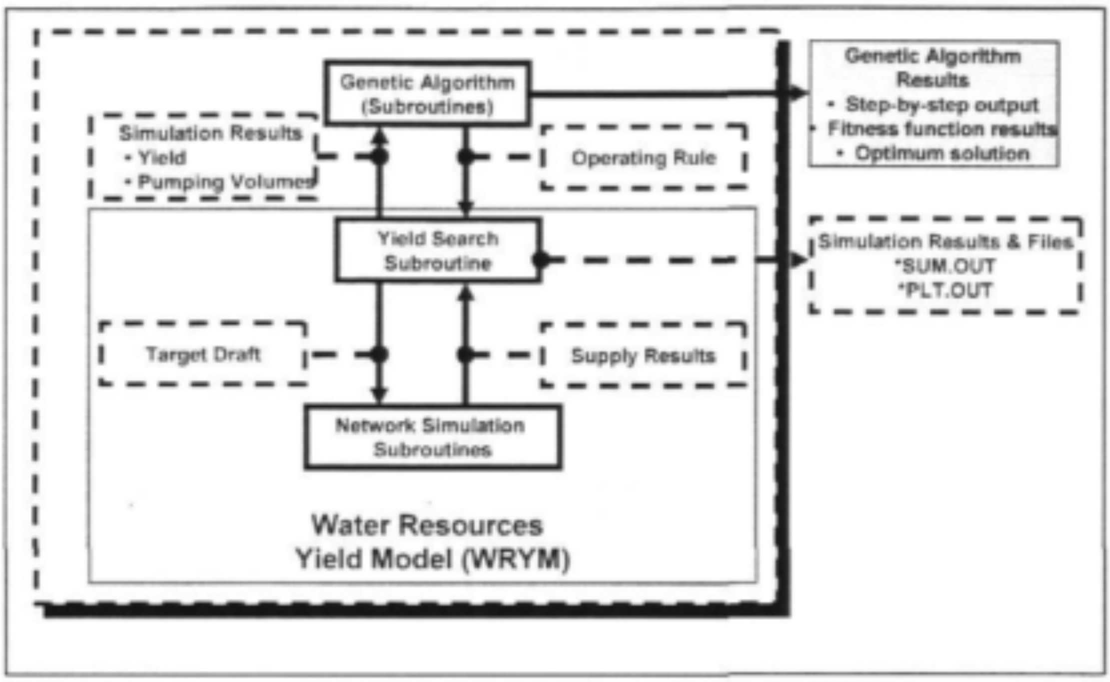


Figure 5.6: Schematic representation of the GA linked to the WRYM

As indicated, the Genetic Algorithm Subroutine will interact with the Yield Search Subroutine by imposing the operating rule parameters onto the system and it will receive back the yield results and relevant pumping volumes. The purpose of the Yield Search Subroutine is to control the Network Simulation Subroutine through an iterative search process to find the yield, historical firm yield or stochastic yield at a given reliability. The Yield Search routine will output relevant simulation results to the existing output files (*SUM.OUT and *PLT.OUT). The step-by-step results from the GA will present output to a new file, summarising the process and behaviour of the GA for evaluation purposes.

What is apparent from the discussions presented above is that the elements (channels and reservoir zones) that should form part of the optimisation process are dependant on the characteristics of the particular water resource system and have to be defined as an input to the GA algorithm. For example, the dummy dams presented in **Figure 5.1** would not form part of the optimisation process and should therefore be excluded from the list of elements to be optimized. The data structure of the GA should therefore allow the system annalists to select the water resource elements for optimisation based on their understanding of the constraints and flexibilities that exist in the water resource system. This will be achieved through a separate input file that will contain the parameter for the GA Subroutine.

6. CONCEPTUAL DEVELOPMENT OF GENETIC ALGORITHMS FOR APPLICATION OF PIPE NETWORK OPTIMISATION

6.1 Background

One of the research products of this project is the conceptual development of procedures to use GAs for network analyses and scheduling. A major problem for water supply authorities is to select those components in a network that should be changed, increased or replaced to ensure a sustainable service to the consumers.

Although this problem does not sound difficult, it is in fact a very complex optimisation problem. The pipe network, pump system (scheduling, capacity and configuration), reservoir system (size, water levels) are all variables that have to be taken into account. The complexity is mainly caused by two factors; the discrete nature of the variables and the size of the solution space. Most optimisation techniques can only be applied to problems that have continuous variables, unlike the pipe diameter variables in the network optimisation problem, which are only available in discrete sizes. The size of the solution space (the total number of possible solutions to the problem) for the network optimisation problem can be calculated as the number of possible pipe diameters to the power of the number of pipes in the network. For example, if a network has 100 pipes and each pipe can have 7 possible diameters, the size of the solution space is 7^{100}. Although a 100 pipe network cannot be considered large, the number of possible solutions for this network is so great that it is totally impractical, if not impossible, to evaluate them all.

The discrete nature of the network optimisation problem and the size of the solution space make it virtually impossible to apply any of the conventional optimisation techniques to find the global optimum. GAs filled this gap and it is now accepted by most experts that GAs is the best technique for network optimisation.

Water supply authorities are also faced with problems besides that of network design. Other problems include network calibration, operation and reliability. During the implementation of this research product, the research aims were extended beyond network analyses and scheduling to also include other applications of GAs in water distribution systems.

6.2 Conceptual Problem Formulation

Seen holistically, GAs can be used to optimise various different parameters in water distribution systems. In this section, a conceptual model is developed for a system that will be able to optimise any of these parameters and combinations of these parameters using different types of GAs. It should be emphasised that the conceptual model is very general and that to attempt to implement it in a single phase will be risky and impractical. The conceptual model should rather be seen as a framework for further work on GAs. Achievable parts of the framework should be identified and implemented one at a time. The framework will, however, ensure that these individual projects could be integrated to ensure that maximum benefit is achieved.

Various types of optimisation can be identified for water distribution systems. These types can be grouped into design, operation, calibration, level-of-service, monitoring systems and testing. For each type of optimisation, the main objective function, possible variables and the main constraints are defined and discussed in the remaining part of this section. The types of optimisation are summarised in **Table 6.1**.

Although the types of optimisation are discussed separately, it is possible, and often desirable, to combine a number of these objectives when solving practical problems. It is, for instance, prudent to also consider the operational (energy) cost when designing a new pumping line. The formulation of the optimisation problem for any real-life problem will depend on both the water distribution system and the objectives of the study. The number of ways that the types of optimisation can be combined is very large. For this reason, the types of optimisation are discussed separately as a first step. Allowance is made in the conceptual model formulation section to combine the required types of optimisation in a coherent way to solve a particular problem.

In many cases the types of optimisations are in conflict with each other. For example, attempts to minimize operational cost will generally place the system in a more vulnerable state and less able to handle abnormalities such as pipe bursts, thus reducing the level of service (Jowitt, Garrett Cook and Germanopoulus, 1988). In such cases it is necessary to strike a balance between the objectives. This can be done by either defining the balance before hand (for example using an objective function equally weighed for cost and level of service) or by using multi-objective optimisation.

Table 6.1: Types of optimisation for water distribution systems

Optimisation type	Objective	Possible variables	Main Constraints
Design	Minimise cost.	Pipe layout Pipe diameters Pipe rehabilitation	Min level of service Available diameters Rehabilitation options Available budget, LCC
Operation	Minimise operational cost.	Pump controls Reservoir levels Sources and capacity	Min level of service Number of pump switches Source capacity Pump capacity
Calibration	Minimise difference between model and observed values	Pipe roughness Pipe diameter Valve settings Leakage Demands	System layout Available data
Level-of-service	Maximise level of service, e.g. pressure, water quality or reliability	All above	System configuration Budget
Monitoring system design	Minimise cost of monitoring system	Number of monitoring points Positions of monitoring points	System configuration Budget
Network Testing	Find critical sets of events that may cause a system to fail.	Fires Pipe failures Power failures Contamination events	System configuration Number of simultaneous events

6.2.1 Design

In design optimisation, the objective is to find the most cost-effective solution to a particular design or rehabilitation problem. This is the most common application of optimisation in water distribution systems. Examples include:

- Determine the optimal pipe diameters for a new network (utilizing Life Cycle Costing).
- Determine the optimal pipe layout for a new network.
- Determine the optimal combination of pipe rehabilitation (cleaning, slip lining, etc.), pipe replacement and new pipes to upgrade an existing system for future demands.

The main constraints are the following:

- The minimum level of service required, normally defined in terms of the minimum pressure under peak demand conditions. Other level of service parameters can also be used, for example the maximum water age or minimum flow velocity in the network.
- The available pipe diameters are restricted by the commercially available diameters. A minimum pipe diameter is normally specified.
- The rehabilitation options are restricted by various factors, such as the in situ pipe material, ground cover (for instance paving) and cost.
- The available budget plays a determining role in some cases and may limit the available options and the final level of service that can be supplied.

6.2.2 Operation

The goal of operational optimisation is to minimize the operational cost of water distribution systems. Pump operating cost make up a large proportion of the expenses of water supply companies, which is why operational optimisation is often described as pump scheduling optimisation. A conceptual method for operational optimisation using GAs was developed and is discussed in **Section 6.3**.

6.2.3 Calibration

Calibration of water distribution system models is the process of ensuring that the model results are realistic, i.e. the model results reflect the true situation in the field. Pressure and flow data are measured in the field and compared to the network model output under similar conditions. To calibrate the model, uncertain parameters such as pipe roughness and water demands are adjusted until a good correlation is achieved between the model results and the measured data.

Network calibration can be defined as an optimisation problem with the objective of minimizing the difference between measured values and model results. The optimisation variables can include pipe roughness values, pipe diameters, demands (including leaks) and even valve settings.

The main constraints include the network layout and the available calibration data.

6.2.4 Level of Service

Handling level of service parameters is often difficult since they are a complex amalgam of factors including consumer demand, pressures, pumps and leakages (Jowitt et al, 1988). In most cases level of service objectives are in direct conflict with other objectives such as minimizing operational or capital costs. Some of the more common level of service objectives are:

- Water quality objectives (Forzinin, 1991 and Percia, Oron and Mehrez, 1997). Water quality is often a function of the time water spends in the system before being consumed. Large service reservoirs may often increase retention time and thus negatively influence water quality. Different sources may also provide different water qualities that have to be matched to the quality requirements of the system users.

- System reliability and security of supply (Forzini, 1991). Reliability and security of supply is often achieved by increasing the levels of storage in the system, which is normally in conflict with water quality requirements.

- System robustness objectives. System robustness is measured by the ability of the system to handle stochastic variations such as fluctuations in system demand (Ko, Oh and Fontane, 1997) and emergencies such as pipe failures or fires. The distribution system should be able to provide for these events and recover to normal operation swiftly following such emergencies. Jarrige, Harding, Knight and Howes (1991) stated that although optimisation is normally used for a range of normal operation conditions, it is particularly beneficial under critical situations, when even trained staff cannot rely on their personal experience to get the appropriate answer to enable them to manage and control the system. It was found that operators will often over-prioritise security and consequently increase the costs under these conditions.

6.2.5 Monitoring System Design

Various online monitoring systems have been implemented to measure level of service in water distribution systems. Parameters include pressure, flow rate, chlorine residual values and others.

The objective of monitoring may be to ensure that minimum pressures are maintained, manage unaccounted-for-water, continuous automated system calibration and warning against possible intrusion of contaminants into the system.

The objective of the optimisation is to determine the optimal position of the monitoring stations to ensure the most effective information gathering. Constraints, such as the system configuration and budget (or number of monitoring stations that can be installed) form an integral part of the optimisation process.

6.2.6 Network testing

An area that has not yet found adequate coverage in water distribution systems is the testing of the system to identify the most critical events or combinations of events that may result in system failure. For instance, the pipe between the source and supply network is much more critical for service provision than a pipe in the network where water can be supplied along many paths. Events that may be investigated include fires, pipe failures, power failures and contamination events (including possible terror attacks).

The effect of single events are normally easy to determine. However, it has been shown that combinations of events, which are not critical as single events, can be highly critical in combination (Anderson, Bonabeau and Scott, 2004). Optimisation can assist in identifying the most critical combinations of events and subsequently designing more robust water supply systems.

6.3 Conceptual Optimisation Model Formulation

It is necessary to have a high-level perspective when designing a conceptual optimisation model for water distribution systems. It will not be possible or desirable to develop the whole model in a single exercise, due to the complexity and size of this task. However, it is important to have a holistic view of the problem and solution to ensure that individual models can eventually merge seamlessly to provide the maximum benefit.

The conceptual model consists of four main modules, namely a user interface, optimisation module, simulation module and control module. The model is shown diagrammatically in **Figure 6.1**.

Figure 6.1: Diagrammatic representation of the conceptual optimisation model

The user controls the modelling process via the user interface. Allowance should be made for selecting the type of optimisation (or combination of optimisation types) to be performed, as well as the simulation and optimisation models to use.

The control model handles the optimisation process: It takes input from the user via the user interface, links with the required simulation and optimisation modules to run the simulation and then returns output via the user interface. The types of optimisation are added to the control module as they are developed, and are linked to allow more than one type of optimisation to be performed simultaneously.

Simulation modules include the tools used for simulation of water distribution systems to do various types of simulation. EPANET and Ooten are available software tools that can be used for hydraulic and water quality simulations, while Mocasim is used to perform stochastic analyses of water distribution systems. Other simulation modules can be added once they become available.

Finally, the optimisation modules allow for different types of optimisation to be performed including single and multiple objective GAs, fuzzy GAs and others.

Sections of the conceptual model described above have been added to the GA Utility Program software package for network analyses, which is described in detail in Chapter 8.

7. CONCEPTUAL DEVELOPMENT OF A METHOD FOR OPERATIONAL OPTIMISATION

7.1 Introduction

Pump operating costs make up a large proportion of the expenses of water supply utilities (Van Zyl, 2001). It is thus important that the operation of water distribution systems is optimized to minimize energy cost, while maintaining minimum standards of service provision and reliability. Likeman (1993) estimated that for complicated systems, cost savings in the order of 15 % can be achieved using formal, computer-based scheduling aids.

In this section, the problem of operational optimisation is discussed in detail, and a conceptual model for GA operational optimisation is developed.

7.2 Problem Definition

7.2.1 Goals and Objectives

The main objective of operational optimisation is to provide an acceptable level of service to the customer within the system constraints and legal regulations, while minimizing the operational cost (Brdys and Ulanicki, 1994; Likeman, 1993). These goals are in conflict with each other to a large extent. Attempts to minimize operational cost will generally place the system in a more vulnerable state and less able to handle abnormalities such as pipe bursts, thus reducing the level of service (Jowitt et al., 1988).

It is clear from the above discussion that it is important to set appropriate optimisation objectives. Not only does this determine the potential benefits of the analysis, it also influences the speed and complexity of the calculations, and thus of the computer resources required. It is important to produce the simplest possible meaningful statement for expressing the optimisation objectives and physical operational limitations of the system (Quevedo Cembrano, Wells, Péres and Argelaguet, 1999).

A balance has to be struck between costs and risk. Cost will play a dominant role in most operational optimisation problems, but it is also important to take level of service factors into account.

7.2.1.1 Operational Cost Objectives

According to Tarquin and Dowdy (1989) there are two reasons why considerable energy conservation may be possible in municipal water supply systems:

- Networks are often created by being pieced together through a series of expansion projects brought about by changing populations and urban sprawl. The overall operational optimisation of the system is often not considered when these extensions are made.

- It is common for large pumps to remain in service for 20 years or more. Wear and tear and poor maintenance can result in situations where the pump efficiencies are reduced to such an extent that different operation procedures could ensure significant reductions in water pumping costs. In most water distribution systems, the greatest potential reductions in operational cost can be made by scheduling the pumps in the system to minimize electrical energy costs (Ormsbee and Reddy, 1995). For this reason the term *pump scheduling* is often used as substitute for operational optimisation. However, other possibilities for reducing the operational cost include:

- Using the cheapest water source.

- Minimizing water losses. Night-time pumping increases pressure and thus leakage in the system and this should be included in the evaluation of the operational plan (Jowitt et al., 1988).

- Maximizing the stability of pump operation, i.e. minimizing the variation in pump load which will move the pump's working point away from the maximum efficiency point (Ko et al., 1997).

- Minimizing the number of pump switches. Lansey and Awumah (1994) found that pump switching is in many cases as important, if not more so, than operating cost in pump scheduling.

- Minimizing the maximum power demand over a given period. Electricity costs normally consist of a *consumption charge*, based on the actual electricity consumption, and a *capacity charge*, based on the maximum power consumption in a given period. Minimizing the maximum power demand would thus also minimize the capacity charge.

The most common operational optimisation problem deals with the operation of pumps to minimize the total energy cost. For this reason the conceptual GA development will be restricted to controlling pump operation to minimize energy costs. The price of water, water losses and the capacity charge will not be included in the initial method development.

7.2.1.2 Level of Service Objectives

Handling the level of service in the operational optimisation problem is not always easy since the level of service is made up of a complex amalgam of factors including consumer demand, pressures, pumps and leakages (Jowitt et al., 1988).

In most cases level of service objectives are in direct conflict with the objective of minimizing the operational cost. Some of the more common level of service objectives are briefly discussed below.

- Water quality objectives (Forzini, 1991; Percia et al., 1997). Water quality is often a function of the time water spends in the system before being consumed. Large service reservoirs may often increase retention time and thus negatively influence water quality. Different sources may also provide different water qualities which have to be matched to the quality requirements of the system users.

- System reliability and security of supply objectives (Forzini, 1991). Reliability and security of supply is often achieved by increasing the levels of storage in the system, which is in direct conflict with the water quality requirement.

- System robustness objectives. System robustness is measured by the ability of the system to handle stochastic variations such as fluctuations in system demand (Ko et al., 1997) and emergencies such as pipe failures or fires. The distribution system should be able to provide for these events and recover to normal operation swiftly following such emergencies.

Jarrige et al. (1991) stated that although optimisation is normally used for a range of normal operating conditions, it is particularly beneficial under critical situations, when even trained staff cannot rely on their personal experience to get the appropriate answer to enable them to manage and control the system. They found that operators will often over-prioritise security and consequently increase operating costs under these conditions.

Level of service objectives will be handled as constraints. In other words, the user will set up a water distribution system with minimum levels of service and optimise the operation of the system for this system. The effect of level of service on the operational cost can be ascertained by running the optimisation for a new set of minimum levels of service.

7.2.2 Problem Formulation

Operational optimisation problems may vary greatly in size and complexity, and different optimisation techniques may be appropriate for different problems. However, for all optimisation problems it is possible to define an objective function, decision variables and constraints (Brdys and Ulanicki, 1994).

The objective function is the function that has to be minimized in the optimisation. For water distribution systems, the objective function is normally defined by the operational cost of the system over a period of time (typically 24 hours). The operational cost, in turn, is calculated by doing a dynamic simulation of the system. Level of service objectives can be incorporated in the objective function in a multi-objective optimisation, but it is more common to incorporate these objectives as constraints.

The decision variables used in operational optimisation may vary considerably depending on the available control systems. Pumps can be switched directly at certain times, or indirectly for example when service reservoir levels or pressures in the system reach certain trigger levels.

Various constraints are imposed on the operation of water distribution systems. These vary from physical constraints such as conservation of mass and energy to constraints imposed by regulation or policy, such as many level of service constraints. Constraints may be included in the operational optimisation as either hard or soft constraints:

- *Hard constraints*. Hard constraints have to be satisfied for a solution to be valid. These constraints are very rigid in nature and solutions are rejected even if a single constraint is only marginally violated.

- *Soft constraints*. Solutions that violate soft constraints are not excluded from the set of possible solutions, but penalized using an appropriate penalty function. Limited multi-objective optimisation may be performed by weighting the penalty functions in relation to their importance in the operational process. This method can, however, only be used if the relative importance of the different objectives is known beforehand.

In analysing a particular aspect of water distribution systems it is important not to lose sight of the wider implications and interactions with the rest of the system. Operational optimisation over a 24-hour period may, for instance, result in an operational policy that is far from optimal in the long term by ignoring increased maintenance as a result of a great number of pump switches.

There are also strong links between design and operation and a small alteration to a design may have large implications for operational optimisation. However, these interactions fall largely outside the scope of this study.

7.2.3 Complexity of the Problem

Operational optimisation of water distribution systems is a complex task. A number of factors significantly complicate operational optimisation, including (Nitivattananon, Sadowski and Quimpo, 1996; Yu, Powell and Sterling, 1994; Shamir, 1985):

- The underlying complexity of the problem. Water distribution systems often have complex physical structures with mainly nonlinear elements and this makes global optimal control difficult and time-consuming to achieve. The objective function is often nonlinear and the control variables are implicitly and nonlinearly related to the state variables. The existence of reservoirs and non-constant electricity tariffs introduces very strong dynamic features in the control problem.

- Computational requirements. The mathematical problem can be very large in the number of nonlinear constraints and decision variables, especially when large systems and long planning periods are considered. This translates into large computer resource requirements and long solution times which may limit the size of problem that can be addressed.

- Mixed-integer nature of the problem. When pump speeds are considered fixed, the solutions for pump discharges are a discrete set of feasible operating points. There may also be limitations in the number of times pumps are switched. In Mathematical Programming, the computational complexity of discrete problems is much greater than that of continuous ones. The computational complexity of the latter is polynomial, whereas the former's is exponential (Ulanicki et al., 1997).

- Complex electricity charge. Apart from the normal charge for electricity consumed, suppliers also levy a fee based on the maximum power consumption in a given demand period, known as a capacity charge. This charge is difficult to incorporate into a daily optimal scheduling problem, while a long control period (say of a month) would often make the scheduling problem dimensions too large for practical analysis.

Various approaches have been suggested to reduce the complexity or size of the operational optimisation problem. However, in many cases this limits the optimisation method to a single distribution system (Nitivattananon et al., 1996) and much effort is required to apply the method to each new system to be analysed.

7.2.4 Optimisation techniques

Various optimisation techniques have been applied to the operational optimisation problem, including linear programming (Burnell, Race and Evans, 1993; Jowitt and Germanopolous, 1992), nonlinear programming (Yu et al., 1994; Chase and Ormsbee, 1993), dynamic programming (Nittivattanonnon et al., 1996; Lansey and Awumah, 1994), fuzzy logic (Angel, Hernandez and Agudelo, 1999), nonlinear heuristic optimisation (Ormsbee and Reddy, 1995; Leon, Martin, Elena and Luque, 2000), flexible constraint satisfaction (Likeman, 1993) and genetic algorithms (Savic et al., 1997, Boulos, Orr, De Schaetzen, Chatila, Moore, Hsiung and Thomas, 2001).

The vast majority of methods listed cannot handle the complexities of the optimisation problem discussed in the previous section. In order to apply these methods, the optimisation problem is simplified by implementing assumptions, discretization or heuristic rules. Such simplification makes it easier for specific optimisation methods to determine an optimal solution, but introduces bias into the solution by excluding a large number of potentially good solutions.

Genetic algorithms (GAs) are not bound by most of the limitations of other optimisation methods and can handle the complexities of the operational optimisation problem without simplifications. This gives GAs a substantial advantage over the other optimisation methods in finding the optimal solution. The proven efficiency and robustness of GAs make them the ideal optimisation method for operational optimisation (Van Zyl, 2001).

7.2.5 Practical Implementation Issues

Optimized operational plans for water distribution systems only find meaning once they are successfully implemented in the system. One factor that is consistently mentioned in papers on the implementation of operational schemes is that of acceptance by the system operators. This factor often frustrates the optimisation project to such an extent that it is never realized. Various reasons have been put forward for lack of operator acceptance (Hsu, Louie and Yeh, 1992):

- most operators have not been directly involved in the development of computer models and are thus not entirely comfortable with using them
- most papers on operations deal with simplified systems, are difficult to adapt for use in real systems and are poorly documented from a practical point of view;
- institutional constraints exist that impede user-research interaction; and
- operational models are often not user-friendly and not conducive to changes.

Shamir (1985) suggested that if software for optimal control is to be successfully developed, it should be introduced cautiously, delegating to automatic control only well tested and proven parts of the overall operational control.

7.3. Conceptual Development of GA Optimisation Methodology

7.3.1 Optimisation Variables

The optimisation variables used in this study are the parameters through which the operation of the pumps in the system is controlled.

The two most common types of controls are reservoir level controls and time-based controls. In reservoir level controls, control actions are triggered at certain reservoir water levels. For example, a pump will be switched on when a reservoir is empty and switched off when the reservoir is full. Reservoir level controls are normally used in pairs, with the one control triggering an action (such as switching a pump on) and the other control triggering an opposing action (switching the pump off). The lower of the two trigger levels is called the *low level control*, and the higher the *high level control*.

In time-based controls, pumps are switched on and off at certain times of the day. Only reservoir level and time-based controls will be considered in the study. The reason for this is that these types of controls are widely used in practice, due to their simplicity and proven robustness.

7.3.2 System Constraints

To ensure the viability of solutions in the optimisation process, certain constraints have to be imposed on the variables and the operational behaviour of the system.

7.3.2.1 Variable Constraints

Reservoir level controls must be constrained to the valid reservoir water levels. The maximum value it can take is determined by the reservoir's full water level, and the minimum value by the reservoir's emergency storage requirement. In pairs of controls, the high level control variable must also be higher than the low level-control variable.

7.3.2.2 Operational Constraints

The operational cost, or total pump energy cost, of each set of variables is calculated by doing an extended period simulation of the system. To ensure convergence on a viable solution, two operational constraints have to be applied. The first constraint is that the reservoir water levels must balance over the run. In other words, the levels at the end of the run should not be lower than the levels at the start of the run. The second constraint is a limit on the number of pump switches in a 24-hour run, to avoid increased maintenance costs due to excessive wear-and-tear on the pumps.

The operating constraints can be implemented as soft constraints by adding penalty costs to the objective function. The reservoir level penalty cost should be based on the total deficit volume at the end of a simulation, and the pump penalty cost on the number of pump switches. To be effective, the penalty unit costs have to be high enough to assist convergence on a viable solution, but not so high that they eliminate potentially good solutions.

7.3.3 Objective Function

The objective function is simply the total energy cost of a simulation plus the reservoir level and pump switch penalty costs. It will be necessary to experiment with different penalty unit costs to ensure effective convergence.

7.3.4 Genetic Algorithm Formulation

It is proposed that a well established GA library is used for implementing the optimisation. Important elements of the GA are the encoding of the problem, calculation of the solution fitness values and the termination conditions.

7.3.4.1 Encoding

The optimisation variables of the GA will be encoded in bounded, real-valued chromosomes, with each gene representing a single variable. The number of genes in the chromosome will thus be equal to the number of variables used. To decode a chromosome into a set of reservoir level controls, the process is simply reversed. **Figure 7.1** illustrates the coding and decoding of an GA optimisation with two reservoir level controls and four time-based controls.

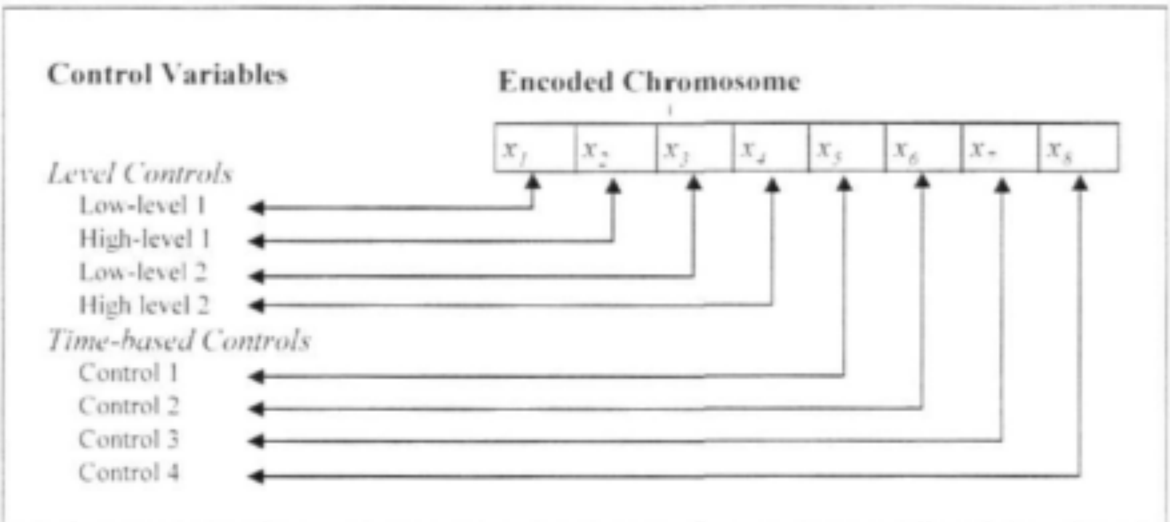


Figure 7.1: Encoding and decoding of control variables to and from a chromosome

7.3.4.2 Fitness Function

The GA code that was used was developed for maximizing fitness values. Since the operational optimisation problem is a minimization problem, the fitness function had to be defined so that it will increase with a decrease in the objective function value. This can be achieved by using the relationship:

$$Fitness = \frac{1}{Objective}$$

7.3.4.3 Termination Conditions

In practice the time available for decision-making in the operation of water distribution systems is limited. Moreover, the implementation of a given operational strategy is restricted by the accuracy of instrumentation and uncertainties in demand forecasting. It is therefore more important to find a good solution in a short period of time than to find the true global optimum.

The most important termination condition was thus defined as the time available for decision-making.

7.3.4.4 Parameters

A previous study (Van Zyl et al., 2003b) applied the following parameters for operational optimisation of water distribution systems (see **Table 7.1**):

Table 7.1: Parameters for operational optimisation of water distribution systems

GA Parameter	Value
Type	steady-state
Parent selection	SRSWR
Crossover	one point random
Replacement	weakest organism
Crossover rate per event	90 %
Mutation rate per organism	90 %
Population size	35 organisms

The above values can be used as starting point for the development of the GA, and can be adjusted after experimentation to ensure the best GA performance.

7.4 Conclusions

In this chapter, a conceptual method for GA operational optimisation of water distribution systems was developed. It was concluded that GAs are the only optimisation technique capable of handling the true complexities of the operational optimisation problem. An objective function, constraints and parameters for the GA were proposed.

8. EXTENSIONS OF THE GA WATER UTILITY PROGRAMS

8.1 Introduction

This section provides a description of the implementation and procedures that had been incorporated in the development of a simple pipeline optimisation program called GAPOP and a network optimisation program called GANEO. The programs grouped together as the GA Water Utility Programs, see **Figure 8.1**, and are to be used as an educational tool to transfer the use of genetic algorithms in the water field.

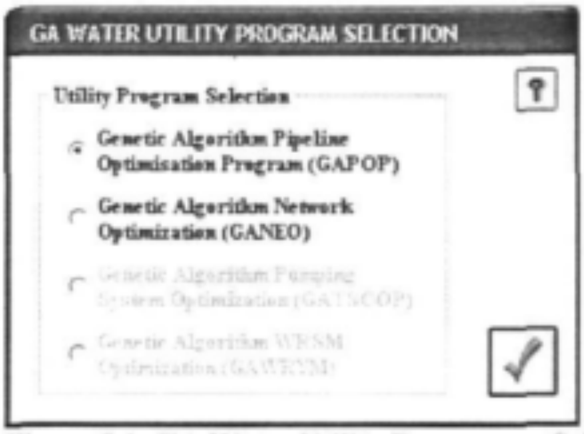


Figure 8.1: GA Water Utility Programs – Selection options

It is important that the operation of water distribution systems is optimized to minimize energy cost, while maintaining minimum standards of service provision and reliability. Likeman (1993) estimated that for complicated systems, cost savings in the order of 15 % could be achieved using formal, computer-based scheduling aids. In Chapter 7, the problem of operational optimisation is discussed in detail, and a conceptual model for GA operational optimisation is developed which is being implemented in the Genetic Algorithm Pumping System Optimization (GATSOP) program. The GATSCOP program is unfortunately not yet ready for distribution

8.2 Genetic Algorithm Water Utility Programs

8.2.1 GAPOP

The Genetic Algorithm Pipeline Optimisation Program (GAPOP) was developed during the previous research study. Further extensions/improvements that have been made include the following:

- The GAPOP program has been extended to handle 100 pipes.
- The hydraulic engine that was developed in Visual Basic has now been replaced with the EPANET engine. This improved the stability of the software as well as increasing the analysis capabilities. The time of executing an number of iterations has also drastically reduced, i.e. the analysis runs are faster.

  A detailed description of all the input screens and results was provided in the previous report and is not repeated here (Van Vuuren et al., 2001).

8.2.2 GANEO

The objectives of the network optimisation program, called GANEO are:

- Provide access for the novice in the use and understanding of the working of GA techniques;
- Determine an optimal solution quickly for a water distribution system

The conceptual design as discussed in the Chapter 6 was implemented in this program. The steps a designer would follow to optimize a water distribution system is as follows:

- Step 1: Using EPANET, create a working network model (*.net file) and export the water distribution system for importing in GANEO (*.inp file)
- Step 2: Create a new project in GANEO and importing the EPANET model
- Step 3: Create a pipe selection file in GANEO
- Step 4: Set the boundaries and GA parameters in GANEO
- Step 5: Run the optimisation analysis procedure
- Step 6: Evaluate the alternative options and export the result back to EPANET

Step 1: Using EPANET, create a working network model (.net file) and export the water distribution system for importing in GANEO (*.inp file)*

The free public domain software package EPANET is used to set-up a network. Once it has been successfully analysed to ensure that it is a working hydraulic model it must be exported to an input file that can be read by the GANEO program. This is done by clicking on *File | Export | Network....* and saving it with a *.inp file extension, see **Figure 8.2**.

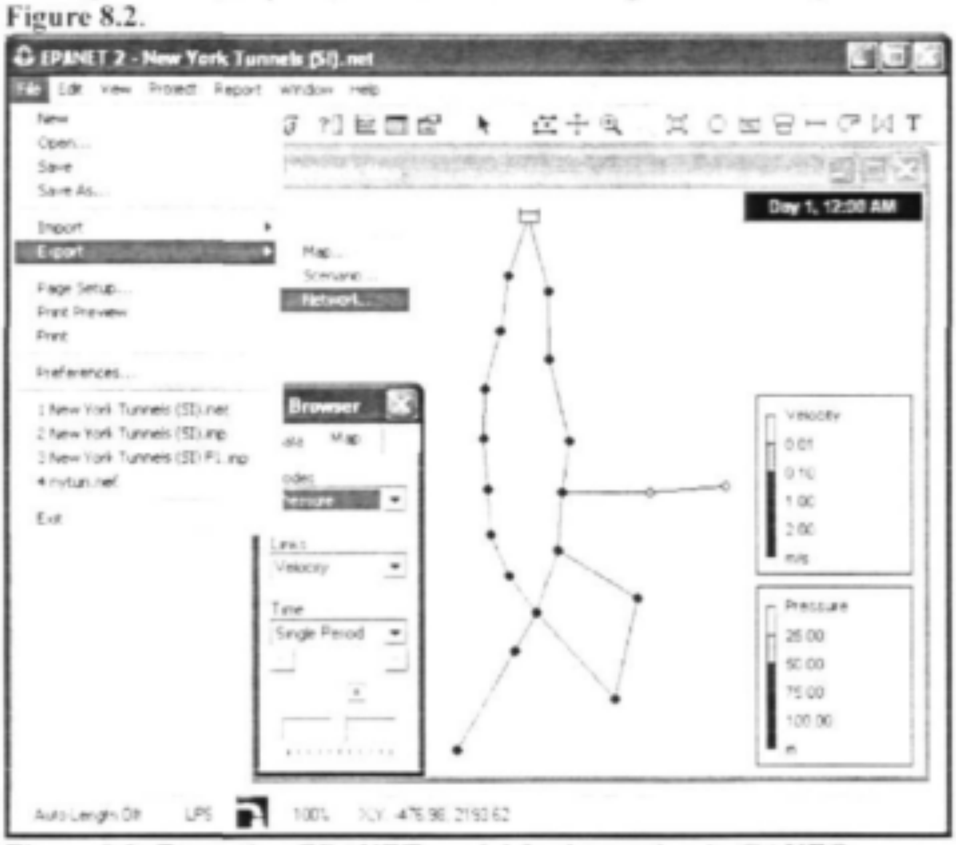


Figure 8.2: Exporting EPANET model for importing in GANEO

Step 2: Create a new project in GANEO and importing the EPANET model

To create a new project in GANEO that will store all the input data as used in the GA process select *File | New project* and enter a relevant filename such as "New York Tunnels Example (SI).gan".

To import/load the created EPANET model into GANEO click on the *Load file* button (see **Figure 8.3** below) to load the input file (for example "New York Tunnels (SI).inp). The created file can also be viewed in its text format by clicking on the *View file* button or a graphical layout of the system by clicking on *View Network*.

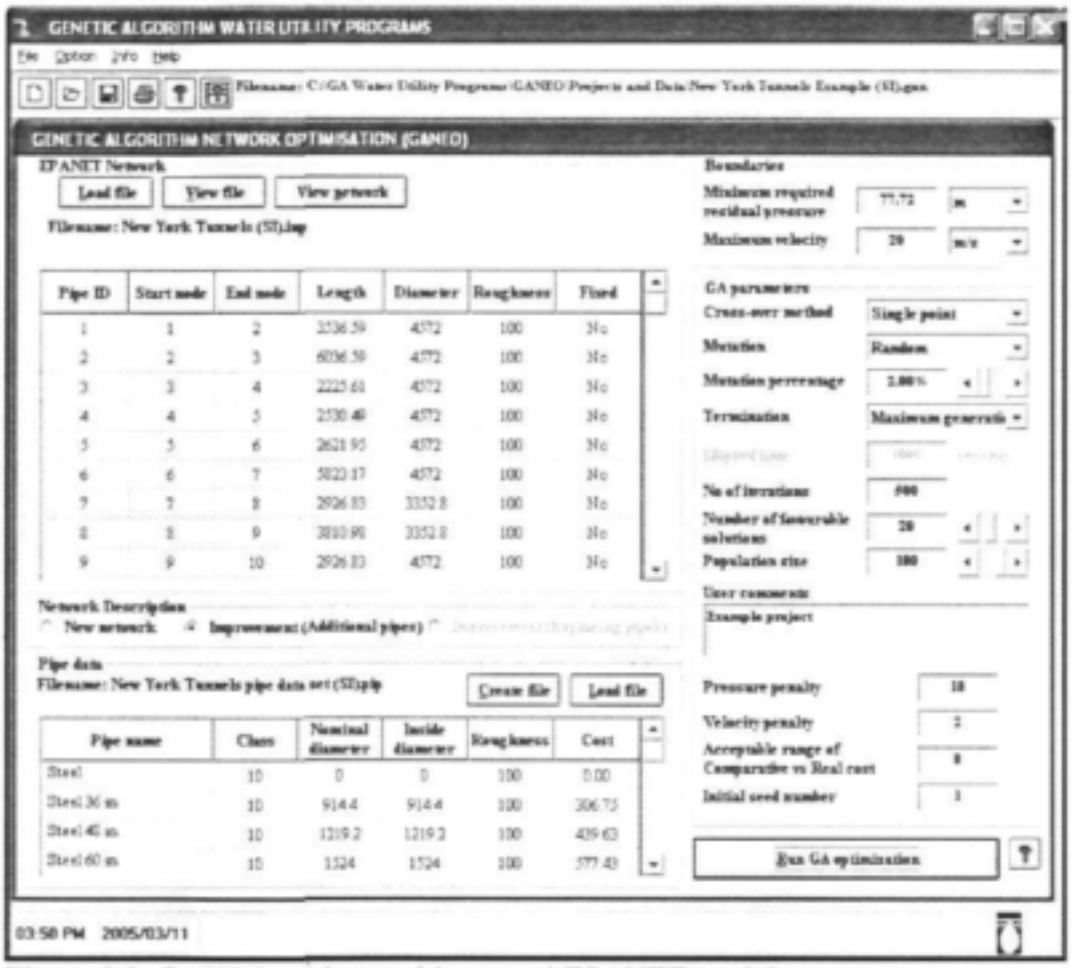


Figure 8.3: GANEO project and imported EPANET model

As can be seen in **Figure 8.3**, the EPANET file is shown in table format, detailing the pipe IDs, nodes, lengths, diameters and roughness values. The last column, with heading "Fixed", is used to allow the user to control the optimization process of specific pipes by preventing the program from changing a pipe diameter. Pipes 3 and 4 have been fixed by double clicking in that column. In other words the program will keep the diameters of pipes 3 and 4 unchanged throughout the GA process. Existing pipes or connection can in this way be excluded from the optimization process.

The program currently allows for two types of networks to be analysed, "New network" and "Improvement (additional pipes)" as shown in **Figure 8.4** below.

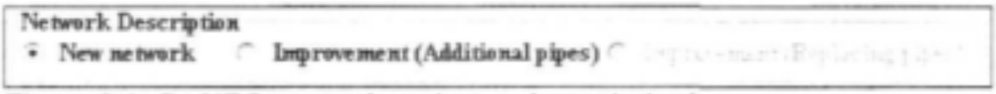


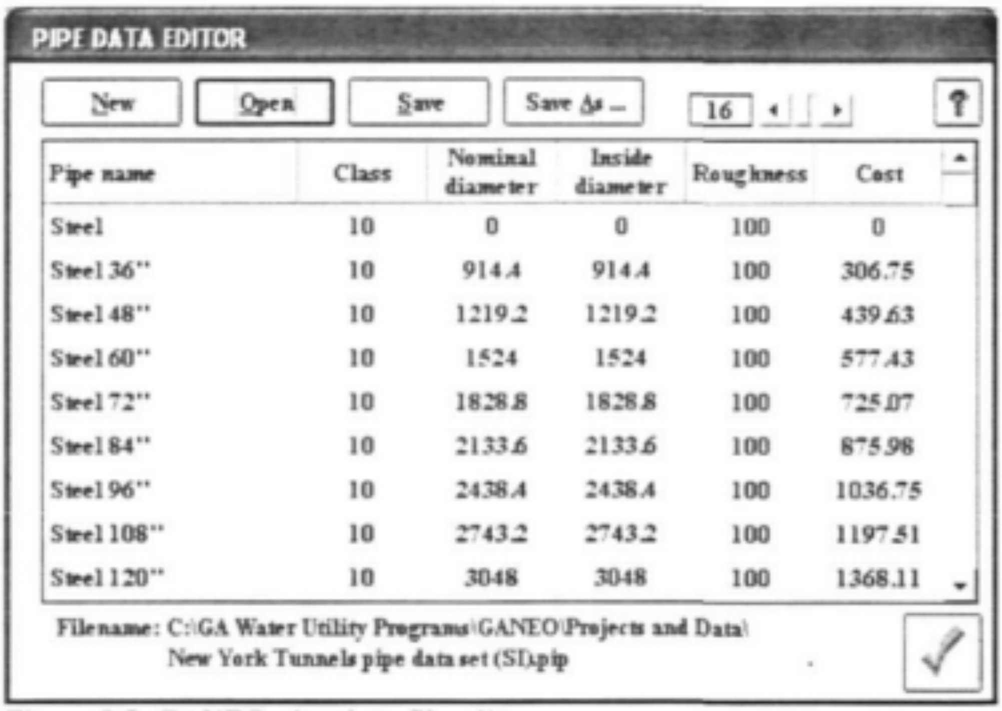
Figure 8.4: GANEO network options to be optimized

The main difference between the various options is in the end objective. When designing a new network, pipe diameters needs to be determined for the entire system. When there are demand shortages or pressure deficiencies in a system a number of parallel pipes at strategic positions in the network will improve the system. The existing pipes are thus not replaced but additional pipes are added to the network. The last option, "Improvement (Replacing pipes)", which will be added later allows for improvement of the system by replacing certain pipes in the network with newer pipes.

Step 3: Create a pipe selection file in GANEO

A pipe data set is required for the GA procedure. From this data set the program will select the various pipe options and populate the network system with these. In the fitness function the cost based on the obtained pressures and velocities versus that required will be calculated and used as a criteria in selecting the optimum system. This data set is in the form of a pipe data file and is created by clicking on the *Create File* button as shown in **Figure 8.3**.

The data input required for this input file is shown in **Figure 8.5**.



Pipe name	Class	Nominal diameter	Inside diameter	Roughness	Cost
Steel	10	0	0	100	0
Steel 36"	10	914.4	914.4	100	306.75
Steel 48"	10	1219.2	1219.2	100	439.63
Steel 60"	10	1524	1524	100	577.43
Steel 72"	10	1828.8	1828.8	100	725.07
Steel 84"	10	2133.6	2133.6	100	875.98
Steel 96"	10	2438.4	2438.4	100	1036.75
Steel 108"	10	2743.2	2743.2	100	1197.51
Steel 120"	10	3048	3048	100	1368.11

Figure 8.5: GANEO pipe data file editor

The same pipe data set can thus be used for various projects and does not require re-entering of all the pipe data.

Step 4: Set the Boundaries and GA parameters in GANEO

Once the network system and the pipe selection data set have been loaded, the boundaries and GA parameters must be selected.

Boundaries

The boundaries that have to be entered are required to test the fitness function, see **Figure 8.6**. The user is required to enter the minimum residual pressures for the water distribution system. Once a hydraulic analysis has been performed the fitness function will test all the nodes in the system to establish if these nodes meet the minimum pressure criteria. In a similar way the fitness function will test the obtained velocities against the maximum velocity that was entered. If there are any failures of pressure or velocity the program will apply a penalty factor to the capital cost depending on the degree of failure. The preferred units of the constraints can be set on the drop down boxes next to the input boxes.

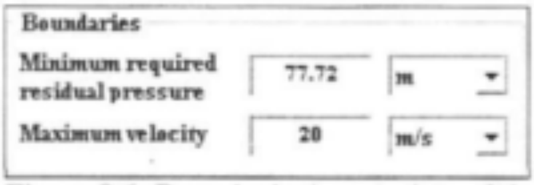


Figure 8.6: Boundaries/constraints of the system

GA parameters

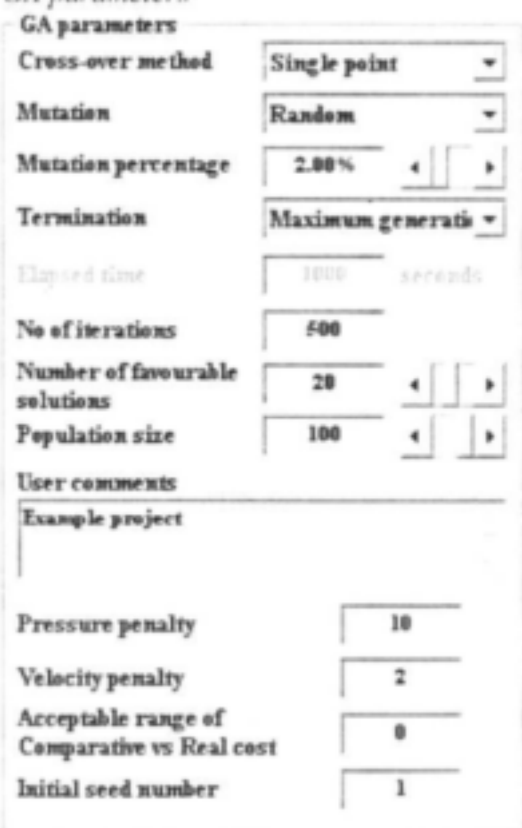


Figure 8.7: GA parameters of the system

- Selection procedures

  There are a number of different selection procedures for picking the parents from the populations that have been generated or analysed, based on their fitness values which are then used by the crossover and mutation processes to produce two offspring for the new population. The higher the fitness values the higher the probability of that chromosome being selected for reproduction. Some of these selection procedures are:

  - *Top mate*

    The first parent is selected by the fitness order. The second parent is selected randomly.

  - *Roulette wheel (rank/cost)*

    This selection method calculates the fitness (cost), ranks the chromosome (solutions) and according to their rank defines the chance of a chromosome to be selected.

  - *Tournament*

    With this selection method, a small subset of chromosomes is selected and the one with the best fitness will become a parent.

  - *Random*

    This is a very simple method where the parents are simply selected randomly.

In the GANEO program the selection procedure was hard coded as Roulette wheel selection. The chromosomes are ranked according to their fitness and based on their proportional probability are individually selected as parents.

The following GA parameters are required to define the GA process. (**Figure 8.7**):

- Cross over method

  After two parents have been selected by the selection method, crossover takes place. The crossover in carried out according to the crossover probability. The crossover methods implemented in the GANEO program are:

  - *Single point*

    A random crossover point is selected. The first part of the first parent is swapped with the first part of the second parent to make the offspring.

  - *Double point*

    The double point crossover operator differs from the single point crossover in the fact that two crossover points are selected randomly.

  - *Uniform*

    In the uniform crossover each gene is selected randomly, either from the first parent or from the second one.

- Mutation procedure
  Mutation is the genetic operator that randomly changes one or more of the chromosome's gene. The purpose of the mutation operator is to prevent the genetic population from converging to a local optimum and to introduce to the population new possible solutions. The mutation is carried out according to the mutation probability. The mutation methods implemented in the GANEO program are:
  - *Random*
    The random mutation operator exchange's a random selected gene with a random value within the range of the gene's minimum value and the gene's maximum value.

  - *Min-max*
    The min-max mutation operator exchange's a random selected gene with the gene's minimum value or with the gene's maximum value, selected randomly.

- Mutation percentage
  The mutation percentage sets the probability for a string to be mutated. The probability should be between 0% and 100%. If set to 0% no mutation will take place and if set to *100%* the search procedure becomes a complete random process since every obtained solution will be mutated and thus not carrying over any characteristics of the previous analysis. The default mutation rate is 2%.

- Termination
  The termination method determines when the genetic process will stop evolving. The termination methods implemented in the GANEO program are:
  - *Maximum generations*
    The genetic process will end when the specified number of generation's have evolved as entered by the user.

  - *Elapsed time*
    The genetic process will end when a specified time has elapsed.
    If the number of iterations (generation) has been reached before the specified time has elapsed, the process will end.

  - *No change in fitness*
    The genetic process will end if there is no change to the population's best fitness for a specified number of iterations (generations). If the number of iterations (generations) has been reached before the specified number of iterations with no changes has been reached, the process will end.

- No of iterations
  Enter the maximum number of generations that should be carried out by the program. Once the maximum number of generations has been reached the GA process will end. The program does however allow the user to run a further process starting, where the previous process has stopped i.e. continuing where the previous process ended.

- Number of favourable solutions
  The program only stores a limited number of the best solutions, called favourable solutions. The user has the option of selecting how many favourable solutions should be stored. The default value is 20.

- Population size
  The population size can be set to a value between 4 and 100. The population size is equal to the number of possible solutions that will be generated with every generation (iteration). From these parents new offspring will be made. The default value is 20.

- User comment
  This is just a general user comment, which can be used to describe the current optimization being performed.

- Pressure penalty
  When a node does not meet the pressure criteria the system is penalised. The penalty component (cost) is added to all the pipes that supply that specific node based on their relative contribution in terms of flow. The user defines this variable, but the program will suggest certain default values. This penalty cost will be a multiple of the actual cost for each pipe. The penalty that is applied is dependent on the degree that a solution fails. The following example will demonstrate the calculation of the penalty cost component:

  Example of penalty calculation
  Consider 3 pipes, A and B, supplying Node 1, and C continuing from Node 1 to the next node (**Figure8.8**).

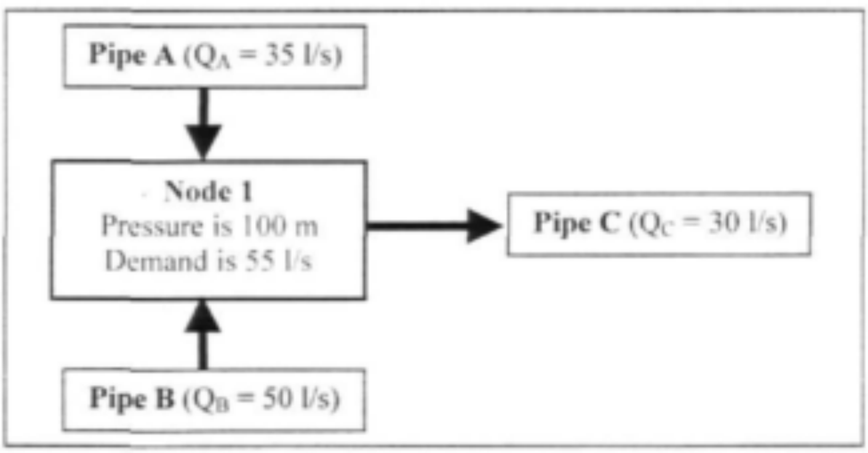


Figure 8.8: Demonstration of determining penalty cost for a node pressure deficiency

The node pressure is 21 m, 3 m less than the specified minimum residual pressure of 24 m, and the demand at the node is 55 l/s. Pipe A and B supplying Node 1 would then both be penalised using the following formula:

$$PC_{H\,min-i} = C_{pp}^{PFP} \; x \; L_p \; x \left(\frac{P_{required} + \left(P_{required} - P_{calculated}\right)}{P_{required}} \right) x \left(\frac{Q_i}{Q_{node}} \right)$$

Where:

PC_{Hmin-i} = Penalty cost for pipe "i" due to minimum pressure
C_{PP} = Cost of the pipe per unit length
PFP = User specified penalty factor for pressure
L_P = Length of the installed pipe
$P_{calculated}$ = Calculated pressure at node
$P_{required}$ = Minimum residual pressure required at node
Q_i = Flow in pipe "i"
Q_{node} = Total flow into node

If for example the penalty factor (PF) is 0.5, pipe cost (C_P) for pipe A is R150/meter and the pipe length (L_P) is 500 m the additional penalty cost can be calculated as follows:

$$PC_{H\min-i} = (150)^{0.5}\ x\ (500)\ x \left(\frac{24 + (24-21)}{24}\right) x \left(\frac{55}{85}\right) = R44\ 57.71$$

The pipe cost of pipe A is R75 000.00 without the penalty and R79 457.71 with additional penalty cost.

The aim of the penalty cost structure as defined above is to penalise a system more the greater the pressure deficiency is and to add some proportional distribution of the flows in the pipes to the cost. The more water a specific pipe supplies to the node the greater the importance of the pipe. The higher the user specified penalty factor (PF) is the higher the cost component will be.

Please note that the pressure penalties are subject to an if–then–else statement, that means if the values fall within the specified intervals, no penalties would be applied.

In the case of a network improvement optimization it might happen that no new pipes are supplying the node that is experiencing pressure failures. In this case the pipe with lowest cost (except zero) in the pipe data set is used.

In the case of a node pressure becoming negative the PFP is set to 50 thus adding a heavy penalty cost component to the system.

- Velocity penalty

  A similar approach is followed with pipes not meeting the maximum velocity constraint. The velocity penalty is calculated as follows:

$$PC_{V\max-i} = C_{PV}^{\ PFV}\ x\ L_P\ x \left(\frac{V_{\max} + (V_{calculated} - V_{\max})}{V_{\max}}\right)$$

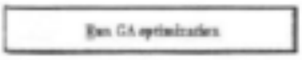
Where:

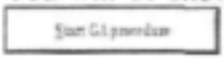
PC_{Vmax-i} = Penalty cost for pipe "i" due to maximum velocity
C_{PV} = Cost of the pipe per unit length
PFV = User specified penalty factor for velocity
L_P = Length of the installed pipe
$V_{calculated}$ = Calculated velocity in pipe
V_{max} = Maximum velocity in pipe

- Acceptable range of Comparative vs Real cost
  When calculating the optimum system good solutions are sometimes discarded even if it is only due to an insignificant pressure failure at one node. An attempt to keep these "near good solutions" a part of the GA process a new factor was incorporated. Before discarding a solution a further check is done on the Comparative vs Real cost. If these are the same it implies that there were no penalty costs added and thus no failures occurred. If the Comparative and real cost differs by only 1% it means that a "near good solution" was obtained and it is left to the user to decide whether or not to accept this solution. The closeness/acceptable variation between the Comparative and Real costs can be set by changing this variable.

- Initial seed number
  The GA procedure is driven by a random number generation component. The software utilizes a pseudo random generator and for any given initial seed, the same number sequence is generated because each successive call to the random function uses the previous number as a seed for the next number in the sequence. The user can select any initial seed to start the optimization process.

Step 5: Run the optimisation analysis procedure

Click on the 

button on the Input Screen (see **Figure 8.3**). You will be shown the Optimisation Screen as shown in **Figure 8.9**. Click on the 

button to start the procedure.

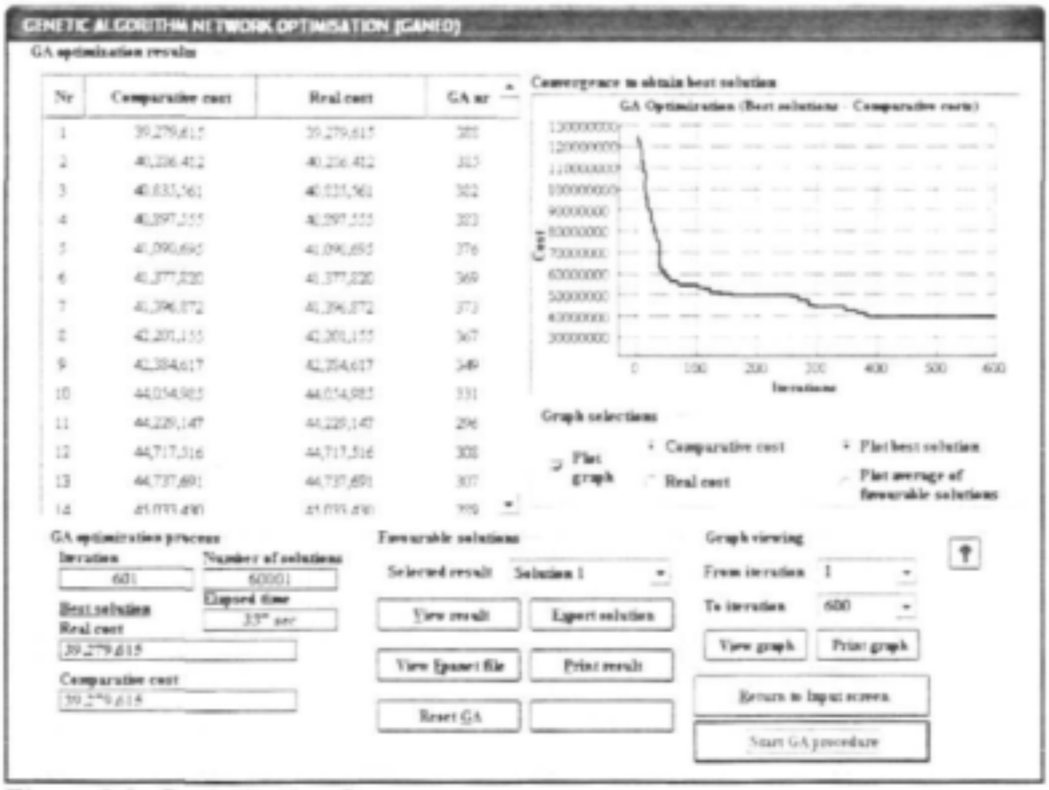


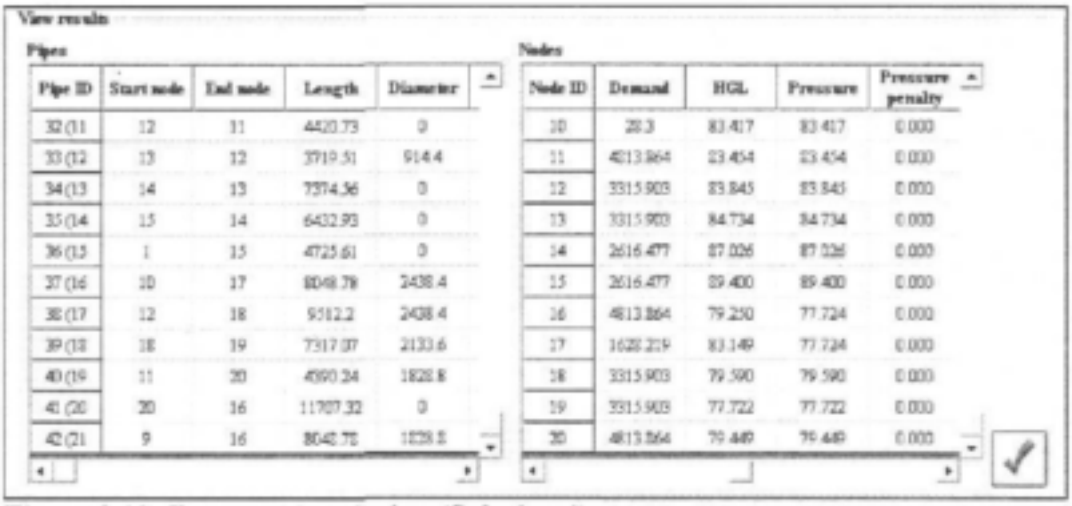
Figure 8.9: Optimization Screen

The program will run the Genetic Algorithm developed for optimising a network and show the progression thereof on the convergence graph, which can be activated with the *Plot Graph* selection box.

In the left bottom corner the iteration (generation) number, number of solutions, elapsed time and real and comparative costs are shown. The table on the left shows the most favourable solutions.

Step 6: Evaluate the alternative options and export the result back to EPANET

After the optimisation is complete the user can select any of the favourable solutions and view the specific solution's result, as shown in **Figure 8.10,** by using the drop down list of Favourable solutions. The results will provide all the information of the system such as pressures, velocities, applied penalties and very importantly what pipe is required and where in the system.


View results

Pipes

Pipe ID	Start node	End node	Length	Diameter
32 (11	12	11	4420.73	0
33 (12	13	12	3719.51	914.4
34 (13	14	13	7374.36	0
35 (14	15	14	6432.93	0
36 (15	1	15	4725.61	0
37 (16	10	17	8048.78	2438.4
38 (17	12	18	9512.2	2438.4
39 (18	18	19	7317.07	2133.6
40 (19	11	20	4390.24	1828.8
41 (20	20	16	11707.32	0
42 (21	9	16	8042.78	1828.8

Nodes

Node ID	Demand	HGL	Pressure	Pressure penalty
10	28.3	83.417	83.417	0.000
11	4813.864	83.454	83.454	0.000
12	3315.903	83.845	83.845	0.000
13	3315.903	84.734	84.734	0.000
14	2616.477	87.026	87.026	0.000
15	2616.477	89.400	89.400	0.000
16	4813.864	79.250	77.724	0.000
17	1628.219	83.149	77.724	0.000
18	3315.903	79.590	79.590	0.000
19	3315.903	77.722	77.722	0.000
20	4813.864	79.449	79.449	0.000

Figure 8.10: Favourable solution (Solution 1)

The user also has the option to export the selected solution, thus creating an EPANET input file of the proposed solution. The solution can thus be analysed and engineered as required using the EPANET hydraulic analysis software package.

8.2.3 GATSCOP

A further extension to the GA Water Utility Programs is the Genetic Algorithm Pumping System Optimization (GATSCOP) program.

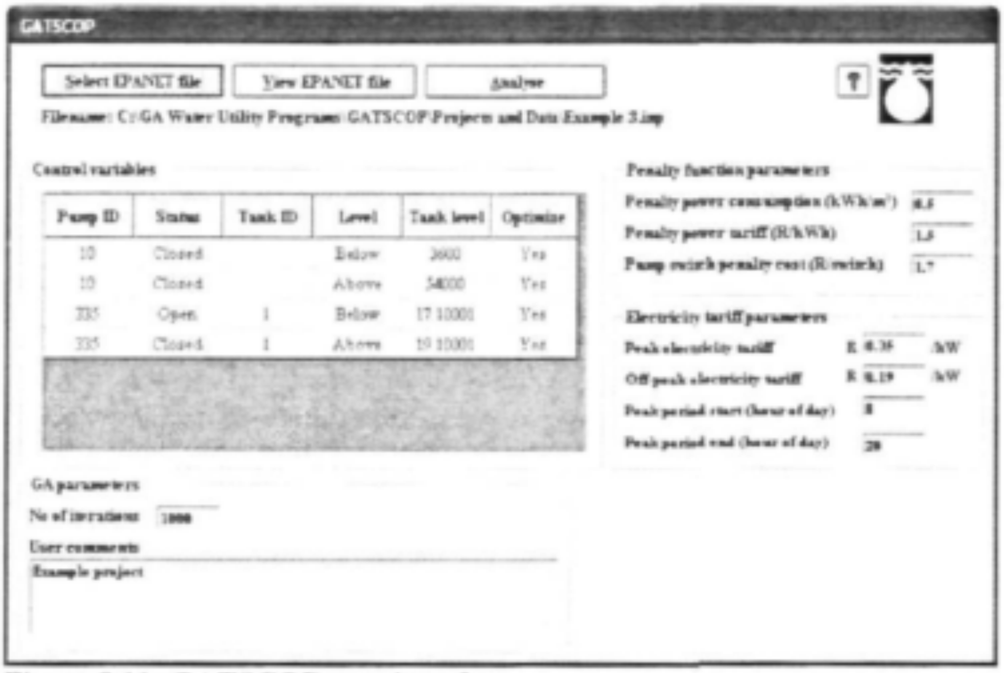


Figure 8.11: GATSCOP user interface

The user will have the option to load an EPANET water distribution model, which contains various controls, pumps, reservoirs and their status or levels, see **Figure 8.11**. The penalty function parameters can be set as well as the electricity tariff parameters. The GA operational optimisation as detailed in Chapter 7 will then be executed to optimize the operational control variables of the system. This software will be completed in the next phase of the research study.

8.3 Examples

The developed Genetic Algorithm procedures had to be tested against benchmark problems in order to test its functionality and efficiency. The GANEO program was tested against some common networks for which many optimizations have been performed, these include traditional and heuristic methods. Two of the more common ones are the New York Tunnels (Example 1) and Hanoi (Example 2) detailed below (Centre for Water Systems, 2004).

8.3.1 GANEO – Example 1: New York systems problem

8.3.1.1 Description of system

The New York water supply system has been studied by a number of researchers in the pipe network optimisation field. The aim of the various studies was to determine the most economical design to improve the existing system of tunnels that constituted the main water distribution system.

Figure 8.12 is a general layout of the system indicating the pipes (**Table 8.1**), nodes (**Table 8.2**) and supply reservoir. The existing system was found to be inadequate due to aging and an increase in demands for the anticipated demands in terms of the pressure requirements.

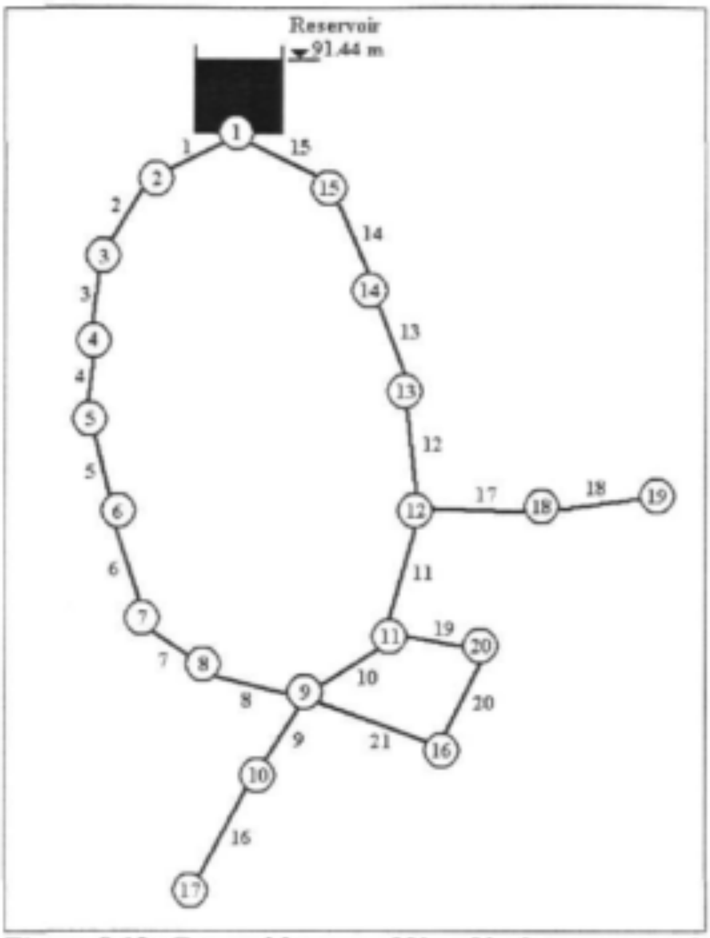


Figure 8.12: General layout of New York water supply system

Table 8.1: Pipe data of the existing system/network

Pipe	Length (m)	Diameter (m)
1	3536.59	4.57
2	6036.59	4.57
3	2225.61	4.57
4	2530.49	4.57
5	2621.95	4.57
6	5823.17	4.57
7	2926.83	3.35
8	3810.98	3.35
9	2926.83	4.57
10	3414.63	5.18
11	4420.73	5.18
12	3719.51	5.18
13	7347.56	5.18
14	6432.93	5.18
15	4725.61	5.18
16	8048.78	1.83
17	9512.20	1.83
18	7317.07	1.52
19	4390.24	1.52
20	11707.32	1.52
21	8048.78	1.83

The method utilized to improve the system was to place parallel pipes between certain nodes. There are 15 available diameters (**Table 8.3**), which could be used as well as a so-called "do nothing" option for each of the 21 pipes in the system.

Table 8.2: Node data of the existing system/network

Node	Minimum required pressure level (m)	Demand (m^3/s)
1	91.44	Reservoir
2	77.72	2.62
3	77.72	2.62
4	77.72	2.50
5	77.72	2.50
6	77.72	2.50
7	77.72	2.50
8	77.72	2.50
9	77.72	4.81
10	77.72	0.03
11	77.72	4.81
12	77.72	3.32
13	77.72	3.32
14	77.72	2.62
15	77.72	2.62
16	79.25	4.81
17	83.15	1.63
18	77.72	3.32
19	77.72	3.32
20	77.72	4.81

The ground elevation is 0.0 m.

Table 8.3: Available pipes from which the pipe selection will be made

Pipe no	Diameter (mm)	Unit pipe cost ($/m)	Reference diameter (inches)
1	0.0	0.00	0"
2	914.4	306.75	36"
3	1219.2	439.63	48"
4	1524.0	577.43	60"
5	1828.8	725.07	72"
6	2133.6	875.98	84"
7	2438.4	1036.75	96"
8	2743.2	1197.51	108"
9	3048.0	1368.11	120"
10	3352.8	1538.71	132"
11	3657.6	1712.60	144"
12	3962.4	1893.04	156"
13	4267.2	2073.49	168"
14	4572.0	2260.50	180"
15	4876.8	2447.51	192"
16	5181.6	2637.80	204"

A Hazen-Williams friction factor of C_H = 100 is assumed for both the old tunnels and the new pipes.

The only system constraint that this network has is that the pressure at each node should be equal to the values as specified in **Table 8.2**.

It has been indicated that the conversion from imperial units to metric units could result in small differences and subsequent changes in optimum solutions. For comparison purposes the system was analysed using both unit systems.

Although the system is fairly unsophisticated and small in comparison to the internal distribution system there are 1,93 x E25 or 16^{21} possible solutions for this system. It is thus impossible to analyse every single network improvement alternative.

8.3.1.2 Previous results obtained

Previous researchers have investigated the system extensively and obtained the following solutions that met the defined fitness function of minimum cost and required pressure at every node. Due to different interpretations of the Hazen-Williams equation different researchers obtained different solutions (Savic and Walters, 1997).

Table 8.4: Previous results

Pipe	Schaake and Lai (1969)	Quindry et al. (1981)	Gessler (1982)	Morgan and Goulter (1985)	Fujiwara and Khang (1990)	Savic and Walters (1997)	Evolver (2000)
	Diameter (inches)	Diameter (inches)	Diameter (inches)	Diameter (inches)	Diameter (inches)	Diameter (inches)	Diameter (inches)
1	52,02	0,00	0,00	0,00	0,00	0,00	168,00
2	49,90	0,00	0,00	0,00	0,00	0,00	0,00
3	63,41	0,00	0,00	0,00	0,00	0,00	0,00
4	55,59	0,00	0,00	0,00	0,00	0,00	0,00
5	57,25	0,00	0,00	0,00	0,00	0,00	0,00
6	59,19	0,00	0,00	0,00	0,00	0,00	0,00
7	59,06	0,00	100,00	144,00	73,62	108,00	96,00
8	54,95	0,00	100,00	0,00	0,00	0,00	0,00
9	0,00	0,00	0,00	0,00	0,00	0,00	0,00
10	0,00	0,00	0,00	0,00	0,00	0,00	0,00
11	116,21	119,02	0,00	0,00	0,00	0,00	0,00
12	125,25	134,39	0,00	0,00	0,00	0,00	36,00
13	126,87	132,49	0,00	0,00	0,00	0,00	0,00
14	133,07	132,87	0,00	0,00	0,00	0,00	0,00
15	126,52	131,37	0,00	0,00	0,00	0,00	0,00
16	19,52	19,26	100,00	96,00	99,01	96,00	72,00
17	91,83	91,71	100,00	96,00	98,75	96,00	96,00
18	72,76	72,76	80,00	84,00	78,97	84,00	84,00
19	72,61	72,64	60,00	60,00	83,82	72,00	72,00
20	0,00	0,00	0,00	0,00	0,00	0,00	0,00
21	57,82	54,97	80,00	84,00	66,59	72,00	72,00
Cost ($M)	**78,09**	**63,58**	**41,80**	**39,20**	**36,10**	**37,13**	**42,63**

Some of the researchers did not use discreet values for the pipe diameters and for this reason direct comparison of the solutions are not always possible. All of the solutions in **Table 8.4** met the requirement of minimum pressure at every node, based on the various interpretations of the Hazen-Williams equation, except Fujiwara and Khang (1990).

8.3.1.3 GANEO solution

The first step was to set-up the system in EPANET to be able to perform a hydraulic analysis. Once the system was hydraulically analysed the problem nodes could be evaluated.

The next step was to proceed to optimise the system utilizing the GANEO program. This is explained by means of a step-by-step tutorial below.

The tutorial provides an introduction on using GANEO to optimize the network model setup in EPANET. The topics to covered include:

- Creating a new project and opening an EPANET model
- Creating a pipe selection file
- Setting the Boundaries and GA Parameters
- Running the optimisation analysis procedure
- Evaluating the alternative options and exporting the results

In this example the simple distribution network shown in **Figure 8.12** (New York Tunnel system) will be optimized. It consists of a source reservoir from which water is gravitated into a loop pipe network. If the system was correctly set-up in EPANET then proceed using the created network file for example "New York Tunnels.net".

Step 1: Return to EPANET and create the input file required for GANEO

1. If EPANET is not already running then launch it from the Windows Start menu.
2. Select *File* | *Open* to open the project for example "New York Tunnels.net".
3. Change the elevations of two nodes in the system, Node 16 and 17. This can be done by selecting Node 16 into the Property Editor. Double click on the node, or select it from the Browser and click the Browser's *Edit* button. Change the elevation of Node 16 to 1.53 m and the elevation of Node 17 to 5.43 m. The reason for this is described in *Step 4* later in this tutorial.
4. Save the network by selecting from the *File* menu the *Save* option.
5. Select *File* | *Export* | *Network…* to export the project into a format, which will be readable in GANEO i.e. an input file with extension *.inp.
6. Once the input file has been created, *.inp file (for example "New York Tunnels (SI).inp") exit EPANET and launch the GA Water Utility Programs from the Windows Start menu. Select the Genetic Algorithm Network Optimization (GANEO) option.

Step 2: Creating a new project and importing an EPANET model

1. The first task is to create a new project in GANEO. Select *File | New project* to create a new project.
2. Click on the *Load file* button (see **Figure 8.13** below) to load the created EPANET input file (for example "New York Tunnels.inp). The created file in can also be viewed in its text format by clicking on the *View file* button.

EPANET Network

Load file | View file | View network

Filename: New York Tunnels (SI).inp

Pipe ID	Start node	End node	Length	Diameter	Roughness	Fixed
1	1	2	3536.59	4572	100	No
2	2	3	6036.59	4572	100	No
3	3	4	2225.61	4572	100	No
4	4	5	2530.49	4572	100	No
5	5	6	2621.95	4572	100	No
6	6	7	5823.17	4572	100	No
7	7	8	2926.83	3352.8	100	No
8	8	9	3810.98	3352.8	100	No
9	9	10	2926.83	4572	100	No

Figure 8.13: Loading EPANET network file

The opened data file should be similar to that shown in **Figure 8.C** above.

In this water distribution network the aim was to improve the pressures by adding in parallel pipes. The program allows the user to make this distinction, see **Figure 8.14** below. Select the *Improvement (Additional pipes)* option.

Network Description

○ New network ◉ Improvement (Additional pipes) ○ Improvement (Replacing pipes)

Figure 8.14: Network description

Step 3: Creating a pipe selection file

1. During the analysis of the New York Tunnels system it was decided that the method utilized to improve the system was to place parallel pipes between certain nodes. There are 15 available diameters (see **Table 8.3**), which could be used as well as a so-called "do nothing" option for each of the 21 pipes in the system. To create this pipe data file click on the *Create File* button as shown in **Figure 8.15**.

Pipe data
Filename: New York Tunnels pipe data set (SI).pip

Create file | Load file

Pipe name	Class	Nominal diameter	Inside diameter	Roughness	Cost
Steel	10	0	0	100	0.00
Steel 36 in	10	914.4	914.4	100	306.75
Steel 48 in	10	1219.2	1219.2	100	439.63
Steel 60 in	10	1524	1524	100	577.43

Figure 8.15: Created Pipe Data input file

2. The pipe data must be entered in the same units as the loaded EPANET input file. In other words the diameters are entered in millimetres (mm) for this example and the roughness values are the Hazen Williams friction factor ($C_H = 100$).
3. If a data file was previously created it can be directly loaded into the project by clicking on the *Load file* button.

Step 4: Setting the Boundaries and GA Parameters

Once the EPANET network was loaded and the pipe data file from which the additional pipes will be selected is incorporated in the project, the boundaries and GA parameters.

Boundaries

The GANEO program allows the user to enter the minimum pressure allowed in the systems as well as the maximum velocity.

1. The boundary criteria that were set for this network is that the minimum pressure in the system should be as indicated in **Table 8.2**. As can be seen the minimum pressure varies for different nodes. To accommodate this in the program the ground elevations had to be increased for nodes 16 and 17. Since the nodes required a higher pressure than the rest of the nodes, increasing the ground elevation will result in a solution, which will meet the minimum pressure requirement at every node of 77.72 m. The elevations of nodes 16 and 17 were increased from 0 m to 1.53 m and 5.43 m respectively. In other words if the minimum pressure of 77.72 m is obtained at node 17 the real pressure is in fact 83.15 m.
2. Enter the *Minimum required residual pressure*, 77.72 m and select the correct units from the drop down list.
3. The maximum velocity is set to 2.5 m/s. This criteria usually assists in obtaining "beter" size diameters. If the velocity is higher than 2.5 m/s the program will start to penalise the solution.
4. Enter the *Maximum velocity*, 2.5 m/s and select the correct units from the drop down list.

GA Parameters

Enter/select the following GA Parameters:

1. Cross over method: Select *Single point*
2. Mutation method: Select *Random*
3. Mutation percentage: Scroll to *2.00%*
4. Termination: Select *Maximum generations*
5. No of iterations: Enter *600*
6. Number of favourable solutions: Scroll to *20*
7. Population size: Scroll to *100*
8. User comment: Enter *Example project*
9. Pressure penalty: Enter *10*
9. Velocity penalty: Enter *2*
10. Acceptable range of Comparative vs Real Cost: Set to *0*
11. Initial seed number: Enter *123*

All the required data has now been entered. Click on the *Save* button below the menus or under the *File* menu click on *Save As*. In the Save As dialog that appears, select a folder and file name under which to save this project for example "New York Tunnels.gan". Once the project is saved the optimisation analysis procedure can be performed.

The entered data should look similar to screen capture shown in **Figure 8.16.**

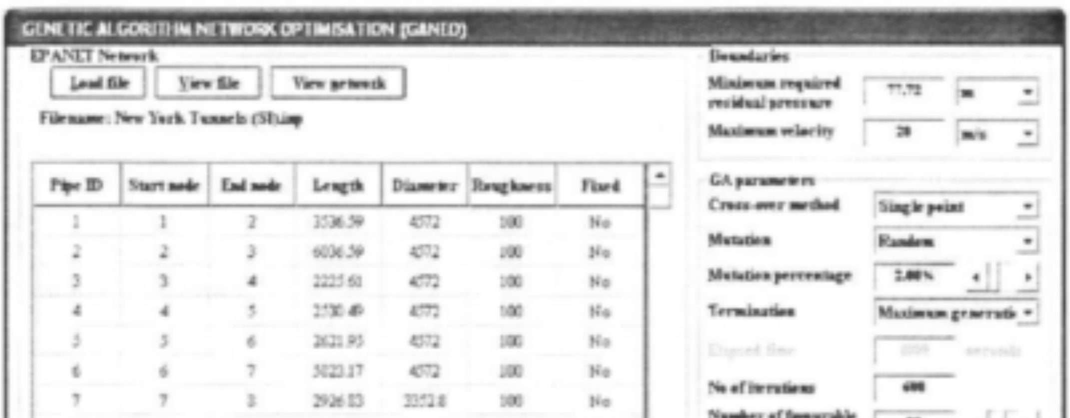
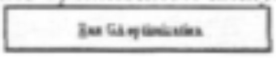
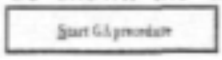


Figure 8.16: New York Tunnels input data screen

Step 5: Running the optimisation analysis procedure

Click on the  button on the Input Screen (see **Figure 8.16**). You will be shown the Optimisation Screen as shown in **Figure 8.17**. Click on the  button to start the procedure.

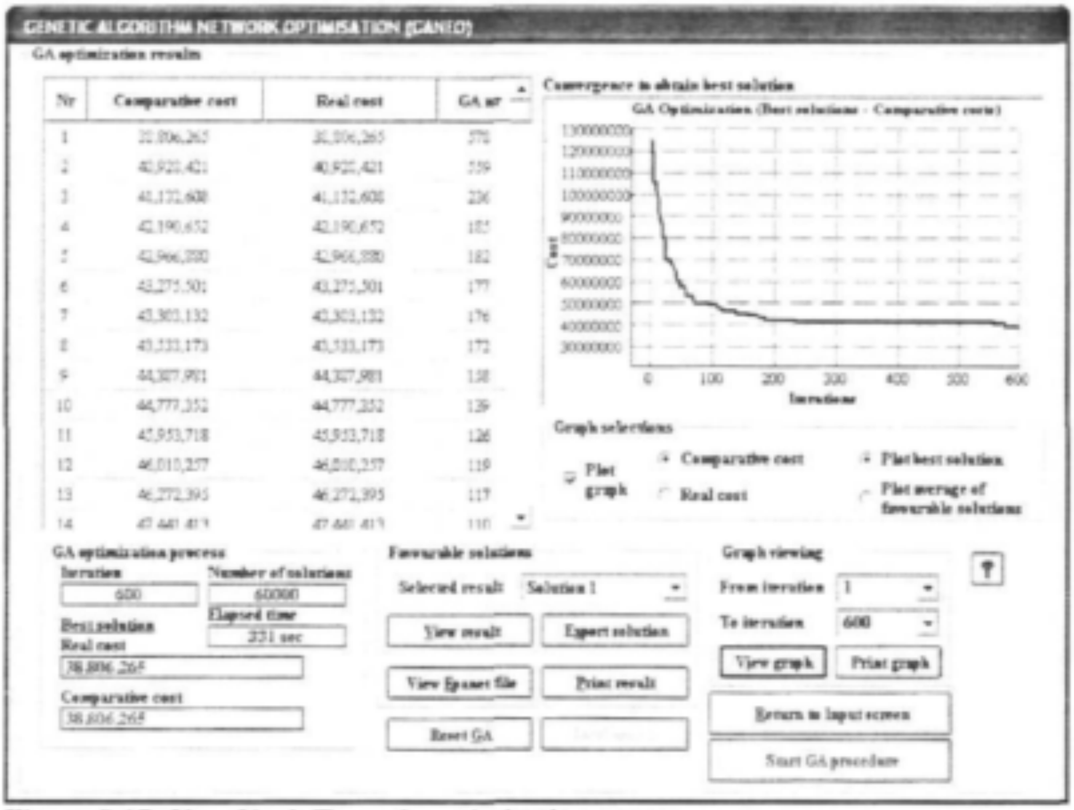


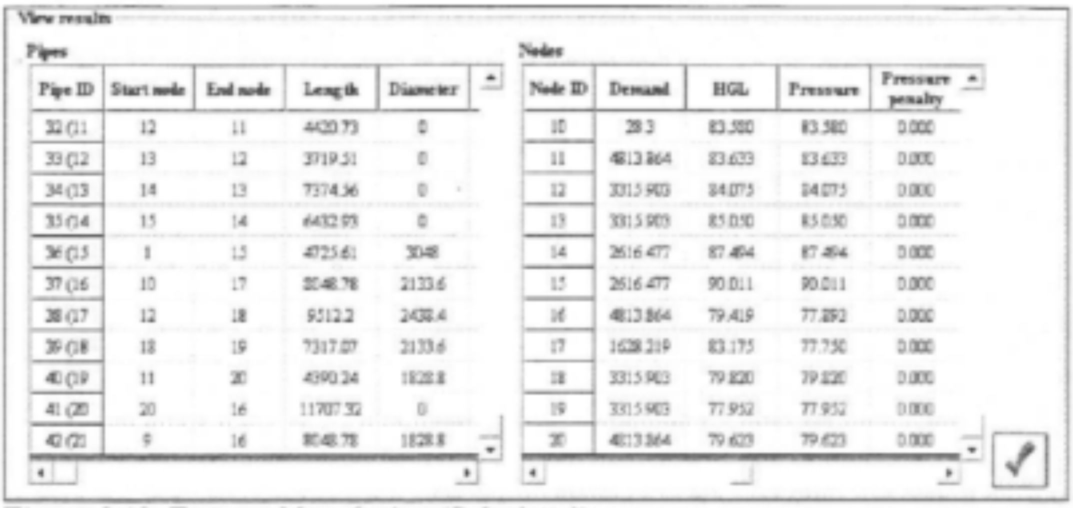
Figure 8.17: New York Tunnels optimisation screen

The program will then run the Genetic Algorithm developed for optimising a network and show the progression thereof on the convergence graph.

In the left bottom corner the iteration number, number of solutions, elapsed time and real and comparative costs are shown. The table on the left shows the 20 best solutions.

Step 6: Evaluating the alternative options and exporting the results

After the analysis is complete the user can select any of the 20 best solutions and view the favourable result as shown in **Figure 8.18**, by using the drop down list of best solutions. The user also has the option to export the selected solution, thus creating an EPANET input file of the proposed solution.



View results

Pipes

Pipe ID	Start node	End node	Length	Diameter
32 (11	12	11	4420.73	0
33 (12	13	12	3719.51	0
34 (13	14	13	7374.36	0
35 (14	15	14	6432.93	0
36 (15	1	15	4725.61	3048
37 (16	10	17	8048.78	2133.6
38 (17	12	18	9512.2	2438.4
39 (18	18	19	7317.07	2133.6
40 (19	11	20	4390.24	1828.8
41 (20	20	16	11707.32	0
42 (21	9	16	8048.78	1828.8

Nodes

Node ID	Demand	HGL	Pressure	Pressure penalty
10	28.3	83.580	83.580	0.000
11	4813.864	83.633	83.633	0.000
12	3315.903	84.075	84.075	0.000
13	3315.903	85.050	85.050	0.000
14	2616.477	87.494	87.494	0.000
15	2616.477	90.011	90.011	0.000
16	4813.864	79.419	77.893	0.000
17	1628.219	83.175	77.750	0.000
18	3315.903	79.820	79.820	0.000
19	3315.903	77.952	77.952	0.000
20	4813.864	79.623	79.623	0.000

Figure 8.18: Favourable solution (Solution 1)

8.3.1.4 Conclusion

The optimum obtained with the GANEO program was $M38,806. The number of iterations (generations) it took to obtain this optimum was 578. The optimization process took 331 seconds (600 iterations) utilizing a 1.8 Ghz Celeron computer with 256 MB ram.

The convergence to the optimum is shown in **Figure 8.19** below. As can be seen the initial reduction in total cost of the system occurs fast after which this improvement process slows down.

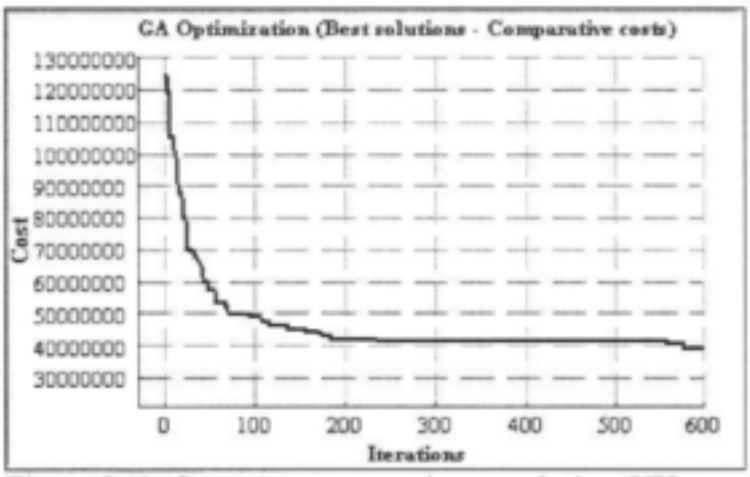


Figure 8.19: Convergence to optimum solution (NY tunnel system)

The nodes with its pressure closets to the minimum required are node 16 and 17 with pressure elevations of 77.893 m and 77,750 m respectively. The 5 best solutions, based on different initial seed values, and their costs are shown in **Table 8.5**.

A comparison of the optimum obtained with the GANEO program and that obtained by other researchers is shown in **Table 8.6**. As can be seen the GANEO program provides promising results.

Table 8.5: The 5 best GANEO solutions (using different seed values)

Pipe	Solution for New York Tunnel system optimisation using GANEO				
	Solution 1	Solution 2*	Solution 3*	Solution 4*	Solution 5*
1	0,0	0,0	2438,4	0,0	0,0
2	0,0	0,0	0,0	0,0	0,0
3	0,0	0,0	0,0	0,0	0,0
4	0,0	0,0	0,0	0,0	0,0
5	0,0	0,0	0,0	0,0	0,0
6	0,0	0,0	0,0	0,0	0,0
7	0,0	3657,6	3048,0	0,0	3048,0
8	0,0	0,0	0,0	0,0	0,0
9	0,0	0,0	0,0	0,0	0,0
10	0,0	0,0	0,0	0,0	0,0
11	0,0	0,0	0,0	0,0	0,0
12	0,0	0,0	0,0	0,0	1524,0
13	0,0	0,0	0,0	0,0	0,0
14	0,0	0,0	0,0	0,0	0,0
15	3048,0	0,0	0,0	3657,6	0,0
16	2133,6	2438,4	2133,6	1828,8	2438,4
17	2438,4	2743,2	2438,4	2438,4	2743,2
18	2133,6	1828,8	2133,6	2133,6	1828,8
19	1828,8	1828,8	1828,8	2743,2	1828,8
20	0,0	0,0	0,0	0,0	0,0
21	1828,8	1828,8	1828,8	1524,0	1828,8
Cost ($M)	**38,806**	**39,073**	**40,012**	**40,105**	**40,212**
Initial seed	**123**	**1**	**12**	**20**	**1**

Note: Diameters are in millimetres,
* Solutions 2 to 5 - Pressure penalty set to 3 and number of iterations set to 2000

Table 8.6: Comparison with previous researchers results

Pipe	Schaake and Lai (1969)	Quindry *et al.* (1981)	Gessler (1982)	Morgan and Goulter (1985)	Fujiwara and Khang (1990)	Savic and Walters (1997)	Evolver (2000)	GANEO (2004)
	Diameter (inches)							
1	52,02	0,00	0,00	0,00	0,00	0,00	168,00	0,00
2	49,90	0,00	0,00	0,00	0,00	0,00	0,00	0,00
3	63,41	0,00	0,00	0,00	0,00	0,00	0,00	0,00
4	55,59	0,00	0,00	0,00	0,00	0,00	0,00	0,00
5	57,25	0,00	0,00	0,00	0,00	0,00	0,00	0,00
6	59,19	0,00	0,00	0,00	0,00	0,00	0,00	0,00
7	59,06	0,00	100,00	144,00	73,62	108,00	96,00	0,00
8	54,95	0,00	100,00	0,00	0,00	0,00	0,00	0,00
9	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
10	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
11	116,21	119,02	0,00	0,00	0,00	0,00	0,00	0,00
12	125,25	134,39	0,00	0,00	0,00	0,00	36,00	0,00
13	126,87	132,49	0,00	0,00	0,00	0,00	0,00	0,00
14	133,07	132,87	0,00	0,00	0,00	0,00	0,00	0,00
15	126,52	131,37	0,00	0,00	0,00	0,00	0,00	120,00
16	19,52	19,26	100,00	96,00	99,01	96,00	72,00	84,00
17	91,83	91,71	100,00	96,00	98,75	96,00	96,00	96,00
18	72,76	72,76	80,00	84,00	78,97	84,00	84,00	84,00
19	72,61	72,64	60,00	60,00	83,82	72,00	72,00	72,00
20	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
21	57,82	54,97	80,00	84,00	66,59	72,00	72,00	72,00
Cost ($M)	**78,09**	**63,58**	**41,80**	**39,20**	**36,10**	**37,13**	**42,63**	**38,81**

Note: The pipe diameters in millimetres as obtained in GANEO were converted to inches

8.3.2 GANEO – Example 2: Hanoi system

8.3.2.1 Description of system

The Hanoi system is a new network that should be optimized. It has a single fixed head source at elevation of 100 m. There are 34 pipes and 32 nodes as shown in **Figure 8.20** and listed in **Tables 8.7** and **8.8**.

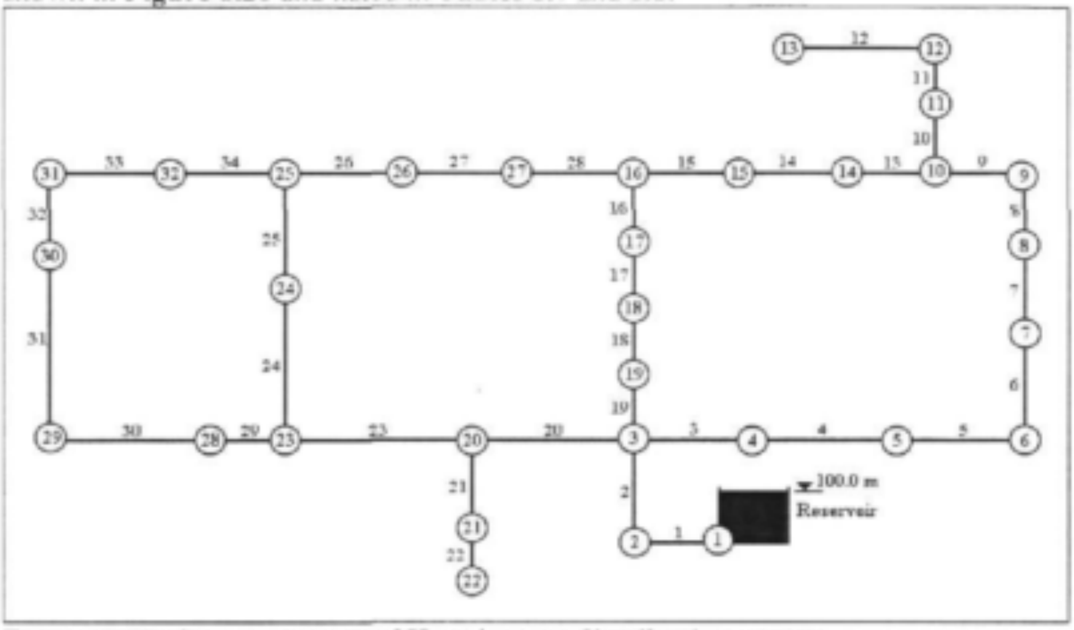


Figure 8.20: General layout of Hanoi water distribution system

Table 8.7: Pipe lengths of the existing system/network

Pipe	Length (m)	Pipe	Length (m)
1	100	18	800
2	1350	19	400
3	900	20	2200
4	1150	21	1500
5	1450	22	500
6	450	23	2650
7	850	24	1230
8	850	25	1300
9	800	26	850
10	950	27	300
11	1200	28	750
12	3500	29	1500
13	800	30	2000
14	500	31	1600
15	550	32	150
16	2730	33	860
17	1750	34	950

Table 8.8: Node data of the existing system/network

Node	Demand (m^3/h)	Node	Demand (m^3/h)
2	890	18	1345
3	850	19	60
4	130	20	1275
5	725	21	930
6	1005	22	485
7	1350	23	1045
8	550	24	820
9	525	25	170
10	525	26	900
11	500	27	370
12	560	28	290
13	940	29	360
14	615	30	360
15	280	31	105
16	310	32	805
17	865		

The ground elevation is 0,0 m.

The pipes that could be utilized in the design of the system are shown in **Table 8.9**.

Table 8.9: Available pipes from which the pipe selection will be made

Nr	Diameter (inches)	Diameter (mm)	Cost/meter length
1	12	304	45.73
2	16	406	70.40
3	20	508	98.39
4	24	609	129.33
5	30	762	180.75
6	40	1 016	278.28

A Hazen-Williams friction factor of $C_H = 130$ is assumed for all new pipes

The head constraint for this system is 30 m i.e. the pressure everywhere in the system should be greater than 30 m. The costs as indicated in **Table 8.9** were calculated based on the following relationship.
Cost calculation: $1.1 \times D^{1.5} \times L$

This network was optimized by creating a network with pipes of size 0,0001 mm. It is however not a requirement to use 0,0001 mm pipes since GANEO will populate all the pipes of the network with the possible options as shown in **Table 8.9**.

8.3.2.2 Previous results obtained

The Hanoi network (a water distribution network in Vietnam) has been optimized by other researchers as shown in **Table 8.10**. Fujiwara and Khang (1990) did not use discreet values for the pipe diameters and for this reason direct comparison of the solutions are not always possible.

Table 8.10: Comparison with previous researchers results

Pipe	Fujiwara and Khang (1990)	Savic and Walters (1997)	Savic and Walters (1997)
	Diameters (inches)		
1	40	40	40
2	40	40	40
3	38.8	40	40
4	38.7	40	40
5	37.8	40	40
6	36.3	40	40
7	33.8	40	40
8	32.8	40	40
9	31.5	40	30
10	25	30	30
11	23	24	30
12	20.2	24	24
13	19	20	16
14	14.5	16	16
15	12	12	12
16	19.9	12	16
17	23.1	16	20
18	26.6	20	24
19	26.8	20	24
20	35.2	40	40
21	16.4	20	20
22	12	12	12
23	29.5	40	40
24	19.3	30	30
25	16.4	30	30
26	12	20	20
27	20	12	12
28	22	12	12
29	18.9	16	16
30	17.1	16	16
31	14.6	12	12
32	12	12	12
33	12	16	16
34	19.5	20	20
Cost ($M)	**5,354**	**6,073**	**6,187**

8.3.2.3 GANEO solution

Similar to Example 1, EPANET is used to set-up the network for the hydraulic analysis of the water distribution system. Once the system is hydraulically analysed to ensure the proper working thereof the tutorial as detailed in Example 1 is used to optimize the system.

Step 1: Return to EPANET and create the input file required for GANEO

1. Set-up the network and export the network into a format, which will be readable in GANEO by selecting *File | Export | Network...* to for example "Hanoi.inp".
2. Exit EPANET and launch the GA Water Utility Programs from the Windows Start menu. Select the Genetic Algorithm Network Optimization (GANEO) option.

Step 2: Creating a new project and importing an EPANET model

1. Create a new project in GANEO and load the created EPANET input file i.e. "Hanoi.inp".
2. In this water distribution network the aim was to design the network by determining the optimum diameters of all the pipes in the system. The program allows the user to make this distinction, see **Figure 8.21** below. Select the *New network* option.

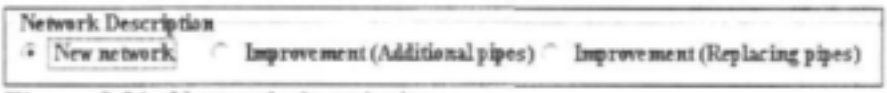


Figure 8.21: Network description

Step 3: Creating a pipe selection file

1. For the analysis of the Hanoi water distribution system the available pipes are as shown in **Table 8.9**. There are 6 available diameters, which could be used for each of the 34 pipes in the system.
2. The pipe data must be entered in the same units as the loaded EPANET input file. In other words the diameters are entered in millimetres (mm) for this example and the roughness values are the Hazen Williams friction factor (C_H = 130).

Step 4: Setting the Boundaries and GA Parameters

The boundaries and GA parameters used for this network is shown below:

Boundaries
The GANEO program allows the user to enter the minimum pressure allowed in the system as well as the maximum velocity.

1. The boundary criteria in terms of pressure that this system should adhere to is 30 m. Enter the minimum required residual pressure of 30.0 *m* and select the correct units from the drop down list.
2. The maximum velocity is set to 10 m/s since no criteria has been previously used for this boundary by other researchers. Enter the maximum velocity of *10 m/s* and select the correct units from the drop down list.

GA Parameters

Enter/select the following GA Parameters:

1. Cross over method: Select *Single point*
2. Mutation method: Select *Random*
3. Mutation percentage: Scroll to *3.00%*
4. Termination: Select *Maximum generations*
5. No of iterations: Enter *2000*
6. Number of favourable solutions: Scroll to *20*
7. Population size: Scroll to *100*
8. User comment: Enter *An example of a new network*
9. Pressure penalty: Enter *3*
10. Velocity penalty: Enter *2*
11. Acceptable range of Comparative vs Real Cost: *Set to 0*
12. Initial seed number: *3*

All the required data has now been entered. Click on the *Save* button below the menus or under the *File* menu click on *Save As*. Once the project is saved the optimisation analysis procedure can be performed.

The entered data should look similar to the screen capture as shown in **Figure 8.22**.

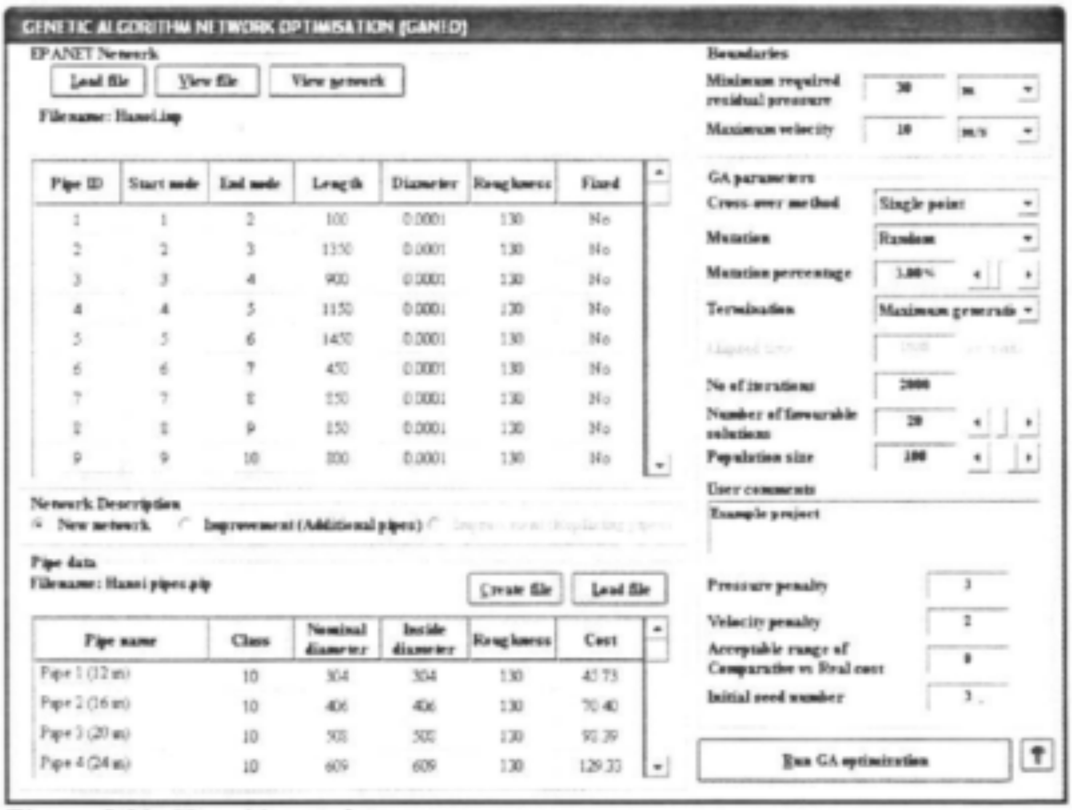


Figure 8.22: Hanoi input data screen

Step 5: Running the optimisation analysis procedure

Click on the Run GA optimisation button on the Input Screen (see **Figure 8.22**). You will be shown the Optimisation Screen as shown in **Figure 8.23**. Click on the Start GA procedure button to start the procedure.

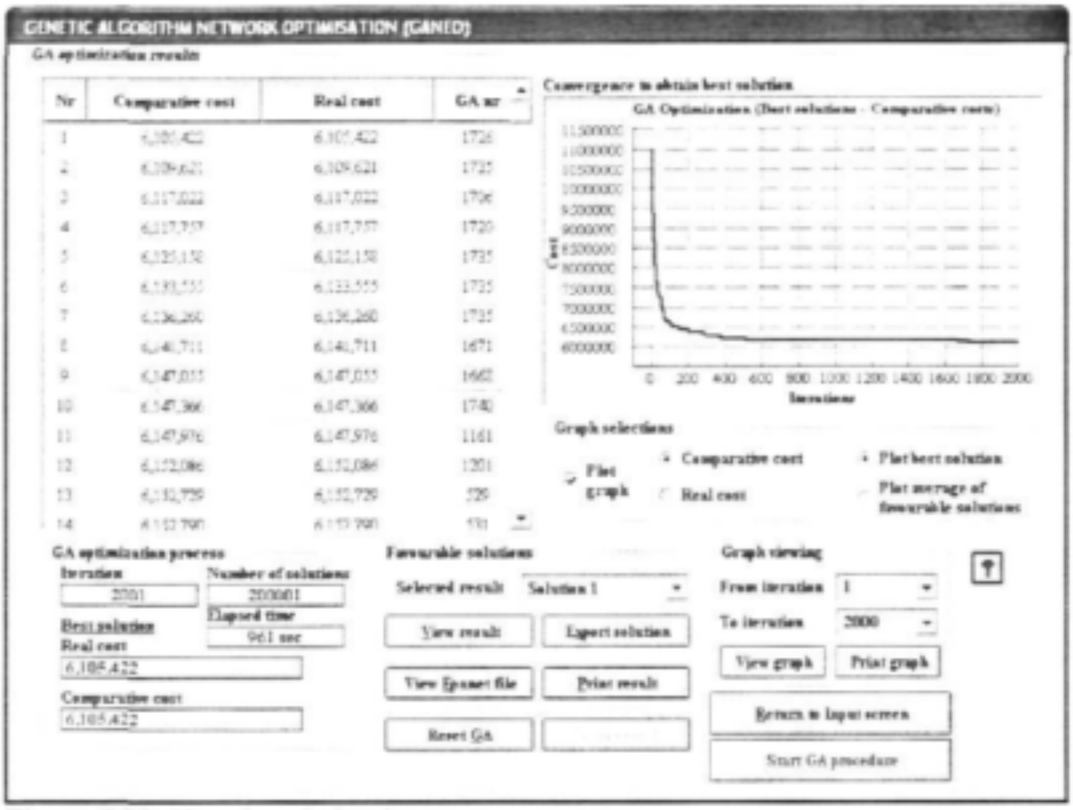


Figure 8.23: Hanoi optimisation screen

The program will then run the Genetic Algorithm developed for optimising a new network and show the progression thereof on the convergence graph.

Step 6: Evaluating the alternative options and exporting the results

After the analysis is complete the user can select any of the 20 best solutions and view the favourable results.

8.3.2.4 Conclusion

The optimum obtained with the GANEO program was $M6.105. The number of iterations (generations) it took to obtain this optimum was 1726. The optimization process took 961 seconds (for 2000 iterations) utilizing a 1.8 Ghz Celeron computer with 256 MB ram.

As can be seen the initial reduction in total cost of the system occurs rapidly after which the improvement process slows down.

The pressures that were obtained with this optimum solution obtained are shown in **Table 8.11**.

Table 8.11: Pressures of the Hanoi system

Node	Pressure (m)	Node	Pressure (m)
2	97.141	18	45.030
3	61.670	19	58.712
4	57.046	20	50.591
5	51.317	21	41.242
6	45.293	22	36.010
7	43.887	23	44.487
8	42.227	24	38.864
9	40.909	25	35.252
10	39.945	26	31.570
11	38.386	27	30.563
12	34.941	28	38.884
13	30.713	29	30.005
14	36.911	30	30.299
15	32.031	31	30.586
16	31.045	32	33.058
17	38.241		

As can be seen in **Table 8.11** the solution meets the minimum pressure criteria of 30 m.

The 3 best solutions obtained with the GANEO program and their costs are shown in **Table 8.12**, which is also compared with that of the previous researchers. As can be seen the GANEO program provided very good results.

Due to the interpretation of the Hazen-Williams C-value, different researchers obtained different solutions. This is discussed in detail in Savic and Walters (1997).

Table 8.12: The 3 best GANEO solutions compared with previous researchers' results

Pipe	Solutions for Hanoi water distribution system					
	Fujiwara and Khang (1990)	Savic and Walters (1997)	Savic and Walters (1997)	GANEO Solution 1	GANEO Solution 2	GANEO Solution 3
1	40	40	40	40	40	40
2	40	40	40	40	40	40
3	38.8	40	40	40	40	40
4	38.7	40	40	40	40	40
5	37.8	40	40	40	40	40
6	36.3	40	40	40	40	40
7	33.8	40	40	40	40	40
8	32.8	40	40	40	40	40
9	31.5	40	30	40	40	40
10	25	30	30	30	30	30
11	23	24	30	24	24	24
12	20.2	24	24	24	24	24
13	19	20	16	20	20	20
14	14.5	16	16	12	12	12
15	12	12	12	12	12	12
16	19.9	12	16	12	12	12
17	23.1	16	20	20	20	20
18	26.6	20	24	20	20	20
19	26.8	20	24	24	24	24
20	35.2	40	40	40	40	40
21	16.4	20	20	20	20	20
22	12	12	12	12	12	12
23	29.5	40	40	40	40	40
24	19.3	30	30	30	30	30
25	16.4	30	30	30	30	30
26	12	20	20	20	20	20
27	20	12	12	12	12	16
28	22	12	12	12	12	12
29	18.9	16	16	16	16	16
30	17.1	16	16	12	12	12
31	14.6	12	12	12	12	12
32	12	12	12	16	20	20
33	12	16	16	16	16	16
34	19.5	20	20	24	24	24
Cost ($M)	5.354	6.073	6.187	6.105	6.109	6.117
Initial seed				3	3	3

Note: The pipe diameters in millimetres as obtained in GANEO were converted to inches

8.4 Conclusion

The EPANET software used for the hydraulic analysis of the systems is well-accepted, well-tested analysis software used in various other hydraulic analysis packages. The GAPOP and the newly developed GANEO program provide good results as shown in the sections above. These programs are useful tools to assist water engineers in the design of single pipeline systems such as bulk main transmission lines and water distributions systems. The GANEO program can design a new network as well as provide suggestion on how to improve an existing network. The programs will assist in the further transfer of knowledge on the use of genetic algorithms in the water field.

9. CONCLUSIONS

Based on the conceptual developments as described in this report and the development of the first version of an interface with EPANET, it can be concluded that the GA optimization technique has great benefits and can be successfully applied to water distribution systems and water resource assessment.

During Phase 2 of the study, the GAPOP program, which was developed during Phase 1, was upgraded, to handle more pipes and utilizing the powerful EPANET hydraulic engine. The program determines the best pipe selection based on the life Cycle Cost. (LCC).

The conceptual development of GA procedures for water distribution systems, was incorporated in the software program called GANEO and yielded extremely promising results. GANEO is capable of optimising complex in a short time.

These programs (GAPOP and GANEO) were grouped together and are called the Genetic Algorithm Water Utility Programs. Further developments and features in the GA field will be added to this software model in the next Phase of the study. During a recent Pipeline Design Course, a practical computer session attended by more than 100 people from industry utilized the GA Water Utility Programs software package.

The request for the program and feedback highlighted the need for such an optimisation tools.

These programs can be downloaded from either the WRC or the University of Pretoria web sites (links below).

WRC - http://www.wrc.org.za
University of Pretoria - http://www.up.ac.za/academic/civil/divisions/water/ga.html

10. RECOMMENDATIONS

This study concentrated on the potential applications of GAs in the following fields:

- Water resources assessment
- Network analyses and scheduling

It is recommended that the following aspects should be researched/developed in future studies:

- Conceptual development of a genetic algorithm multi-objective optimization model for the optimization of water distribution systems specifically for rural developments, combining design and operational objective functions, including Life Cycle Costing analysis of networks.
- Further development and improving of the GA Water Utility Programs software model
- Testing of parameter values and alternative operators for the GA (mutation percentage, population size, crossover methods etc.).
- Utilization of the developed GA Water Utility Programs in the design of a new water distribution system in South Africa to illustrate/demonstrate its applicability and benefits.
- Develop a generic inter reservoir optimisation model.

11. REFERENCES

Anderson, C., Bonabeau, E. & Scott, J.S. (2004) "Evolutionary testing as both a testing and redesign tool: a study of a shipboard firemain's valve and pump controls." *Proceedings of the Congress on Evolutionary Computation* (CEC2004), Portland, OR, June 20-23, 2004.

Angel, P. L., Hernandez, J. A., and Agudelo, R. (1999). "Fuzzy expert system model for the operation of an urban water supply system." *Computing and Control for the Water Industry*, D. A. Savic and G. A. Walters, eds., Exeter, UK.

Boulos, P. F., Orr, C. H., De Schaetzen, W., Chatila, J. G., Moore, M., Hsiung, P., and Thomas, D. (2001) "Optimal pump operation of water distribution systems using genetic algorithms." *AWWA Distribution System Symposium*, San Diego, CA.

Brdys, M., and Ulanicki, B. (1994). *Operational Control of Water Systems: Structures, Algorithms and Applications*, Prentice Hall.

Burnell, D., Race, J., and Evans, P. (1993). "Overview of the trunk scheduling system for the London Ring Main." *Water Science and Technology*, 28(11-12), 99-109.

Centre for Water Systems (2004). *Benchmark Networks for Design and Optimisation of Water Distribution Networks*. University of Exeter. [Online]. Available: http://www.projects.ex.ac.uk/cws/ [2004, December 15]

Chase, D. V., and Ormsbee, L. E. (1993). "Computer-generated pumping schedules for satisfying operational objectives." *Journal of the American Water Works Association*, 85(7), 54-61.

Davis, L. (1991). *Handbook of genetic algorithms*, Van Nostrand Reinhold, New York.

De Jong, K.A. (1975): "An analysis of the behaviour of a class of genetic adaptive systems." Ph.D. Thesis, University of Michigan, USA.

DWAF. (1987). *Yield analysis – terms and procedures.* Report No. PC000/00/6986 of the South African Department of Water Affairs and Forestry. Compiled by BKS Water Resources Associates as part of the Vaal River System Analysis, Pretoria, SA.

DWAF (1999). *Water Resources Yield Model: User Guide* – Release 4.1.1. Report No. P0000/00/001/98 compiled by BKS (Pty) Ltd for the Department of Water Affairs and Forestry, Pretoria, SA.

DWAF (2005). *Water Resources Yield Model: Procedural Manual.* Internal report of the Department of Water Affairs and Forestry, Pretoria SA, compiled by WRP Consulting Engineers (Pty) Ltd.

Forzini, Y. (1991). "Optimized pumping program." *Civil Engineering Systems*, 8, 207-211.

Goldberg, D.E. (1989). *Genetic Algorithms in search, optimisation and machine learning*. 2nd ed., Addison-Wesley, Reading, Mass.

Goldberg, D. E., and Kuo, C. (1987). "Genetic Algorithms in Pipeline Optimization". *Journal of Computing in Civil Engineering*, 1(2), 128-141.

Goldberg, D. E., and Voessner, S. (1999) "Optimizing global-local search hybrids." *GECCO*, 220-228.

Holland, J.H. (1975). *Adaptation in Natural and Artificial Systems*. University of Michigan Press. 2nd Edition 1992

Hsu, N.-S., Louie, P., and Yeh, W.-G. (1992). "Modeling of a large-scale water distribution system." *Water Resource Planning and Management: Saving a Threatened Resource - In Search of Solutions*, Proceedings of the Water Resources Sessions at Water Forum, ASCE, Baltimore, USA, 598-603.

Jarrige, P.-A., Harding, T., Knight, D., and Howes, D. (1991). "Using optimisation for integrated water network management." *Civil Engineering Systems*, 8, 241-245.

Jaszkiewicz and Ziembinski. (2004). MOMHLib++ Open source Multiple-Objective MetaHeuristics Library in C++ http://www-idss.cs.put.poznan.pl/~jaszkiewicz/MOMHLib/, Accessed 26 May 2004

Jowitt, P., Garrett, R., Cook, S., and Germanopoulos, G. (1988). "Real-time forecasting and control for water distribution." *Computer Applications in Water Supply*, B. Coulbeck and C. Orr, eds., John Wiley & Sons, Letchworth, Enland, 329-355.

Jowitt, P. W., and Germanopoulos, G. (1992). "Optimal pump scheduling in water-supply networks." *Journal of Water Resources Planning and Management*, 118(4), 406-422.

Ko, S. K., Oh, M. H., and Fontane, D. G. (1997). "Multiobjective analysis of service-water-transmission systems." *Journal of Water Resources Planning and Management*, 123(2), 78-83.

Lansey, K. E., and Awumah, K. (1994). "Optimal pump operations considering pump switches." *Journal of Water Resources Planning and Management*, 120(1), 17-35.

Leon, C., Martin, S., Elena, J. M., and Luque, J. (2000). "Explore: hybrid expert system for water networks management." *Journal of Water Resources Planning and Management*, 126(2), 65-74.

Likeman, M. (1993). "Constraint satisfaction methods in water supply scheduling." *Integrated Computer Applications in Water Supply*, B. Coulbeck, ed., Research Studies Press, Taunton, England, 213-225.

Ndiritu, J.G. (2003). "Reservoir system optimisation using penalty approach and multi-population genetic algorithm.", *Water SA*, Vol. 29, 273-280, July 2003.

Nel, D. (1993). *Bepaling van die optimale Stoorkapasiteit van twee Johannesburgse Munisipale Diensreservoirs*. MEng dissertation. Rand Afrikaans University.

Nel, D., and Haarhoff J. (1996) Sizing Municipal Water Storage Tanks with Monte Carlo Simulation. Aqua 45(4) 203-212.

Nitivattananon, V., Sadowski, E. C., and Quimpo, R. G. (1996). "Optimization of water supply system operation." *Journal of Water Resources Planning and Management*, 122(5), 374-384.

Olivera, R. and Loucks, D.P. (1997). "Operating rules for multi-reservoir systems" *Water Resource Research*, Vol 33, Issue 4, pp 839-852

Ormsbee, L. E., and Reddy, S. L. (1995). "Nonlinear heuristic for pump operations." *Journal of Water Resources Planning and Management*, 121(4), 302-309.

Percia, C., Oron, G., and Mehrez, A. (1997). "Optimal operation of regional system with diverse water quality sources." *Journal of Water Resources Planning and Management*, 123(2), 105-115.

Quevedo, J., Cembrano, G., Wells, G., Pérez, R., and Argelaguet, R. (1999). "Criteria for applying optimisation in water distribution networks." *Computing and Control for the Water Industry*, Research Studies Press.

Robertson, A., and Willet, P. (1995). *Use of Genetic Algroithms in Information Retrieval*, Research and Development Department, British Library, London.

Rossman, L. (2000). "EPANET users manual." Environmental Protection Agency, Risk Recuction Engineering Laboratory, Cincinnati.

Savic, D.A. (2002). "Single-objective vs. Multiobjective Optimisation for Integrated Decision Support", iEMSs 2002, Integrated Assessment and Decision Support, Lugano, Switzerland.

Savic, D., Walters, G., and Schwab, M. (1997). "Multiobjective genetic algorithms for pump scheduling in water supply." *Evolutionary Computing Workshop*, Manchester.

Shamir, U. (1985). "Computer applications for real-time operation of water distribution systems." *ASCE Speciality Conference on Computer Applications in Water Resources*, Buffalo, New York, 379-390.

South, M., Wetherill, G. and Tham, M. (1993). "Hitch-hiker's guide to genetic algorithms." *Journal of Applied Statistics*, 20(1).

Tarquin, A. J., and Dowdy, J. (1989). "Optimal pump operation in water distribution." *Journal of Hydraulic Engineering*, 115(2), 158-168.

Ulanicki, B., Coulbeck, B., and Ulanicka, K. (1997). "Optimization techniques can enhance water network analyses." *3rd International conference on water pipeline systems - Leakage management, network optimization, and pipeline rehabilitation technology*, R. Chilton, ed., Mechanical Engineering Publications, London, 73-82.

Van Vuuren S.J., Van Rooyen, P.G., Fouchè, W.L., Haarhoff, J., 2001. WRC Report No. 1144/1/01, Water Research Commission, Pretoria, South Africa.

Van Vuuren, S.J., 2001. Application of genetic algorithms – determination of the optimal pipe diameters. Water SA Vol. 28 No.2 (April 2002), pp. 217-225

Van Zyl, J.E., (2001) *A Methodology for Improved Operational Optimization of Water Distribution Systems*, Ph.D. Thesis, University of Exeter

Van Zyl, J.E., and Haarhoff, J. (2002). Sensitivity of the Reliability of Bulk Water Distribution Systems to Different Design Parameters. *1st Annual Environmental & Water Resources Systems Analysis (EWRSA) Symposium*, ASCE, 19 – 22 May, Roanoke, Virginia.

Van Zyl, J.E., Borthwick, J., Hardy, A. (2003a). OOTEN: An Object-Oriented Programmers Toolkit for Epanet. *Supplementary Proceedings: Advances in Water Supply Management*. Imperial College, London.

Van Zyl, J.E., Savic, D.A., Walters, G.A. (2003b). "Operational Optimization of Water Distribution Systems Using a Hybrid Genetic Algorithm". *Journal of Water Resources Planning and Management*, Accepted for publication.

Wall, M. (2003) GAlib - A C++ Library of Genetic Algorithm Components. http://lancet.mit.edu/ga/. Accessed 26 May 2004

Yu, G., Powell, R. S., and Sterling, M. J. H. (1994). "Optimized pump scheduling in water distribution systems." *Journal of Optimization Theory and Applications*, 83(3), 463-488.

Zebulum, R. S., Pacheco, M. A., Vellasco, M. M. (2002). "*Evolutionary Electronics: Automatic Design of Electronic Circuits and Systems by Genetic Algorithms.*" CRC Press, Boca Raton.

APPENDIX A

STEP-BY-STEP EXAMPLE OF OPTIMIZING A WATER DISTRIBUTION SYSTEM BY HAND UTILIZING A GENETIC ALGORITHM

WATER DISTRIBUTION SYSTEM OPTIMISATION:
STEP-BY-STEP EXAMPLE

1. Introduction

During this research study it again became apparent that the mere mentioning of the words "*genetic algorithms, genes, strings, offspring, binary code, mutation and pairing*" deterred people from exploiting the exciting potential use of GAs.

It is for this reason that a simple water distribution network will be optimized using a genetic algorithm by means of a step-by-step tutorial to illustrate the simple procedures thereof and the meaning of the processes.

2. Problem definition

The water distribution network shown in **Figure A1** has to be optimized:

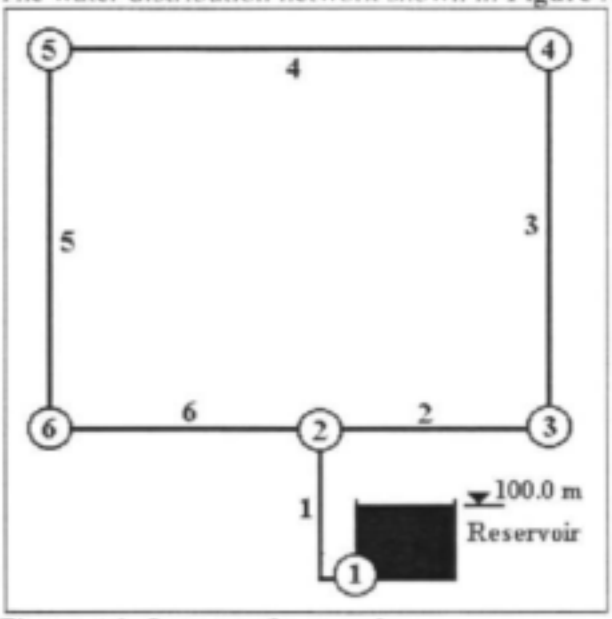


Figure A1: Layout of network

Table A1 and **A2** provide all the relevant information regarding the layout of the system. The available pipe diameters that could be used in the optimization of this system are provided in **Table A3**.

Table A1: Node data

Node	Elevation (m)	Demand (l/s)
1	90	-
2	70	0
3	50	5
4	50	7
5	50	8
6	50	6

Table A2: Pipe data

Pipe	Length (m)
1	500
2	750
3	1200
4	1500
5	1200
6	750

Table A3: Available pipe diameters (uPVC)

Pipe	Diameter (mm)	Pipe cost (R/m)
1	90	32
2	110	48
3	160	90

The absolute roughness of the uPVC pipes (**Table A3**) is 0.03 mm.

In this optimization procedure the GA process will be done by hand calculation and the EPANET software package will be used to perform the hydraulic analyses.

The objective of this system is to obtain a minimum pressure of 24 m at every node.

The cost of a pipe that does not meet the minimum pressure requirement will be calculated using the following equation:

$$PC_{H\,min-i} = C_p^{\,b}\ x\ L_p \qquad \text{and} \qquad b = \frac{P_{required} + \left(P_{required} - P_{calculated}\right)}{P_{required}}$$

Where:

PC_{Hmin-i} = Cost of pipe "i" with added penalty due to minimum pressure requirement not being achieved

C_P = Cost of the pipe per unit length

L_P = Length of the installed pipe

b = Penalty factor (b = 5 if pressure is negative)

$P_{calculated}$ = Calculated pressure at node

$P_{required}$ = Minimum residual pressure required at node

3. Step-by-step GA procedure

The typical GA procedure will be followed:

Step 1: Create the network using EPANET

In EPANET set-up the network consisting of 6 pipes, 5 nodes and the reservoir as shown in **Figure A2**.

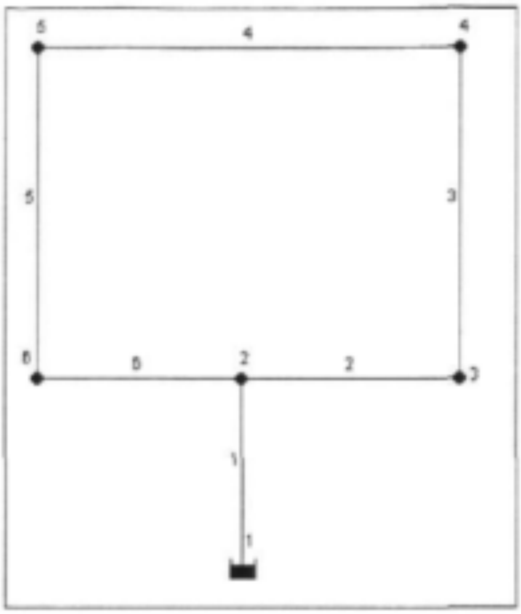


Figure A2: EPANET network of system

Step 2: Create the initial string (seed)

- Make a random selection of pipe diameters from **Table A3** for the pipe network.
- Repeat this four times, which will be the number of populations used in the GA (see **Table A4**).

Table A4: Initial strings (Generation 1)

Pipe	Diameter (mm)			
	String 1	String 2	String 3	String 4
1	90	160	160	160
2	110	110	90	160
3	110	90	160	110
4	160	90	90	110
5	110	90	110	160
6	160	110	110	160

Step 3: Perform hydraulic analysis on each of the population strings

Using EPANET, modify the existing network and analyse to obtain the pressures at each of the nodes (**Table A5**).

Table A5: Calculated pressures of first generation strings

Node	Pressure (m)			
	String 1	String 2	String 3	String 4
1	10.00	10.00	10.00	10.00
2	-51.23	25.40	25.40	25.40
3	-39.76	33.98	20.83	44.13
4	-44.10	14.68	20.03	40.34
5	-44.04	14.42	20.49	40.92
6	-33.80	32.65	29.95	42.70

Step 4: Determine the fitness of each string (solution)

It can clearly be seen from **Table A5** that String 1 produced a solution, which will not meet the pressure criteria. Strings 2 and 3 are significantly better solutions although these do still not meet the minimum pressure requirement of 24 m at nodes 2 to 6. String 4 meets the minimum pressure criteria having pressures greater than 24 m.

When a node does not meet the minimum pressure requirement the pipes supplying that node should be penalised. If nodes have negative pressures such as in the solution of String 1 the nodes should be penalised extensively to highlight the poor results thereof.

The costs of all four strings are calculated (see **Table A6** to **A9**)

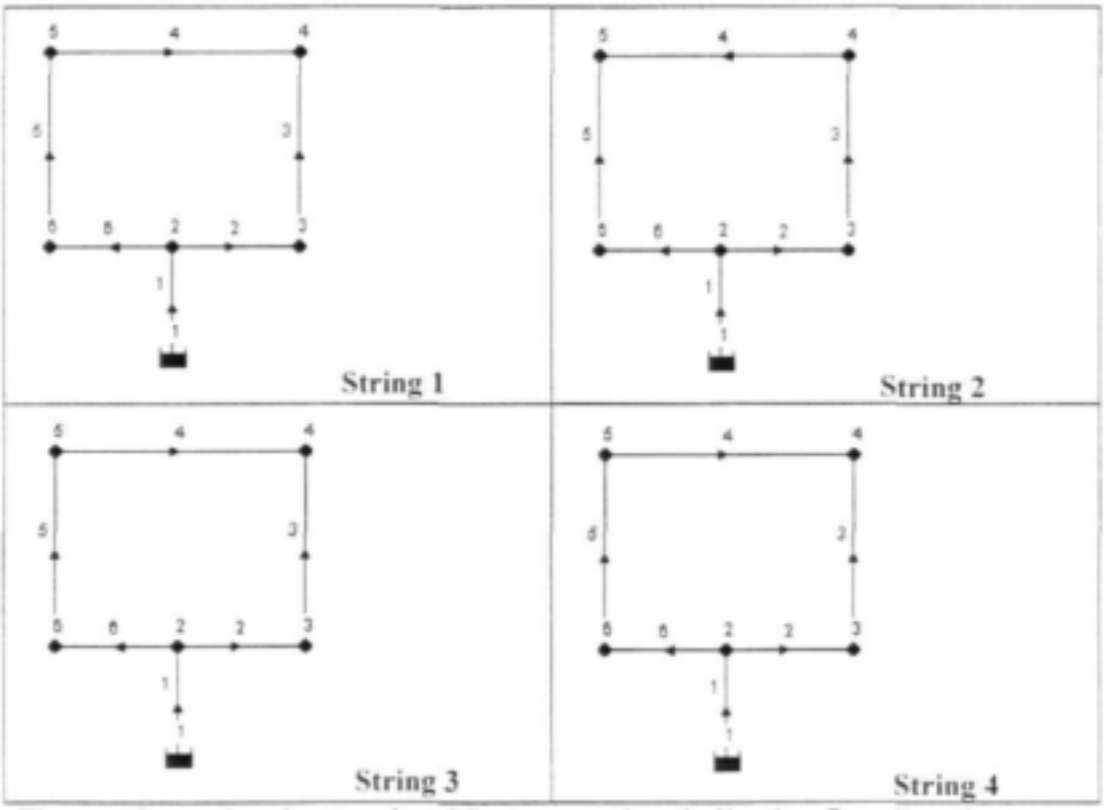


Figure A3: Analysed networks of first generation (indicating flow direction)

Table A6: Cost of string 1 (Generation 1)

Pipe	String 1 (Generation 1)					
	Diameter (mm)	Pipe cost (R/m)	Supply node	Penalty factor (b)	Length (m)	Total pipe cost (R)
1	90	32	2	5	500	1.678×10^{10}
2	110	48	3	5	750	1.911×10^{11}
3	110	48	4	5	1200	3.058×10^{11}
4	160	90	4	5	1500	8.857×10^{12}
5	110	48	5	5	1200	3.058×10^{11}
6	160	90	6	5	750	4.429×10^{12}
Total						$\mathbf{1.411 \times 10^{13}}$

All the pressures in String 1 were negative and thus the penalty factor of b = 5 were applied.

Table A7: Cost of string 2 (Generation 1)

Pipe	String 2 (Generation 1)					
	Diameter (mm)	Pipe cost (R/m)	Supply node	Penalty factor (*b*)	Length (m)	Total pipe cost (R)
1	160	90	2	1	500	45 000.00
2	110	48	3	1	750	36 000.00
3	90	32	4	1,388	1200	147 513.28
4	90	32	5	1,399	1500	191 446.28
5	90	32	5	1,399	1200	153 157.03
6	110	48	6	1	750	36 000.00
Total						**609 116.59**

The penalty factor for, for instance pipe 3 was calculated since the pressure at node 4 of 14,68 m was below the required 24 m. Pipe 3 supplying node 4 should thus be penalised with a factor *b*.

$$b = \frac{P_{required} + \left(P_{required} - P_{calculated}\right)}{P_{required}}$$

Table A8: Cost of string 3 (Generation 1)

Pipe	String 2 (Generation 1)					
	Diameter (mm)	Pipe cost (R/m)	Supply node	Penalty factor (*b*)	Length (m)	Total pipe cost (R)
1	160	90	2	1	500	45 000.00
2	90	32	3	1	750	24 000.00
3	160	90	4	1,165	1200	227 346.38
4	90	32	4	1,165	1500	85 156.56
5	110	48	5	1,146	1200	101 462.14
6	110	48	6	1	750	36 000.00
Total						**518 965.08**

Table A9: Cost of string 4 (Generation 1)

Pipe	String 2 (Generation 1)					
	Diameter (mm)	Pipe cost (R/m)	Supply node	Penalty factor (b)	Length (m)	Total pipe cost (R)
1	160	90	2	1	500	45 000.00
2	160	90	3	1	750	67 500.00
3	110	48	4	1	1200	57 600.00
4	110	48	4	1	1500	72 000.00
5	160	90	5	1	1200	108 000.00
6	160	90	6	1	750	67 500.00
Total						**417 600.00**

Rank the strings in terms of their costs **Table A10**.

Table A10: Ranking positions (Generation 1)

Ranking position	*String*	*Cost*
1	String 4	R417 600.00
2	String 3	R518 965.08
3	String 2	R609 116.59
4	String 1	R1.411 x 10^{13}

In this GA procedure a decision is made to retain 75% of the top ranked solutions of the generation in other words 25% will be discarded. The generation consists of 4 strings and thus 1 string will be discarded. Four new strings will now be generated from the remaining 3 strings.

As can clearly be seen in **Table A10** String 1 is the worst solution and thus will be discarded. The probabilities of the remaining strings are now calculated to create the roulette wheel. String 4 has the lowest cost and thus should have the highest probability of being used again in generating the new offspring (**Figure A4**).

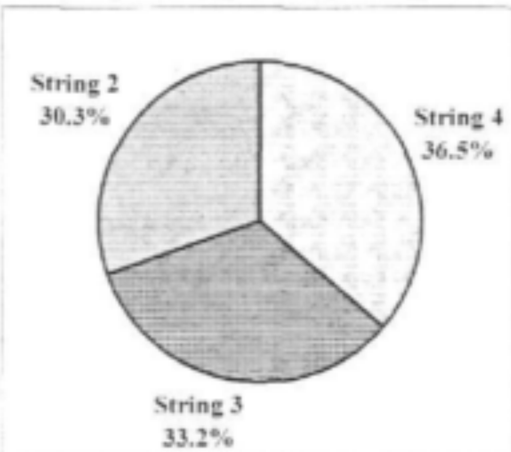


Figure A4: Roulette wheel of probabilities

The probabilities can be calculate with the following equation:

$$\text{Probability (K)} = \frac{\left(1 - \frac{\text{String cost(K)}}{\text{Total cost}}\right)}{n-1}$$

Where:

K = string number
n = number of remaining strings (3 in this case)

Probability of string 2 = 0.303
Probability of string 3 = 0.332
Probability of string 4 = 0.365

Step 5: Generate new strings

Through a random process, or the spin of the roulette wheel, the 4 new strings for the second generation are created. Since String 4 has the highest probability value, largest portion of the wheel, the chance will be the greatest that it is represented the most in the new generation.

Spinning the wheel yielded the following four new strings:

Table A11: New strings (Generation 2)

Spin	String
1	String 4
2	String 3
3	String 4
4	String 2

If these new strings would be analysed again the same solutions as detailed in **Table A5** will be obtained and thus the next step in the GA process should first be performed i.e. crossover and mutation to generate new strings.

Step 6: Crossover (pairing)

The four new strings in the network, as determined with the spinning of the roulette wheel, each consists of the 6 pipes as shown in **Table A12.**

Table A12: New strings' pipe data (Generation 2)

String	Pipe					
	1	2	3	4	5	6
New string 1	160	160	110	110	160	160
New string 2	160	90	160	90	110	110
New string 3	160	160	110	110	160	160
New string 4	160	110	90	90	90	110

The crossover is where the genes of the strings will be transferred between parents at a specific point in the string.

Firstly randomly determine which of the parents should be paired to obtain the offspring strings (children). Secondly **randomly** determine the position for the crossover to occur in each of the pairs.

It was decided to pair the strings from spin 1 and 2 and at a single position after pipe 4 and the second pairing will be between the strings from spin 3 and 4 at a single position after pipe 2. The newly paired strings will thus be as shown in **Table A13**.

Table A13: Newly paired strings (Generation 2)

String	Pipe					
	1	2	3	4	5	6
New string 1	**160**	**160**	**110**	**110**	110	110
New string 2	160	90	160	90	**160**	**160**
New string 3	**160**	**160**	90	90	90	110
New string 4	160	110	**110**	**110**	**160**	**160**

Step 7: Mutation

To force the solution to include gene strings from the total solution space and to steer away from the local optimum a mutation is performed. Randomly determine if mutation should occur (mutation probability normally < 2 %). Each pipe of each solution string has in other words the probability of mutating through the random process. For this example the random process only indicated that pipe 2 (90 mm diameter pipe) in String 2 will change. Again randomly select from the available pipe diameters (**Table A3**), excluding the current pipe diameter, a pipe diameter that will replace the current pipe diameter. The result of the mutation is shown in **Table A14**.

Table A14: Mutated strings (Generation 2)

String	Pipe					
	1	2	3	4	5	6
New string 1	160	160	110	110	110	110
New string 2	160	**110**	160	90	160	160
New string 3	160	160	90	90	90	110
New string 4	160	110	110	110	160	160

Step 7: Repeat the GA process

The newly paired and mutated strings can now be used in the hydraulic analysis to determine the fitness of this second generation. This process is repeated until sufficient convergence to an optimum has been achieved.

Using EPANET, the existing network is modified and analysed to obtain the pressures at each of the nodes (**Table A15**).

Table A15: Calculated pressures of second generation strings

Node	Pressure (m)			
	String 1	String 2	String 3	String 4
1	10.00	10.00	10.00	10.00
2	25.40	25.40	25.40	25.40
3	43.09	38.34	43.37	39.68
4	32.54	37.84	20.27	37.76
5	31.44	40.50	19.15	39.76
6	35.58	42.50	33.99	42.15

As can be seen in **Table A15** the second generation of solutions already seems an improvement on the first generation. Only string 3 has two nodes that do not meet the minimum pressure requirement of 24 m.

If the costs of the four strings are calculated as detailed in Step 4 the improvement is clearly visible (**Table A16**).

Table A16: Fitness costs of strings (Generation 2)

String	Cost
String 1	R335 700.00
String 2	R412 500.00
String 3	R388 356.84*
String 4	R386 100.00

Note: * This is a penalised cost.

The pressure failure at nodes 4 and 5 of String 3 was marginal (compared with the required 24 m) and thus the penalised fitness cost only increased the real cost of the network from R273 300.00 to R388 356.84. There will thus still be a probability that this string will be used again in the next generation since its cost is not significantly higher than that of the other strings.

The cost of a working system that meets the pressure requirement of 24 m in the first generation was String 4 at **R417 600.00** and this has reduced to **R335 700.00** (String 1) in the second generation.

Repeating this process for a number of generation could possibly provide you an optimum as shown in **Table A17** and **A18** and **Figure A5**, which has a cost of **R280 200**.

Table A17: An optimum solution (pipes)

Pipe	Diameter (mm)	Velocity (m/s)
1	160	1.29
2	110	1.34
3	110	0.81
4	90	0.11
5	110	0.77
6	110	1.40

Table A18: An optimum solution (node pressures)

Node	Pressure (m)
1	10.00
2	25.40
3	33.83
4	26.48
5	26.14
6	32.80

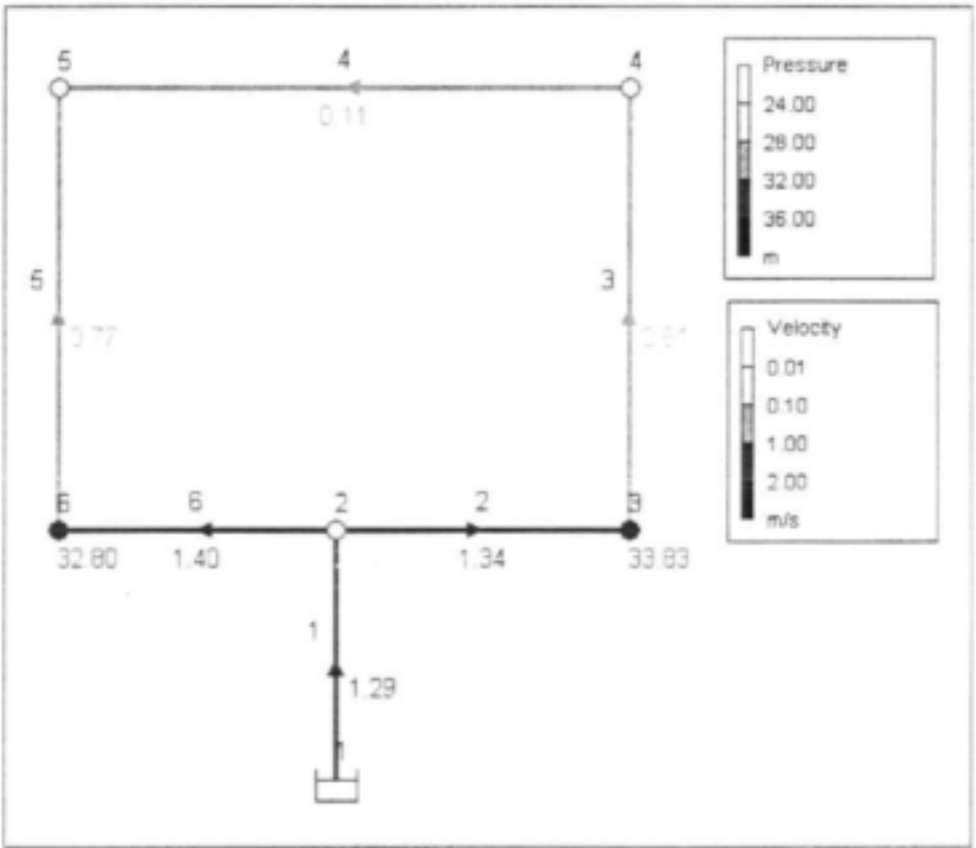


Figure A5: An optimum solution

4. Conclusion

This example that demonstrated a GA optimization by means of hand calculation can also be set-up using EPANET and GANEO and solved in less than 20 seconds.

Two software programs developed during this research study, as detailed in **Chapter 8** (GAPOP and GANEO), are available for further distribution in the water field, to expose designers to the possible uses and benefits of applying genetic algorithms in optimization.

These programs can be downloaded from either the WRC or the University of Pretoria web sites (links below).

WRC - http://www.wrc.org.za
University of Pretoria - http://www.up.ac.za/academic/civil/divisions/water/ga.html

TO ORDER: Contact **Publications** - Telephone No: 012 330 0340
Fax Number: 012 331 2565
E-mail: publications@wrc.org.za

1770053182

Water Research Commission
Private Bag X03, Gezina, 0031, South Africa
Tel: +27 12 330 0340, Fax: +27 12 331 2565
Web: http://www.wrc.org.za