Intelligent Framework for Sewer pipes' Performance Evaluation, Optimal strategy for Rehabilitation& Failures Prediction using GIS, AI and DSS

إطار لنظام ذكى لتقييم أداء شبكات الصرف الصحى ووضع استراتيجية لإحلالها ونموذج للتنبؤ بالانهيارات المستقبلية باستخدام نظم المعلومات الجغرافية والذكاء الاصطناعى ونظم دعم القرار

Samia Mohammed Abdel Halim Emera[*1], Ass. Prof Dr. Mohammed Nour El Dean[*2] , Prof. Dr. Esmail M.F. Naguib[*3] , Prof. Dr. Galal Hassan Galal-Edeen[*4]

الملخص:

تهدف الدراسة الى عمل اطار ذكى للجمع بين نظم المعلومات الجغرافية ونظم التنقيب فى البيانات ونظم دعم القرار وذلك حتى تعطي القرار الصحيح للمخطط في اتخاذه راي و تقليل الفجوة الزمنية للقرار.

وهذا الاطار سوف يكون بصفة عامة صالح لدعم انظمة دعم القرارالمكانية وبصفة خاصة أنظمة البنية التحتية والتى تعتبر شبكات المياه والصرف الصحى من أهم مكوناتها سواء على المستوى المحلى أو العالمى وذلك لما تشكاه من آثار سلبية على النواحى الجتماعية والاقتصادية.والسياسية للأفراد والمجتمع عند حدوث أية انهيارات بها. ولعل مشكلة الصرف الصحى فى مصر من أهم المشكلات التى تواجهها مصر نظرا لعدم وجود نظام يسمح بالتنبوء بوقوع الانهيارات أو تقييم الأداء الهيدروليكى للشبكات ووضع استراتيجية لاحلال وتجديد الشبكات وفقا للاعتمادات المتاحة .

Abstract:

The treatment of water and wastewater utilities problem represents the best usage of our resources on the national economy. One of the problems that face Egypt now is the ad-hoc failure of water &wastewater utilities. The rapid growth of its population yearly increased both water demands and water disposal. Serious strategies for water development & improving water disposal have to be established to prevent this failure which leads to both private and public benefit. Besides Wastewater utilities face an aging workforce, combined with the declining maintenance budget which presents challenges to infrastructure management agencies to sustain acceptable performance levels and to meet the high demands of these systems. Need for new techniques to improve the efficiency of managing civil infrastructure systems is evident. Many infrastructure management software tools have been developed to address the increasing complexity of infrastructure systems. However, most of these tools were developed for stand-alone use, and thus lack an integrated and comprehensive view to the infrastructure management process and the ability to interoperate with other tools. This research presents an integrated component-based framework that aims to integrate different tools and technologies into one coherent environment to enable infrastructure agencies to address various aspects of infrastructure management from an integrative perspective. One of the most important integration is the integration between GIS and AI .So integration with ANN will be used to develop a prediction model for pipe failures and with GA will lead find near optimal solution for sewer rehabilitation according to certain budget.

Keywords:

AI, ANN, GA, GIS.

Accepted June 20, 2012

1. Introduction

Geographical Information Systems (GIS) are becoming the universal backbone in most information system managing data. Using GIS is a great aid in graphically presenting geo-referenced information and providing the necessary tools to accurately locating information using the spatial position as the main indexing mechanism. The user can also retrieve related information without the need to find or enter tedious information as it has a lot of geographic functions such as data integration, mapping, and overlaying, buffering, projection….etc. .Actually it differs from traditional information systems in its ability to work with both *spatial data* and *attribute data*. Spatial data describes the locations and shapes of geographic features like rivers, roads, buildings and streets, while attribute data describes the characteristics of geographic features. GIS also could be thought of as a map accompanied with database (geo-database) holding information about that map [1].

GIS helps store, manage, analyze, manipulate and display data that are linked spatially. In essence, GIS relates database records and their associated attribute data to a physical location in "real" world coordinates, thereby creating a "smart map". Visualization of discrete parts of these data on a GIS map is possible by layering the data into different "themes". GIS applications can then display the intersection of various "themes".

It has a great participation in process of decision making when the decision has a geographical dimension. It is usually used beside DSS as an aiding tool to help decision makers in choosing among several alternatives and preferences in problems which include geographical dimension. While GIS is focused on simple overlay analysis and extracting knowledge and information from statistics, the actual relations between variables tend to be dynamic and uncertain. The difficulties remain concerning the nature of interdependencies between the variables which can be offered by other means.

Asset managers are required to assess critically their agency's operating environment and its competency. Moreover, they are required to make difficult decisions that have long-term and often critical consequences for their agency and/or stakeholders; decision-support tools will improve the accuracy and the soundness of their decisions [2]. With funding scarcity, more and more rehabilitation of sewers are deferred and the overall condition of assets is worsening [3]. [4] has acknowledged that engineers, technical staff, administrators and politicians all benefit if decisions about infrastructure maintenance, repair and renewal are based on reliable data, solid engineering principles and accepted economic values.

2. Problem Definition

From the aforementioned there have been computer-based infrastructure management tools to support activities in various infrastructure management domains such as bridge management, water and wastewater management. Such tools have supported a wide spectrum of functionality such as inventory and condition data, maintenance management. Besides employed a wide range of advanced algorithms for multi-objective optimization, and analytical and simulation techniques to accurately simulate and analyze system performance. However, the majority of these tools were developed to function as stand-alone systems with limited capability to exchange data. Several tools have implemented GIS functionality or supported links with other GIS systems. GIS technology have proved to be an ideal tool and to enable more intuitive and efficient mechanisms to query, explore, visualize, and analyze the infrastructure data in its spatial context. To integrate these tools into GIS-based DSS we need to take GIS technology one step ahead by combining it with AI as GIS can't stand alone in taking decision in spatial problem environment. GIS can't stand alone in taking decision in spatial problem environment .The limitations of GIS in supporting spatial decision making is coming from many reasons as [5], [6].

1. GIS system doesn't take into account the decision making preferences it was designed to only manage spatial data.
2. Current GIS don't provide the assessment and comparison of different scenarios. They give only the solutions which suitable to the given criteria.
3. In several GIS systems analytical functions focus only on buffering and overlaying operation which are limited in case of multiple and conflict criteria and objectives of decision maker(s).
4. GIS designs are not flexible enough to accommodate variations in either the context or the process of spatial decision-making.
5. All current DSS doesn't support the natural of spatial data by any way. It doesn't take into account any spatial data such as: location, references or

pointers. All current DSS systems deal with non-spatial data.

6. The decision maker need always aggregated and summarized data for strategic decision in order to perform good analysis over time which doesn't exist in current GIS systems.
7. Decision maker hope for fast answers, simple user interface and high level of flexibility which not provided by current GIS systems.

So, as result of the previous limitations of GIS in supporting spatial decision makers all researches suggested that GIS must be combined with one of analytical techniques such as GA, ANN ...etc. From this description we can understand that both GIS and DSS can't stand alone in aiding spatial decision so the need of SDSS has become essential. In this work we are concerned with combining GIS with both ANN and GA.

3. Technologies

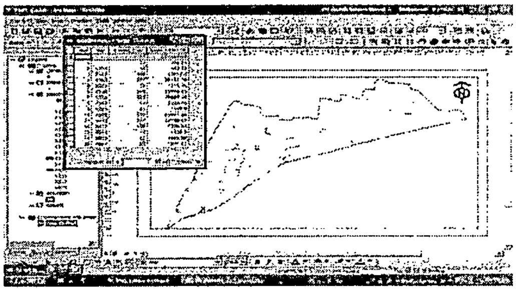
Technologies used in the proposed frame to get powerful SDSS are GIS, ANN, GA.GIS is used for management spatial data, ANN is used to extend functionalities of GIS by adding modeling and simulation capabilities while GA can be of great assistance for examining alternatives since they are designed to evaluate existing potential solutions as well as to generate new (and better) solutions for evaluation. Thus GAs can improve the quality of decision making. The integration forms SDSS which will be used to support decision makers in solving complex spatial problem. The choice of ANN from a large set of modeling techniques because is suitable for problems, whose inputs are both categorical and numeric, and where the relationships between inputs and outputs are not linear or the input data are not normally distributed. In such cases, classical statistical methods may not be reliable enough. Because ANN does not make any assumptions about the data distribution, their power is less affected than traditional statistical methods when data are not properly distributed. The choice of GA from a large set of searching techniques towards near optimal solution because it designs many solutions until no further improvement (no increase in fitness) can be achieved or some predetermined number of generations has evolved or when the allocated processing time is complete. The fit solution in the final generation is the one that maximize or minimize the objective (fitness) function; this solution can be thought of as the GA has recommended choice.

3.1. Geographical Information System (GIS)

It is a special type of information systems which lacks an absolute definition. Really, there is no absolutely agreed upon definition for GIS and this lack of an accepted definition has introduced the problem of misunderstanding about what GIS is, what is its capabilities and its applications. [7] According to [8], [7] and [9] they defined GIS similarly as: *"A collection of computer hardware, software, and geographic data for capturing, managing, analyzing, and displaying all forms of geographically referenced information"*. This means all possible tools which used in processing and analyzing of spatial data into useful information. This information is used in decision making process and in solving complex geographical problems.

By joining both spatial and non-spatial data together we obtained much valuable data rather than view each one separately. In order to be joined correctly, each feature considered a record in database and has a unique ID which uniquely identifies each record. This ID is used to link this feature with its attribute data stored as a data record in a relational database and holds the same ID value of the feature. See Figure (1) which is an example of the joining process:

Fig. (1): Joining spatial and attribute



data

GIS includes set of functions to support the acquisition, storage, querying, analysis and visualization of spatial data. These functions are called "A heart of GIS" according to [7]. They can be categorized into three main types: Input, storage and editing functions, Analysis function and Output functions [8] .GIS uses digital technology, which combines hardware and software to meet user needs in handling spatial data to produce graphical output [10].

3. 2. Artificial Neural Networks

Artificial Neural Network (ANN) is one of soft computing techniques which are a branch of AI. ANNs are powerful tools that use a machine learning approach to quantify and model complex behavior and patterns. It is an information processing paradigm that is inspired by the way biological nervous systems, such as the brain process information. A neural net consists of a large number of simple processing elements called *neurons*. Each neuron is connected to other neurons by means of direct communication links, each with an associated *weight*. The weights represent information being used by the net to solve a problem. Each neural has an internal state, called its *activation function*, which is a function of the inputs it has received. In this research we adopt the multi-layer perceptron (MLP) neural net described by [11] is one of the most widely used ANNs. The MLP consists of three layers: input, hidden, and output as in Figure (2) and thus can identify relationships that are non-linear in nature. ANNs are used for pattern recognition in a variety of disciplines, such as economics, medicine, and landscape classification, image analysis, pattern classification, climate forecasting, mechanical engineering and remote sensing. The use of neural networks has increased substantially over the last several years because of the advances in computing performance and the increased availability of powerful and flexible ANN software.

Any neural network is characterized by (1) its pattern of connection between the neurons (*architecture*), (2) its *activation function* which based on data forms and (3) its method of determining the weights on connections (*training or learning*). [9] and [12]

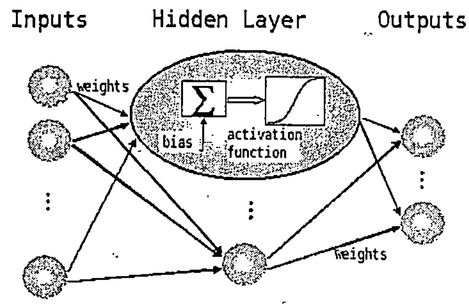


Fig. (2): Components of ANN

3.3. Genetic Algorithm

Evolutionary algorithms are search methods which imitate biological evolution of nature. Evolutionary algorithms are different from previous optimizing algorithms in that they include a search of a set of solutions. Solutions with high capability combine through transaction of elements of solution .Through affecting a little change on an individual elements, solutions also evolution through one solution. Re-combination and evolution for production of new of new solution tending to some areas of the answer are used to which proper solutions identifies in them. Genetic algorithms are particularly suitable for combinational problems with big searching space. All these results show that the performance of GA is satisfactory and efficient even though only a small portion of the total solution space was searched. Thus, GA was employed for determining rehabilitation methods and substitution materials for sewage system rehabilitation in this research due to its three major advantages. First, GA doesn't need complicated mathematical requirements and many model coefficients for optimization are friendly for general users. Unlike many optimization algorithms based on nonlinear programming and optimal control theory, there is no requirement for objective functions to be differentiable for GA [13]. Secondly, the periodicity of evolutionary operators makes GA very effective at searching for global solutions and provides multiple solutions]. GA can quickly and reliably find out optimal solutions [14].

Given the enormous size and combinatorial nature of the solution space and the complexity of the defined objectives and constraints, GA techniques seemed to be appropriate and robust search techniques. Goldberg [14] attributed the robustness of GA to the following: (1) they work with a coding of potential solutions, not the solutions themselves; (2) they search the solution space from a population of points and not a single point; (3) they work directly with the objective function, requiring no additional knowledge about its derivatives or any other auxiliary information; and (4) they direct their search by probabilistic, not deterministic, rules.

3.4. Importance of Integration between GIS and AI

AI'S main field is to simplify calculations. With this in mind, AI is looking to develop systems which trays to imitate human intelligence, without having in claim to understand constitutional procedures. Commonality among all of these algorithms is that despite all previous tradition points of view may stand against ambiguity, skepticism, carelessness, and inaccuracy. AI is a general idiom which includes some of evolutionary algorithms, genetic programming, artificial neural networks, cellular automata and Fuzzy systems. In short, that implementation which combines GIS with AI, we can expect new capabilities in supporting knowledge –based decision making and automation.AI algorithms can solve complex decision making issues, it is clear that AI approach generally in a some cases includes followings which are proceeding traditional methods : 1- data great set or unpredictable non-linearity 2- hidden pattern for decision making is so important, 3- human opinions and ambiguous related factors and decision making conditions. It is also to be mentioned that AI technology does not guarantee more accurate decision making but it is possible to have make decision more consciously [15].

The developed framework will contain the strength of GIS, ANN and GA. So, it will be useful in assessment of decision maker in complex spatial problems. This integration will lead to extend GIS to develop modern GIS (traditional GIS with AI capabilities) and therefore can be used as a Spatial Decision Support System (SDSS).

4. a framework for integration between GIS, ANN and GA

4.1. Architecture

Figure (3) shows the components of the proposed framework.

4.2. Components

A. Database components

• **Geo-databases:** the basic component of the architecture is geo-database where both spatial and attribute data are stored and managed. In complex problem, several geo-databases may be available from different sources and in different formats.

• **Spatial data base management system:** The spatial data base management system (SDBMS) is an extension of the conventional database base management system (DBMS). It is used specially to manage spatial data.

B.GIS component

GIS Basic toolbox functions

This part is related to the functionalities of traditional GIS that support spatial analysis

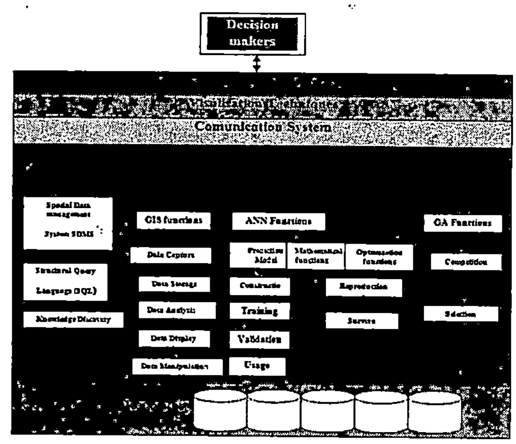


Fig . (3) : Integration framework between GIS, Data Mining and DSS

C. Knowledge Discovery component

• **Spatial data mining:** It has been used successfully to extract relevant pieces of knowledge from huge traditional spatial databases. In particular, spatial data mining (SDM) is a very demanding field that refers to the extraction of implicit knowledge and spatial relationships which are not explicitly stored in geographical databases

D. ANN components

Modeling functions:

Either classification or prediction model is used a set of phases must be followed to get best results. 1st model construction functions are required to build the model in an appropriate form.2nd model training, testing and validation functions are required to measure the accuracy and the performance of the model.3rd using the resultant model should be provided through a systematic way.

Mathematical functions:

A set of strong mathematical functions are required because of the nature of neural network and its learning method and algorithms. Statistical functions are also required to perform a lot of tasks in evaluating models.

Optimization functions

This part is responsible for enhancement the performance of neural network to get best results, Such as normalization and transformation functions. These functions are used to increase the accuracy of the resultant model.

E. Genetic Algorithm

The GA-based optimization model involves ten operation steps, including preprocess, chromosome coding, generating an initial population, estimation of total rehabilitation cost, estimation of rehabilitation efficiency, computation of fitness, generation of the next population, crossover operator, mutation operator, and termination criterion [16].

F-Communication Layer Component

The communication system represents the interface and the equipment used to achieve the dialogue between the user and the SDSS. It permits the decision maker to enter queries and to retrieve the results. It is responsible of the communication between internal components. The movement and transformation of data between GIS components, ANN components, GA components and knowledge discovery components is also from its responsibilities.

G. Visualization techniques and GIS Interface components

GIS visualization techniques are used to present the spatial results of analysis. A map is usually the result of any geographical project so all needed tools for producing the map should be included. Then user interface can be developed by any programming language to provide an easy way for interaction between user and the components.

5. Case Study

5.1. Problem Definition

Pipe lines represent the largest capital investment made by sewers management authorities, yet the condition of these buried assets is generally unknown until some types of failure occur.

Most sewer repair and rehabilitation projects have been performed using "management by crisis" techniques. When the sewer fails and sewage begins to back up or when the sewers odor becomes annoying then the public complains and the problem is given immediate attention. A small investment at critical stage of deterioration can prevent much larger investment for major repair or replacement from being needs later [17]. Therefore there is a need for implementing appropriate maintenance programmers to maintain sewers effectively .Deteriorating sewer service levels combined with limited budgets are forcing agencies to examine many of their current states of practice for sewer maintenance and rehabilitation (M&R) in an attempt to be proactive rather than being reactive. One of the primary reasons for this 'reactive' approach to sewer M&R is the inability to predict the future condition of the sewer pipes .Besides these technical issues, a wide range of managerial, financial, social, and political issues are also involved. Civil infrastructure systems are typically managed by a multi-disciplinary team of stakeholders, each represent a particular view of the system. Understanding and supporting an integrated approach to infrastructure management is a key factor that would determine the success of any technological solution in the domain. An integrative and multidisciplinary approach to infrastructure management [18] is necessary to address the wide spectrum of requirements and constraints of the diverse perspectives of different stakeholders.

5.2. Sewer Problem in Egypt

Problems of Sewage projects in Egypt: are due to many factors: Delay in Execution of projects (delay in confiscation procedures for projects, protracted periods of time for permit procurement erect (Agriculture-irrigation-roads- rail road's, insufficient finance (allocated investments does not exceed 40% from proposed budget), required excavation in cities may reach a depth of 7m requiring dewatering, soil grouting and boarding of trenches, as well as bracing of existing random buildings which are mostly wall bearing type with no foundations in very narrow streets resulting in not executing water and sewerage projects simultaneously, negative effect of overpopulation on development plans which has also direct impact on infrastructure, current Sewage capacity represents 50% of produced water capacity (world rate 80%).Low coverage Ratio of Sewerage Projects in rural areas. Past strategies focused on finishing projects in cities with industrial, commercial, tourist activities and high pollution. Coverage ratio in rural areas will reach 7% by the end of the current plan leaving 93% to be covered in the coming five-year plans. More investment is required in Sewerage projects as: sewerage projects are highly expensive with almost no financial dividend, And Service cost is L.E 4000//m3/ day.

This huge amount of projects, and tens or hundreds of kilometers of sewer lines, that have already been executed or still to come, need a thorough and easy-way-to- reach documentation of all sewer projects with their minute details. Details include location, diameter, depth of pipe, date of execution, material of pipes, slopes or reverse slopes, obstacles faced during execution, and many other factors that are not only helpful but also needed. The need for this information in fact is required in order to assist in any future works in the same area of the sewer pipes. Also, it is needed to anticipate future sewer replacement if any. Instead of waiting until breakage or clogging of a sewer line occurs, the information stored and managed will tell us in many ways how to maintain the sewage lines functioning properly without disruption of the lives of the citizens in their homes and in the streets. It also allows the decision makers to focus on different scenarios for manipulating the sewer network for purpose of rerouting or future residential expansion or development. What is said regarding sewer is 100% applicable to all other utilities? What is presented here for sewer network indeed can be not only applied to other utilities but also extended and amalgamated with these utilities to capture the –almost-full picture of utility management required for a better life for everyone.

5.3. Proposed Application

Most of the underground infrastructure life expectancy depends on the soundness of its original construction, on the amount of its usage in its surroundings, and on the amount and frequency of maintenance. Just as roads must be repaired and resurfaced, as bridges require strengthening to sustain greater loads, and as the steel structures have to be maintained , sewers must be inspected, cleaned, and repaired to perform to their carrying capacities. To solve their sewer pipe breakdown problems, the present sewer pipe system must be rehabilitated and brought up to minimum acceptable standards, so that these problems do not occur in the future.

Currently, sewer problems are often solved by reactive or crisis maintenance by responding to breakdowns. This technique is rarely cost-effective in the long –term.

The application aims to build prediction model using ANN to predict pipe defects based on expert opinions and hard data obtained by monitoring pipe performance over 50-year period. Then it presents a spatial representation of these defects in this time as a visualized map. In the proposed model GIS is used as a good management tool for spatial data. It is used to prepare, organize, display and visualize spatial data while ANN is used to build the prediction model of sewer defects (train-validate-test) using collected data by Closed Circuit Television Camera (CCTV) Camera equipment that has become the most-effective method for viewing sewer pipe condition and is linked to GIS database using the ID attribute. Information gained from CCTV can be used to identify and assess structural problems such as cracks, displaced joints, corrosion, deformation, missing pipe pieces, and collapses as well as to locate both sources of infiltration and hydraulic problems. Ninety five percent of public sewerage network is non-man eatery with pipe diameters too small for effective manual inspection [19].GA was employed for determining rehabilitation methods and substitution materials for sewage system rehabilitation in this research due to its three major advantages. First, GA doesn't need complicated mathematical complicated mathematical requirements and many model coefficients for optimization are friendly for general users. Unlike many optimization algorithms based on nonlinear programming and optimal control theory, there is no requirement for objective functions to be differentiable for GA [13]. Secondly, the periodicity of evolutionary operators makes GA very effective at searching for global solutions and provides multiple solutions [20]. GA can quickly and reliably find out optimal solutions [14].

5.4. Study area description

This study presents case of study .The case is a part of the Greater Cairo for Sanitary Drainage project to provide GCSDC with the potential to become self supporting , and represents one of the concluding phases a significant investment in the Greater Cairo Project .The project comprises Drainage Area Planning (DAP). Develop a GIS Database for the whole of Greater Cairo, including initially the study. Develop a phase printed program for solutions to Structural and Hydraulic deficiencies for the study area (SHAMAL Area).This Pilot DAP was covering the whole of the area of north central Cairo: the 'Shamal' part of East Sector.This area covering an area of over 800 hectares and a population of over 600000.Its bounded to south &East by railways, the West by the River Nile and the North by Ismailia Canal as in Figure (4).The area is very flat, being part of the River Nile flood plain. The ground consists of alluvium soils with a high ground water table. Due to insignificant rainfall in Cairo, the sewer system is classified as Foul.

-Total length of –all sewers-is 225 km

-Total length of -Critical sewers is 49 km

- Total length of sewers greater than 600mm have total length 26.8km

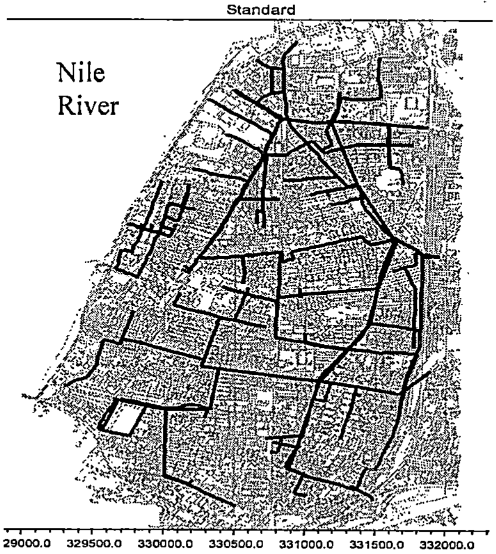


Fig. (4): Sewer Network of Study Area

5.4.1. Definition of Study Area Requirement

- **Initial Critical Sewer Assessment**

As no reliable data was available, data was gathered for 'Significant' sewers, defined as 300 diameters and above.

- **CCTV Outline Requirement**

The requirements for CCTV Survey were for all Critical Sewers, Collector sewers (diameter 600mm and above), local sewers diameter 300mm and above, and in addition any problem sewers. No CCTV had been carried out prior to the study.

- **Modeling Requirement**

The Hydrodynamic model applied for the study is the Sewer GEMS V8i SELECT series 2 (From Bentley). Flow Model. Pipe Flow Model is a computational tool for simulations of unsteady flows in pipe networks with alternating free surface and pressurized flow regimes. The fully dynamic model allows flow features such as backwater effects and Surcharges to be precisely simulated. Inputs to the model are descriptions of the elements in the pipe network and boundary conditions, such as wastewater inflow. Sewer GEMS V8i SELECT series 2 (From Bentley) [21] enables a description of a variety of networks elements, such as standard and irregular shaped pipe cross-sections, manholes, basins, weirs, pumps, gates, flow regulation etc.

The features implemented enable a realistic and reliable simulation of the performance of both existing network systems and, those under design. However the reliability of such a simulation model depends on the available data background and data quality.

- **Flow survey requirements**

One fixed monitor was required to establish the diurnal pattern. Spot measurements were required for all collectors and major flows on local sewers

- **Asset Survey Requirements**

All assets were required to be surveyed. Previous Manhole Surveys had been found to be inaccurate and so resurvey of all significant sewers was required, except where as built drawings were available

- **Additional data requirements**

Information on industrial flows was obtained from the Industrial Flow Department (e.g. Type, Location)

5.4.2. Model Construction and Verification

Because spatial is special, GIS applications require special kind of spatial data.

There are two types of data required for GIS application

- Assets Data
- Application data

Assets data define the physical infrastructure such as water mains and manholes. Assets data form the backbone of a GIS and therefore should be in place before developing any GIS application. Application data are required for creating GIS applications, such as work order management and H&H modelling [22].

A-Model build up

- Network Data

All network data are collected from maps and from on-site inspections. Collected data are entered into a database (the ArcGIS (version 10) database system).

- Creating the Network

Some quality assurance of the network data was needed before starting the modeling activity and the network was accordingly adjusted after the field investigations.

Only the 300 mm diameter and bigger pipes of the Shamal sewerage network were selected.

- Catchments definitions

As a next step, catchments were defined, based on available geographical and population information, In other words, catchment's locations were specified, residential areas and population

- HD boundary conditions

The boundary conditions to the hydrodynamic model consist of inflow of wastewater from households and some additional inflows. The population has been counted for the whole catchments expected the year 2000 population of 603561 PE (PE = Person Equivalents) within the model area and the year 2015 population of 618820 PE.In other word expecting population growth 2000-2015is 2.5%.

- Flow measurements

A comprehensive program of flow measurements within the pilot catchments was completed this record was consequently elaborated to obtain a typical daily flow variation for the catchments. There were no noticeable differences recognized for working days and weekends or/and public holidays.

- Model Statistics

Are listed below: as in Table (1)

Table (1): Model Statistics

parameter	value
Number of	125
Total catchments	416 ha
Residential	521 ha
Number of PE 2000	603 561
Population density	1159 inch./ha
Population growth	2.5%
Waste water	167 l /c/day
Number of nodes in	1017
Number of outlets	6
Number of circular	943
Number of specific	55
Total length of	49 km
Available pipe cross sections (mm)	2870, 2200, 2080,1860.1800,1600, 1500, 1460, 1200, 1000,900,750,600,500, 450.375.300 “horse shoe” profile 1400x1200 1500, 1460, 1200, 1000, 900, 750, 600, 500, 450, 375, 300, “horse shoe” profile 1400x1200,450, 375, 300, “horse hoe”profile1400x120
Overall outflow	1770 l/s
Daily waste water	153 000 m^3

5.4.3 Hydraulic Assessment

Peak dry Weather Flows for year 2001, obtained through the process of model verification, were used for the hydraulic assessment .This was also followed for the year 2015.The hydraulic assessment was made for flows for the year 2001 and 2015.There was no flooding predicted (or observed) and levels of surcharge are within acceptable limits (less than one meter above top of pipe). Specifications determine significant hydraulic problems [23] examples are as follows

Examples of hydraulic problems

- H1: Pipe slope is adverse negative.
- H2: Structure bottom is above pipe invert (s).
- H3: Structure top is below pipe crown.
- H4: Pipe does not meet max cover constraint.
- H5: Pipe does not meet min slope constraint.
- H6: Pipe does not meet min velocity constraint.
- H7: Pipe discharge is above full flow capacity.
- H8: Pipe discharge is above design flow capacity.
- H9: Pipe slope is horizontal.
- H10: Structure is flooded.

We take a sample from the model as Hydrodynamic model to be applied for the study is the Sewer GEMS V8i SELECT series 2 which is completely integrated with the ArcGIS (version 10). Samples of the profiles are listed in the following figures: Figure (5) and Figure (6).

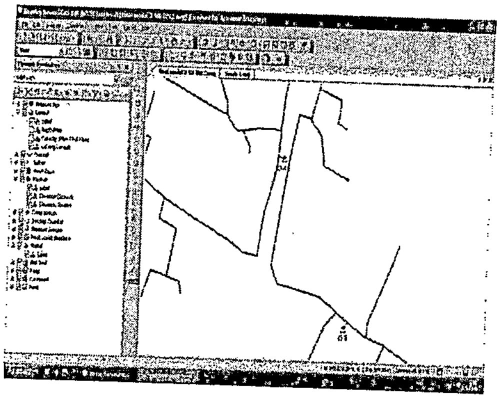


Fig. (5): Sample of Hydraulic Model Profile

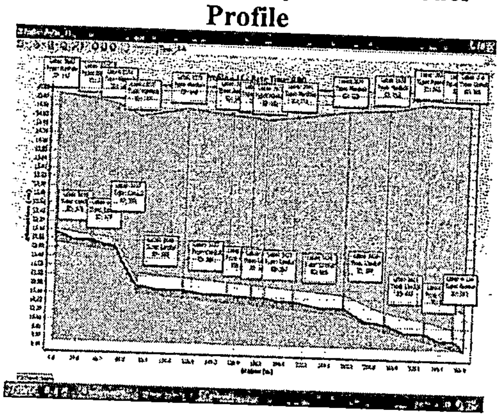


Fig. (6) : Sample of Longitudinal Profile

Advantages of using computerized solutions for common hydraulic problems [24].

- The amount of time to perform an analysis can be greatly reduced.
- Computer solutions can be more detailed than hand calculations.
- Performing a solution manually often requires many simplifying assumptions.
- The solution process may be less error-prone
- Unit conversion and the rewriting of equations to solve for any variable are just two examples of where hand calculations tend to have mistakes, while a well-tested computer program avoids these algebraic and numerical errors.
- The solution is easily documented and reproducible.

Because of the speed and accuracy of a computer model, more comparisons and design trials can be performed in less time than a single computation would take by hand. This results in the exploration of more design options, which eventually leads to better, more efficient designs. Proposed Hydraulic deficiencies are located as in figure (7) as Hydraulic Model database.

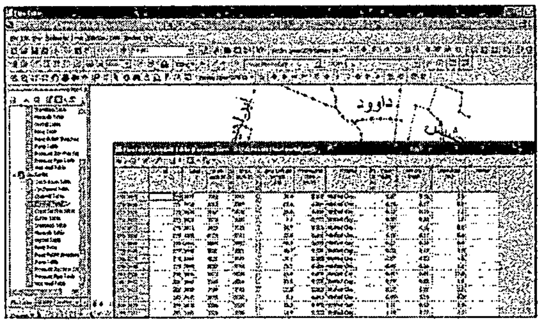


Fig. (7): Hydraulic Model database

5.4.4 Structural Assessment

The third edition of the Sewer Rehabilitation Manual (SRM) [25] recommends thatCritical sewers in the drainage area catchments are graded according to their internal condition in order that pipes with the most severe structural defects can be prioritised for improvement.

The SRM details 5 grades for sewer condition as follows:

Grade 5 Collapsed or collapse imminent .S5

Grade 4 Collapsed likely in foreseeable future. S4

Grade 3 Collapsed unlikely in near future but further deterioration likely **S3**

Grade 2 Min. collapse risk in short term but potential for further unlikely in near future but for further deterioration likely. S2

Grade 1 Acceptable structural condition. **S1**

Video footage from the CCTV surveys was coded according to the SRM and the coded data was then analysed to produce sewer grades for each sewer length as in Figure (8).

Fig. (8): CCTV Inside surfaced pipe showing fracture

The results of this procedure are tabled by Grade for Collectors and Local Sewers in Table (2) below.The predicted results will then represented visually using GIS.

Table (2) Structural Performance Grades for Critical Sewers

	Total CCTV	Grade 1-3	Grade 4	Grade 5
Collector	10 690	10 073	474	143
Local	5 680	5 540	0	140
Total	16 370	15 463	474	283

6. Experiment

The experiment has been done over a real data set of sewer pipes in Egypt with both structural and hydraulic deficiencies. The attributes in the sewer dataset are: Diameter,Range_diameter,Shape,Material,Year_Built,Range_year,Length,Range_length, Slope,Width,Lining, S#, H#. There are 11061 records of sewer pipes of various types of Materials and different diameters built since 1900-2002 with different kinds of structural &hydraulic deficiencies. There are 6883 records containing structural (S#) &hydraulic deficiencies (H#) .Our purpose develop a SDSS to decision makers to help in the strategy of sewer performance (structural defects and hydraulic deficiency) and sewer rehabilitation taking into consideration both technical and economic criteria in decision making. To involve both these incommensurable criteria, we propose a specific model that uses multi-objective optimization model with the help of modified genetic algorithm (MGA) for selection of pipes candidates for renewal according to certain budget. The application aims to build prediction model using ANN to predict pipe defects based on expert opinions and hard data obtained by monitoring pipe performance over 50-year period.. Then it presents a spatial representation of these defects in this time as a visualized map. This forecast of failures allows for the assessment of future utility expenditures using an adequate objective function.

6.1. Using genetic algorithm for optimization

The genetic algorithm (GA) is a search heuristic that mimics the process of natural evolution. This heuristic is routinely used to generate useful solutions to optimization and search problems.

Genetic algorithms belong to the larger class of evolutionary algorithms (EA), which generate solutions to optimization problems using techniques inspired by natural evolution, such as inheritance, mutation, selection, and crossover. In general form, major function of a genetic algorithm is simple. First a set developed from all better solutions. Then more proper method combines with each other to introduce new solutions. Finally new solutions combine with each other to replace weaker primary solutions and this procedure repeated again. Each time the new genetic solution is introduced, a purpose functions will be evaluated, and answer ranking in the current up dated based on compatibility level actively. Such ranking takes place in process which finally the best answers will have the greatest chance to be chosen as parents, just like survival of the fittest principal in the nature. Solutions with proper functions can be selected several times to replacing the election, while inefficient structure may never be selected

6.1.1. Data preparation and data preprocessing

The task of data preparation is to transform the transaction in the database into the format we need so as to improve the quality of data and consequently improve analysis. First we get rid of noisy data. Second, we transform the attribute values in the transaction in the form of codes. There are two sources of data the first presents structural defects taken from CCTV and linked to the GIS data base via the attribute 'pipe-id' the codes of defected pipes is 'T4'.The second represents the hydraulic deficiencies of the pipes taken after running the hydraulic Model using Sewer GEMS V8i which is completely integrated with ArcGIS version 10 taking into account data measured by the Flow-meter in the field for critical parts of the network. The output deficiencies are represented by attributes 'T1' represents (d/D hydraulic radius),'T2' presents (q/Q full %) and 'T3' presents velocity. If any of these attributes does not match the Egyptian code it takes value '1'otherwise the value is '0'. 'T4' presents the grade of structural deficiency of any pipe according to (SRM)'T5' represents the summation of 'T1','T2','T3'(hydraulic deficiencies) and'T4'..Field'8' represents priority given to the pipe according to both structural and hydraulic deficiencies, field '9' represents cost of renewing the defected pipe. Table (3) summarizes these technical conditions.

Table (3) Summary of Technical Conditions

Pipe-Id	T1	T2	T3	T4	T5	priority	Cost
	d/D	q/Q	Velocity	J	Σ3+4+5+6		
	0 OR 1	0 OR 1	0 OR1	1			
	0 OR 1	0 OR 1	0 OR 1	2			
	0 OR 1	0 OR 1	0 OR 1	3			

6.1.2 Problem formulation

The use of a GA requires an accurate definition of design variables, objective functions, and problem constraints. We define a renewal policy as the string formed by the design variables of the pipes that are candidates for renewal. For p representing a set of pipe candidates for renewal, (j E [1, p]) and three renewal alternatives Xj, where (Xj E {1, 2, 3}), for a renewal policy, ($X1$, $X2$,... X_p) are illustrated by the Figure (9). Each chromosome represents a set of defected pipes each with a code representing technical priority

Code 1 is "the mandatory repair only "alternative which do nothing until emergency reparation of a pipe is required (i.e. when a break occurs).

Code 2 is the "replace alternative", (replacing the existing pipe by a new pipe with similar characteristics (i.e. one with the same diameter but necessary one made with the same material)

Code 3 is the "reinforce" alternative which means to reinforce the pipe by replacing it with a new pipe of enhanced diameter

For *n* pipes to be considered candidates for renewal we define a string constraints variable *P1* to *Pn* where *(j Є [1, n])* and three renewal alternatives *Pj* where *(Pj Є {1, 2, 3})*

3	1	2			2	3	1	2	1

P1 p2 p3 pn-1 pn

Fig. (9): layout of a chromosome

Objective functions and constraints

For the considered problem, a GA does not use the objective function itself, but considers a fitness function which evaluates solutions according to several objectives. We define one objective function that describes the available budget for sewer rehabilitation.

```
Function f=my_ fitness(x)
 Global B
 B=29.1*1e6; % Budget
 f=abs (B-sum(x));
```

6.1.3 Running the MGA

In this research the genetic operation was run with crossover probabilities P_c, of 0.85 and the mutation probability, P_m, was 0.002.uniform mutation was adopted to efficiently enlarge the searching space [26].We considered a population of 200 possible strings that expressed possible renewal policies. Each string comprised 50 codes corresponding to the 50 pipes candidates for renewal. The MGA allows assessment of 5000 solutions among all possible solutions equal to 3^{50} .Many scenarios were done using different budgets (what if analysis) to find optimum solution of rehabilitation of sewer pipes. Figure (10) represents a sample of one of these scenarios. The results are presented in table (6) for three runs with different budgets (20, 25 and 29 million L.E) for sewer renewal.

Fig. (10): sample of output of Run3and budget =29.1*10^6

6.1.4 GA Results Analysis

Outputs files from different runs (scenarios) are passed through:

1- GUI to be mapped and visualized using GIS to illustrate pipes to be rehabilitated according to given budgets.

2- Linked to Database of defected pipes through 'Pipe-ID' as shown in Figure (11) and final results are summarized as in Tables(4,5,6) and represents as graph in figure (12).

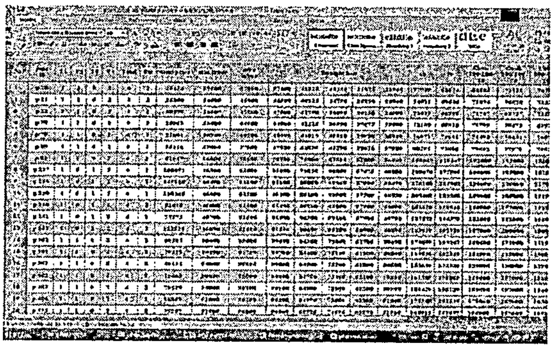


Fig. (11): sample of candidate pipes for rehabilitation according to MGA

Table (4) Results (Run 1) using MGA

Budget 20 million L.E.

Priority Def-acts'	No Of Pipes	Repair-costs-million-LE	Budjet-20 million-LE Run 1	% of achievement
1	28	4174118	4026330	96.5
2	155	8933225	8167696	91.4
3	295	18027809	7805974	43.3
	478	31135152	20000000	64.2

Table (5) Results (Run 2) using MGA

Budget 25 million L.E.

Priority Def-acts'	No Of Pipes	Repair-costs-million-LE	Budjet-25 million-LE Run 2	% of achievement
1	28	4174118	4029999	96.6
2	155	8933225	8240113	92.2
3	295	18027809	12729888	70.6
	478	31135152	25000000	80.3

Table (6) Results (Run 3) using MGA

Budget 29 million L.E.

Priority Def-acts'	No Of Pipes	Repair-costs-million-LE	Budjet-29 million-LE Run 3	% of achievement
1	28	4174118	4048850	97
2	155	8933225	8265354	92.6
3	295	18027809	16685796	92.6
	478	31135152	29000000	93.1

From the abovementioned tables it is obvious that:

a-The percentage of rehabilitation of all the pipes increased with increasing budgets with percentage according to their priorities.

b-Within the same table for each run for the same budget the percentage of rehabilitation of pipes are proportional to their priorities.

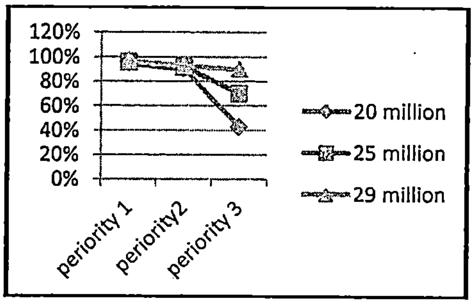


Fig. (12): GA Results obtained for. different budgets and priorities

3- Visualized results of sample of defected pipes (pipes 43, 44) resulted after running MGA and mapped by GIS are in figures (13), (14) respectively.

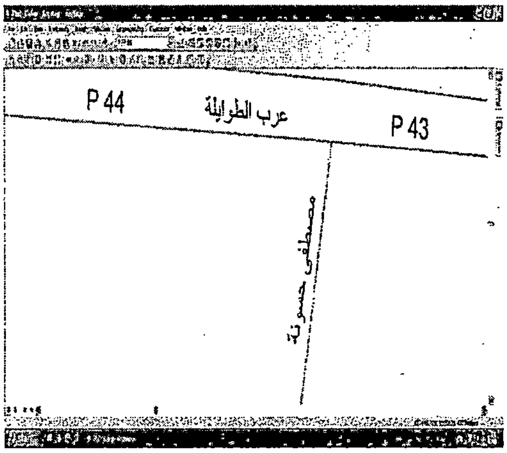


Fig .(13) : Layout of pipe 43, 44

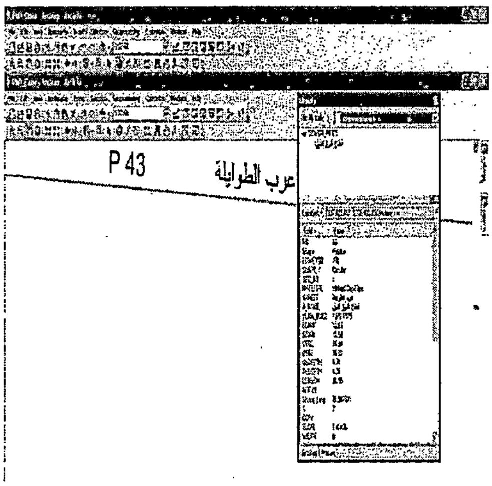


Fig. (14): Details of pipe 43

6.2. Using ANN for prediction

6.2.1. Building Neural Network Condition Rating Models

The developed model used in this research was designed using back propagation (BPNN). Back propagation network is widely used when the model is required to have good generalization capabilities to do accurate predictions. The general methodology adopted to carry out neural network analysis is briefed as: data is collected from different sources and cleaned (data of two different years). 2nd the abstraction of spatial relations is done by using GIS components (map objects components). 3rd the data is transformed into appropriate form for ANN (normalized). 4th Mat lab is used for building a Sewer Failure prediction model using ANN (training-testing-validating). This step is done by linking the Mat lab with desktop application. 5th after finishing the model is used by inserting a dataset of current data and get predicted values as an outputs.6th the output values are passed to GIS components and represented visually as map. The output will be the predicted pipes in the near future depending on the historical and current data and status.

Network Architecture

Architecture: 3-layer perceptron fully interconnected neurons

Algorithm (Training): Back-error Propagation

Objective: To predict Sewer pipelines defects

Input: Factors influencing sewer pipeline structural deterioration training dataset

Output: Estimated defect

Dataset: 4222

Dataset Date: 5-year from 2000 – 2005

Dataset Ratio: Training > Testing > Validation (60:20:20)

How it works?

• ANN adds the intelligent factor into data analysis.

• Non-linear processing through hidden layer using summation and activation function.

• Algorithm:

- Pass training data through network.
- Adjust network weights to reduce error.
- Validate network with out-of-sample data.
- Use network to predict real data
- Adjust network weights to reduce error.
- Validate network with out-of-sample data.
- Use network to predict real data

Figure (15) illustrates a neural network model to SEWER in the study. Input training and testing datasets that are stored in Microsoft Access are called and form a BPNN structure. ANN learning computes the prediction output and is adjusted by an error term with the required output to minimize mean absolute error (MAE), the same way to reduce root mean squared error (RMSE).

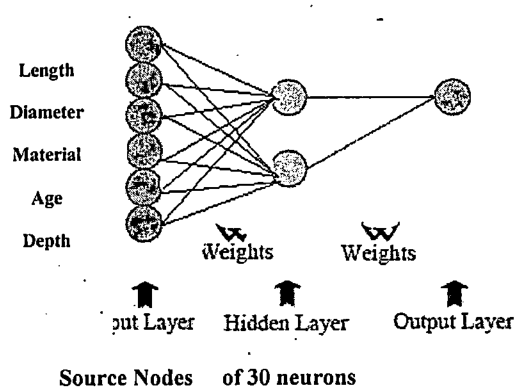


Fig . (15): BPNN model for predicting sewer failures

The developed model used in this research was designed using back propagation (BPNN). Back propagation network is widely used when the model is required to have good generalization capabilities to do accurate predictions

.The general methodology adopted to carry out neural network analysis is briefed in

6.2.2. Model Validation by Mathematical Diagnostics

The developed model is tested with real case data during its validation process as described in the following section.

Data obtained from the Holding Company Cairo were divided into three sets: training data _60%_, test data _20%_, and production data _20%_. The first two sets are used to develop the models _i.e. to train the neural network_, whereas the production data are not exposed to the network during training. This production set of data are used to test the performance of the trained network. In other words, the production set of data are used to validate the developed model and test its accuracy in predicting the condition rating of sewer pipes. The developed BPNN model was validated using twenty percent of the collected data, which the networks were not exposed to during their training. Performance of the model was tested with respect to certain mathematical validation diagnostics as recommended in literature. _

Coefficient of Determination R^2

The second indicator known as the coefficient of determination _R^2_ is the widely used statistical parameter which depicts the degree up to which a model is capable capturing the variability in the structural and operational condition explained by a sewer's attributes: age, diameter, material, length, etc. in the modeled process.

As seen in Figure (16) R=.69175 then R^2 = 0.49 < 1

a- Root-Mean-Square Error

This indicator represents the goodness of fit and can be estimated where RMSE =root-mean-squared error. A model with a RMSE value close to "0" is considered a good fit for its purpose whereas one with a RMSE value close to "1" is not a valid model. In this research as we can see in Figure (33) which represents determination factor R ,Figure (17) presents performance of the developed BPNN using 30 neurons while Figure (18) represents model MSE = 0.111 so RMSE = 0.335 close to "0" so the model is valid.

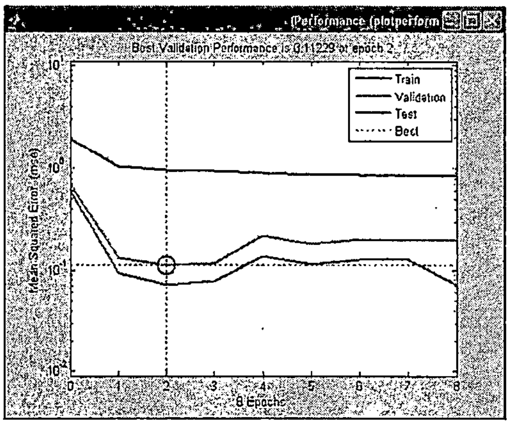


Fig . (16): BPNN model performance

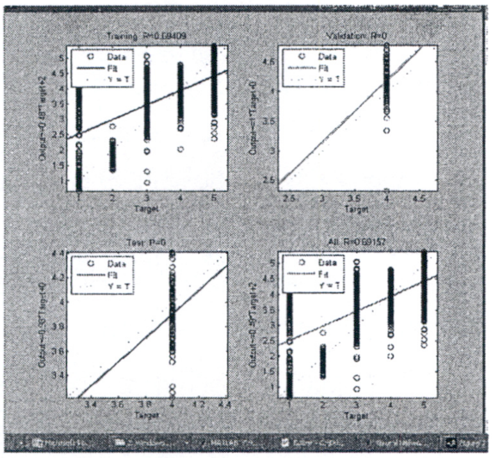


Fig. (17): BPNN model determination factor MSE

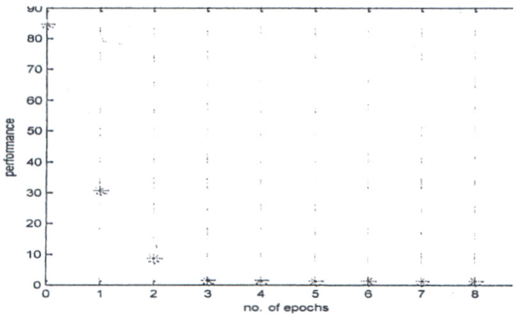


Fig . (18): BPNN model performance (30 neurons)

b- Implementation

The model is implemented in prototype MATLAB that operates within Microsoft Windows. This user-friendly software generates tabular and spatial output and provides user interface with dialog windows, menus and toolbars.GUI can only invoke the whole application through a shell function into C# interface. However, ANN weight file can be linked to spatial layers in MapInfo using ID number.

c- Results

Examples of predicted results for defect: S2 is shown in Figure (19) for pipe-ID 252 and S3 is shown in Figure (20) for pipe-ID 1054.These results were visualized through GUI built in this research.

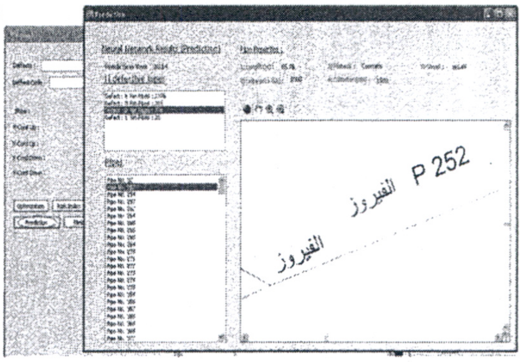


Fig (20) predicted results defects (S3) for pipe-ID 252

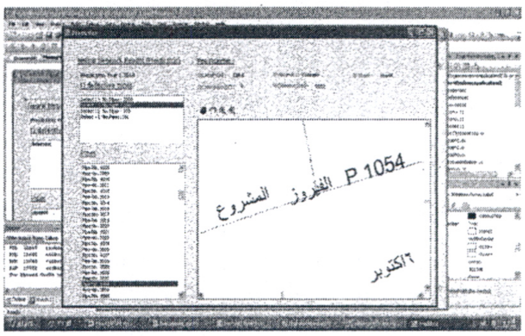


Fig (20) predicted results defects (S2) for pipe-ID 1054

7. Conclusion and future work

In this research, GIS is used in collecting, management and querying of spatial data, the decision process is based on forecasting pipe failure using Artificial Neural Network (ANN) to extend the functionalities of GIS to include modeling and simulation capabilities. Optimal search is frequently required in many GIS applications. However search may becomes very complex when it involves multiple sites, various constraints and multiple objectives GAs can be of great assistance for examining alternatives since they are designed to evaluate existing potential solutions as well as to generate new (and better) solutions for evaluation. Thus GAs can improve the quality of decision making. In this work GA designs many solutions until no further improvement (no increase in fitness) can be achieved or some predetermined number of generations has evolved or when the allocated processing time is complete. The fit solution in the final generation is the one that maximize or minimize the objective (fitness) function; this solution can be thought of as the GA has recommended choice.

8. Contribution

The main contribution of this research is proposing a component GIS-based frame to deal with the problem of water and wastewater networks renewal and the consideration of both technical (assessment of structural deficiencies and hydraulic performance) and economic criteria in decision making. To involve both the incommensurable criteria, we propose a specific model that uses multi-objective optimization with the help of a modified genetic algorithm (MGA), decision process is based on forecasting pipe failure using intelligent prediction model. That implementation which combines GIS with AI, we can expect new capabilities in supporting knowledge-based decision making issues. There is no doubt that tools and technologies used in supporting activities in various infrastructure domain such as inventory and condition data management, maintenance management, performance management, advanced algorithms for deterioration models multi-objective optimization when integrated into one coherent environment as GIS-based frame would facilitate efficient data sharing and exchange, and improve the consistency and quality of infrastructure information, and making decision more consciously.

The frame is innovative because it proposes acceptable renewal policies to be implemented to meet set economic and technical objectives, while respecting technical specification and considering hydraulic and structural criteria for pipe renewal selection. It proposes a set of feasible renewal policies which offer flexibility for water and wastewater utility managers. Besides the frame serves as reference architecture to support the implementation of distributed GIS-based infrastructure management system as many of the developed concepts and techniques are equally applicable to other infrastructure systems

References

1. **Chang Kang-tsung,** "Introduction to geographic information systems", Book, McGraw Hill companies, 3rd edition, ISBN 0-07-060629-3, 2006.

2. **NSW Asset Management Committee. (2003): Total Asset Management Manual. New South Wales.**

3. **GAO (2004),**" Water Infrastructure: Comprehensive Asset Management Has Potential to Help Utilities Better Identify Needs and Plan Future Investments "Report to the Ranking Minority Member, Committee on Environment and Public Works, U.S. Senate, GAO-04-461. Washington DC, US General Accounting Office.

4. **Vanier, D.J. (2001):** Why industry needs asset management tools. Journal of Computing in Civil Engineering, 15(1): 35-43.

5. **Laaribi, J. Chevallier and J. M. Martel, (1997).** "A Spatial Decision Aid: A Multicriterion Evaluation Approach", Article, Computers, Environment and Urban Systems, Vol. 20 - pp. 351-366, 1997.

6. **Salem Chakhar Jean and Marc Martel,** "Enhancing Geographical Information Systems Capabilities with Multi-Criteria Evaluation Functions", Journal of Geographic Information and Decision Analysis, Vol. 7, 2003.

7. **Chang Kang-tsung,** "Introduction to geographic information systems", Book, McGraw Hill companies, 3rd edition, ISBN 0-07-060629-3, 2006.

8. **Demers Michael N.**, "Fundamental of geographic information system", Book, U.S.A, John Wiley & Sons, 2ndedition, ISBN 0-471-31423-4, 2000

9. **Engelbrecht Andries P.,** "Computational Intelligence", Book, South Africa, John Wiley & Sons, 2002.

10. **ESRI,** "The guide to geographic information systems", JAN 15, 2008, http://www.GIS.com.

11. **Rumelhart, D.E., G.E. Hinton, J.L. McClelland,(1986).**"A general framework for parallel distributed processing". In: Parallel distributed processing: explorations

in the microstructure of cognition, Cambridge, MA:The MIT Press, 1: 45–76

12. **Skapura James A. Freeman and David M.,** "Neural Networks Algorithms, Applications, and Programming Techniques", Book, Addison-Wesley Publishing Company, ISBN 0-201-51376-5, 1991.

13. **Sharif, M., and R. Wardlaw. (2000),** "Multi-reservoir Systems Optimization Using Genetic Algorithms Case Study," *Journal of Computing in Civil Engineering* 14, no. 4 255-263.

14. **Goldberg, D.E., (2000).** "The Design of Innovation: Lessons from Genetic Algorithms, Lessons for the Real World," Technological Forecasting and Social Change 64, no. 1 (2000): 7-12.

15. **Gimblett,R.H., Ball, G.L., Guise. A.W, (1994).**"Autonomous-rule generation and assessment for complex spatial modeling". Landscape and Urban Planning. 30, 1.3-26.

16. **Yang, M.D., and T.C. Su, (2007).** "An Optimization Model of Sewage Rehabilitation," Journal of the Chinese Institute of Engineers 30, no. 4 (2007): 651-659.

17. **Roberts, F.L., Kandhal Lee, D. and Kennedy, T.W.(1996).**"Hot mix Asphalt Materials, Mixture Design, and Construction" .NAPA Education Foundation, Lanham, MD.pp.529.

18. Ertl, T. and Haberl, R. (2006), Editorial. Proceedings 2nd International IWA Conference on Sewer Operation and Maintenance. Editors: Ertl, T., Pressl, A., Kretschmer, F. and Haberl, R. SOM 06, Vienna, Austria.

19. **Read, G.F., and Vickridge,I.G. (1997).**" Sewer-Rehabilitation and New Construction.ARNOLD, New York.Toronto.

20. **Caldas, L.G., and L.K.Norford, (2002).**" A Design Optimization Tool Based on a Genetic Algorithm," Automation in Construction 11, no.2 (2002):173-184.

21. Bentley System (2012). (http://www.bentley.com) date of access 15/03/2012.

22. **Shamsi U. M., (Jan, 2005).**"GIS Tools for Water, Wastewater, and Stormwater Systems" CRC Press, Florida

23. (http://www.haestad.com) date of access 29/03/2008

24. **Haestad Methods. Inc, Thomas M.Walski, Thomas E.Barnard, Eric Harold,La Vere B.Merritt,Noah Walker,Brain E.Whitman (2004).**"Waste Collection System Modeling and Design", 1st ed., Heastad Press, Waterbury, CT USA

25. **Holand,J.,(Jan, 1999).**" Program of O&M Support to GOSD – Shamal Drainage

Area Plan", Final Report for DFID-GOSD in accordance with the terms and conditions of Consultancy Services, Hyder Consulting Limited

26. **Chipperfield, A., Fleming, P., Pohlheim, H., and C. Fonseca. (2004).**" *Genetic Algorithm Toolbox User's Guide*, Version 1.2, Department of Automatic Control and Systems Engineering, University of Sheffield, Sheffield, UK, 2004