Optimizing the operation of the Haifa-A water-distribution network

Elad Salomons, Alexander Goryashko, Uri Shamir, Zhengfu Rao and Stefano Alvisi

ABSTRACT

Haifa-A is the first of two case studies relating to the POWADIMA research project. It comprises about 20% of the city's water-distribution network and serves a population of some 60,000 from two sources. The hydraulic simulation model of the network has 126 pipes, 112 nodes, 9 storage tanks, 1 operating valve and 17 pumps in 5 discrete pumping stations. The complex energy tariff structure changes with hours of the day and days of the year. For a dynamically rolling operational horizon of 24 h ahead, the real-time, near-optimal control strategy is calculated by a software package that combines a genetic algorithm (GA) optimizer with an artificial neural network (ANN) predictor, the latter having replaced a conventional hydraulic simulation model to achieve the computational efficiency required for real-time use. This paper describes the Haifa-A hydraulic network, the ANN predictor, the GA optimizer and the demand- forecasting model that were used. Thereafter, it presents and analyses the results obtained for a full (simulated) year of operation in which an energy cost saving of some 25% was achieved in comparison to the corresponding cost of current practice. Conclusions are drawn regarding the achievement of aims and future prospects.

Key words | artificial neural network, genetic algorithm, optimal-control, POWADIMA, water distribution

Elad Salomons (corresponding author)
Alexander Goryashko
Uri Shamir
Grand Water Research Institute, Technion – Israel Institute of Technology,
Technion City Haifa 3200,
Israel
Tel.:+972 4 829 2239
Fax: +972 4 822 4246
E-mail: *selad@optiwater.com*

Zhengfu Rao
School of Civil Engineering and Geosciences,
University of Newcastle,
Newcastle upon Tyne NE1 7RU,
UK

Stefano Alvisi
Dipartimento di Ingegneria,
Università degli Studi di Ferrara,
Ferrara44100,
Italy

INTRODUCTION

Purpose of case study

Having shown that it was possible to formulate a dynamic, near-optimal control process for a small, hypothetical water-distribution network (Rao & Salomons 2007), the next major challenge was to apply the methodology developed to a real network. This would not only provide experience of scale-related issues but also give exposure to some of the idiosyncrasies in network design that are found in practice. With this in mind, a number of urban areas were considered as possible candidates but the geographic convenience of Haifa in northern Israel and an existing relationship with the Municipal Department of Water, Sewage and Drainage provided compelling reasons to choose Haifa for the first of two case studies.

doi: 10.2166/hydro.2006.017

Aims

In applying the control system developed, the initial aim was to quantify the potential operational cost saving that could result, in comparison with current practice. To that end, it would be necessary to run the control system for an extended period of time since it was likely that any cost savings would be a function of the demands which, in turn, vary with the seasons. Therefore, the minimum simulated period would have to be one year, simulation being necessary as, had detailed records even existed, they would have reflected current practice, not near-optimal control. Having determined the potential saving in operational costs, the second aim was to evaluate the operational performance of the control system in terms of service to customers and compliance with operational constraints. Whilst the former comprises the reliability/continuity of supply and delivery pressures, the latter include ensuring that storage tank levels are within the normal operating range and tanks are refilled to a prescribed storage level at a fixed time each morning.

THE HAIFA-A NETWORK

Delineation of the network for the case study

In order to gain some experience of the impact scale has on the control system developed, the original intention was to select two case studies, which had significantly different network sizes. To some extent, this was overtaken by events when the complexity of the two networks selected became apparent, which tended to mask the effects of scale. Nevertheless, only a portion of the Haifa water-distribution network was used for the smaller of the two case studies, which, for want of a better name, is referred to as Haifa-A (Figure 1). The reasons for choosing this particular portion of the overall network include the facts that it is defined uniquely by its two supply points and has negligible influence on other parts of the network as there are no other connections apart from the two supply points. Moreover, most of the required information was readily available since the upgrading of Haifa's Supervisory Control and Data Acquisition (SCADA) facilities has begun in this area. Last but not least, Haifa-A is regarded to be sufficiently challenging for the purposes of rigorously testing the methodology developed.

Description of the Haifa-A water-distribution network

Haifa-A, which comprises about 20% of the overall municipal water-distribution network, is located on the

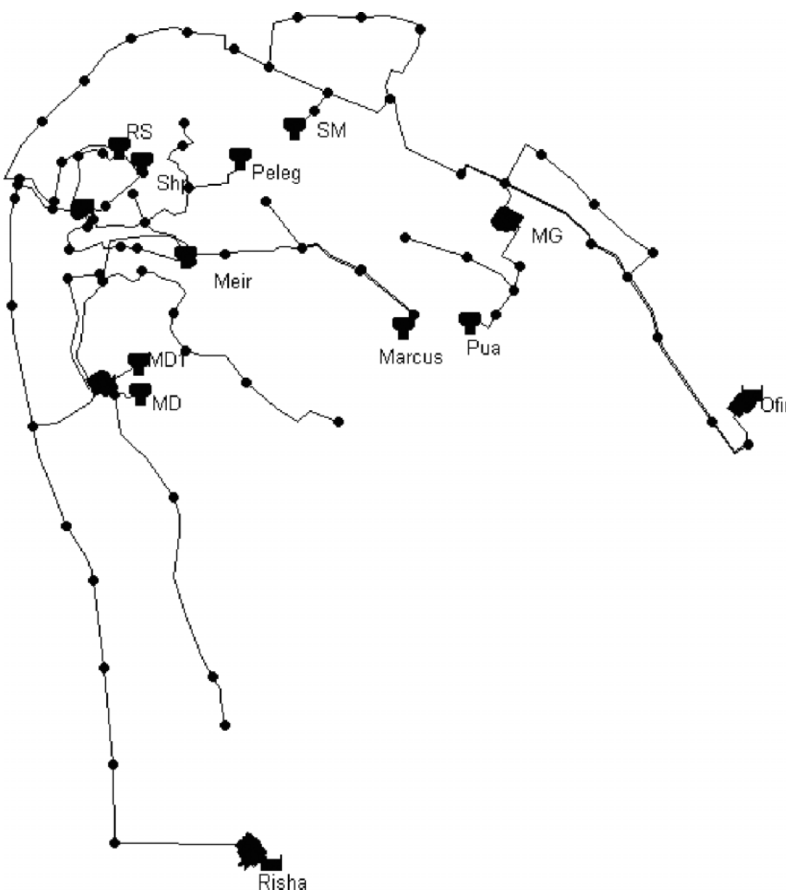

Figure 1 | The Haifa-A water-distribution network.

northwestern tip of Mount Carmel, giving it a pronounced range of topography: the lowest junction is a mere 5 m above mean sea level, whilst the highest has an elevation of 240 m. The network is supplied by two major pumping stations, Vadi-Risha (Risha) in the south and Ofir in the east, which feed the main pipeline that lies at the lower levels. Water is pumped from this pipeline to the higher elevations. Haifa-A serves a population of about 60,000, whose daily average consumption is approximately 13,000 m^3, rising to a peak of some 21,000 m^3/d. The peak hourly demand is estimated to be about 10% of the peak daily demand, that is to say, 2100 m^3/h. The total storage capacity amounts to 12,900 m^3, which represents 61% of the peak daily demand.

Hydraulic simulation model

The Haifa-A network has been modelled using the EPANET hydraulic simulation package (Rossman 2000). The model comprises 126 pipes, 112 nodes, 9 storage tanks and 17 fixed-rate pumps. The pipes range in diameter from 100 mm to 600 mm and have a total length of approximately 41,500 m. The roughness coefficient, which applies to all pipes, is estimated to be 120. Due to the topography, there are 6 separate pressure zones, each supplied by a dedicated set of pumps, which are largely independent one from another. These 6 pressure zones also serve as district metering areas (DMAs), each of which has one or more storage tanks. However, most of the storage capacity (9150 m^3) is located in the lowest zone (DMA1). The remaining 3750 m^3 of storage is distributed between the other 5 DMAs. Of the 17 pumps, which are grouped in 5 pumping stations, 13 are on active duty with the remainder on stand-by. Since the pumps are not identical, each has been modelled individually. Although there are no remote-control valves within the network, there is a valve which closes whenever the Mahane-David (MD) pumping station is activated. However, for the purposes of this exercise, it has been assumed that a pressure-reducing valve (PRV) has been substituted, to make the problem more interesting. A schematic diagram of the network is given in Figure 2.

Current operating regime

At the present time, the network is operated on a local basis, depending on the water levels in the various storage tanks. More specifically, each pumping unit is assigned two water levels for each tank, one at which the pumps are switched on, the other for them to be switched off. Therefore, it can be inferred that energy consumption is not considered to be a high priority, with no special attention being given to the electricity tariff structure. Whilst the present approach is regarded as acceptable from the operators' standpoint

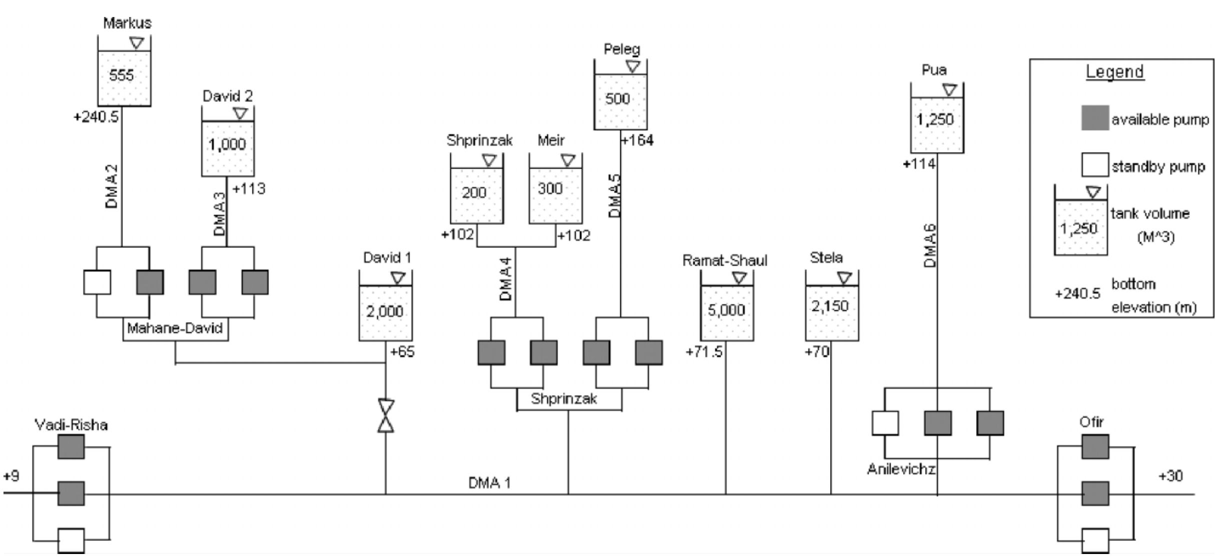

Figure 2 | Schematic representation of the Haifa-A network.

because of its simplicity and robustness, recent changes in business practices have put greater emphasis on energy efficiency. Accordingly, the management of Haifa Municipality's Water, Sewage and Drainage Department not only agreed to be The Technion's subcontractor in this project, but also expressed an interest in the possibility of implementing the results, both of which lend credibility to the exercise from the perspective of the water industry.

FORMULATING THE OPTIMIZATION PROBLEM

Objective function

The basic objective function of real-time, near-optimal control, as applied to water- distribution networks, is to minimize the overall cost of delivering the required amount of water to customers within a given period defined by the time to the operating horizon, subject to maintaining a specified delivery pressure. In this instance, since both sources of supply originate from the National Water Carrier and have the same unit cost of production, minimizing the overall operating cost equates to minimizing the pumping cost. Moreover, a number of additional constraints, such as the requirement that the quantity of water in the larger storage tanks must be at or above a prescribed level at a specified time each morning, have been introduced for reasons that will be explained later.

Electricity tariff structure

One of the first considerations in any optimization problem that involves minimizing energy costs is the form of the electricity tariff structure. In the case of the Haifa-A network, the electricity tariff has a daily pattern with three discrete periods, each having a different charge. Both the duration of these periods and the charges incurred change with the seasons as well as weekends and holidays. Given the complexity of the tariff structure, the only practical solution was to include the charging regime as an hourly look-up table, which is accessed at each update of the operating strategy. All pumping stations were subject to the same electricity tariff except for Anilevich (MG), which has the same fixed rate at all times as a result of a local agreement with the electricity company.

Operating horizon

Urban water-distribution networks generally operate on a daily cycle, the pattern of which may vary in shape for different days of the week and seasons of the year. For this reason, the typical operating horizon is usually 24 h. In some networks where the provision for storage is large, it may be appropriate to consider a longer operating horizon so as to fully exploit the energy tariff structure by continuing to draw on storage over a period of days. However, in the case of the Haifa network in general and Haifa-A in particular, where the storage tanks are comparatively small and can be refilled several times a day if necessary, a 24-h operating horizon is more than adequate and was therefore adopted.

End-point determination

Since the objective function is to minimize pumping costs, the optimization process will draw on storage wherever possible, thereby emptying the storage tanks towards the end of the operating period, unless the end-state is either constrained or assigned a value that mitigates against this likelihood. Failure to impose any restrictions may result in not being able to refill the storage tanks sufficiently for the following day, causing shortages in supply beyond the current operating horizon. There are various ways in which this can be avoided, including:

(i) setting the end-state at the operating horizon for each storage tank to a prescribed water level or constraining it to be above a certain level;
(ii) having a similar constraint not at the end of the operating period but at some intermediate fixed time such as early morning, following the end of the off-peak tariff period;
(iii) giving a positive monetary value to the water levels in the storage tanks at the operating horizon;
(iv) extending the operating horizon to a point where the end state has little impact on operations during the period of concern.

A number of these options were explored to determine their appropriateness for this particular application. Initially, consideration was given to extending the operating period beyond 24 h (option (iv)), with periods of up to 32 h being tested without constraining the end levels in the storage tanks. Whilst this proved to be satisfactory in terms of results, it nevertheless increased the computational burden. The alternative, which was subsequently adopted, required the amount of water in each storage tank to be at or above a prescribed level at a fixed time in the early morning (option (ii)), which is common operational practice to reduce the risk of shortages later in the day. Bearing in mind that the operating horizon and the fixed time will only coincide once every 24 h, this meant that for most of the time there was still a 'loose end' to the operating strategy which would tend to draw on storage rather than start refilling the storage tanks as it approached the next fixed time. Whilst this was not particularly important in the case of the Haifa-A network where the storage tanks could be quickly refilled, it could pose a problem where the storage tanks were larger and/or the pumps smaller. Therefore, the practice of using the water level in each storage tank 24 h previously, as a guide for temporarily anchoring the loose end (a variation of option(i)), was generally adopted, thereby preventing the tanks from emptying as the operating horizon approached the next fixed time.

Time-step

Given that the near-optimal control settings have to be calculated not only for the current situation but also at each time-step up to the operating horizon, the choice of the time-step to be adopted has a profound impact on the computational burden. At the time the decision was made, it was not known what this computational burden would be, nor the impact on the time it would take to compute the operational strategy for the next 24 h. Therefore, a fairly conservative time-step of 1 h was selected as being a compromise between what was desirable and that which could be realistically accommodated. Possible ways of reducing the time-step to make the overall control system more responsive are discussed in the Epilogue of this special edition.

Operational constraints

As a consequence of the stated standards of service, the physical limitations of the network and the requirements of the operational staff, the following constraints have been included within the overall specification of the control system:

(i) a minimum of 25 m water pressure has to be maintained at all demand nodes;
(ii) each storage tank has been assigned a maximum and minimum water level, defining the normal operating range, which is smaller than the actual size of the storage tank;
(iii) the water level in each storage tank has to be at or above a prescribed value at a fixed time in the early morning;
(iv) since the capacity of the electrical connection to some pumping stations is less than that required to operate all of the pumps installed, each pumping station has a maximum limit to its power consumption.

With regard to maintaining a minimum of 25 m water pressure, it was found from the simulation runs that one particular pressure node was invariably lower than the rest and, if the pressure at this node was kept above the minimum, all other pressure constraints within the network are satisfied. As for assigning a minimum operational water level, operational staff wish to ensure there is always a small amount of storage available for firefighting and other emergencies. Similarly, to absorb the effects of communication delays and errors in the SCADA measurements, they also require a maximum operational storage level to prevent the possibility of over-topping. How these values have been determined is not explicit: the values given by the operators have simply been accepted (an evaluation of the trade-off between these safety margins and the additional savings in energy costs could form a separate study). Again, for reasons of supply reliability, the operators also wish to have the storage tanks almost full in the early morning, at the end of the low-tariff period and before the demand increases. This constraint was introduced for 6 out of the 9 storage tanks: it was not imposed on those tanks with a storage capacity of 500 m^3 or less since they do not actually function as storage tanks and can be filled/emptied in an hour or two. It should also be noted that this constraint is not imposed at the same hour throughout the year as the low-tariff period ends at different hours in different months (sometimes 7:00 am, sometimes 8:00 am). Last but not least the energy consumption at some pumping stations is constrained by connection capacity, as in the case of Shprinzak. In these circumstances, the optimization process will seek the best combination of pumps within the maximum power constraint imposed.

APPLICATION OF CONTROL SYSTEM TO HAIFA-A NETWORK

Overview of methodology used

Optimizing the operation of a water-distribution network is a discrete non-linear, non-smooth computational problem, the decision variables comprising each pump's status (on/off) and the setting of the valve, at each time-step up to the operating horizon. Given the current storage levels and demands, the hydraulic relationship between the operating decisions and the resulting pressures/storage levels is extremely complex. Moreover, real-time operation is a dynamic process which automatically 'rolls' forward with each update of the SCADA facilities, incorporating the current state of the network and the revised demand forecasts over the next 24 h. A prerequisite to implementing this concept is an efficient means of calculating the response of the network to different combinations of the decision variables and an effective way of selecting the most appropriate. An outline of the approach adopted can be found in the first paper of this series (Jamieson *et al.* 2007).

In this application, a genetic algorithm (GA) optimizer and an artificial neural network (ANN) predictor have been combined, using a software package referred to as DRAGA-ANN (Dynamic, Real-time, Adaptive Genetic Algorithm – Artificial Neural Network), which was developed as part of the POWADIMA research project (see Rao & Salomons 2007). The approach is based on replicating a conventional hydraulic simulation model by means of an ANN, which is significantly more computationally efficient than using the simulation model directly. Thereafter, the ANN is used in place of the simulation model to determine the feasibility and estimate the cost of each potential solution proposed by the GA optimizer, including any penalties on constraint violations. Besides the consequences of the different control settings, the inputs to the GA optimizer comprise the current state of the network (actual demands and tank storage levels), the operational constraints, the electricity tariff and the demand forecasts up to the operating horizon. After optimizing the control settings for the prevailing situation and each time- step up to the operating horizon, those for the current time-step would be sent via the SCADA facilities for implementation. At this point, the operator has the option of intervening to amend the instructions generated by the optimization process. Then the program waits for the next time-step (in this case 1 h) before scanning the SCADA facilities to establish the revised state of the network and repeating the whole process. In doing so, advantage is taken to 'ground' any discrepancies between the observed values and those predicted at the previous time-step: that is to say, the previously forecast storage-tank water levels at the next time-step are re-set to the measured values at the next scan of the SCADA facilities, so as to minimize any error accumulation.

Developing the ANN predictor

In capturing the domain knowledge of a conventional hydraulic simulation model, the ANN is used as a universal mapping function inasmuch that it relates one multivariate space (the inputs) to another (the outputs). As such, it can be regarded as an input/output model in which a series (layer) of input values (neurons) is connected by arcs to an output layer of neurons via one or more 'hidden' layers, whose functions and weights determine the relationship between the two, even if the data are noisy. In order to construct an ANN for a particular application, its structure and functions need to be postulated before the weights can be estimated by means of a training process, using corresponding sets of input/output vector pairs. Subsequently, separate sets of vectors pairs are used to verify (test) the performance of the ANN in terms of a goodness-of-fit measure. The input–output vector pairs are generated by EPANET for both training and verification. Whilst the numbers of neurons in the input and output layers are fixed by the nature of the application, the number in the hidden layer is to some extent based on trial-and-error, so in that sense, the process is somewhat heuristic. Details of the actual process used to replicate a conventional hydraulic simulation model can be found in the second paper of this special edition (Rao & Alvarruiz 2007).

Decisions relating to how many training and testing sets are required are again subject to opinion and experience, as is the acceptable error in replicating the hydraulic simulation model. However, in formulating the ANN, one always needs to keep in mind the particular purpose for which it was designed, since its real worth in terms of accuracy, robustness and fault tolerance can only be judged in that context. In other words, the absolute accuracy of the ANN is not as important as how it performs in combination with the GA in providing a near-optimal control system for water distribution. Having said that, the accuracy of the ANN has to exceed some minimal threshold, otherwise the control system would almost certainly perform poorly. Experience with the Haifa-A network would suggest that, for real-time control, the root mean square error (*RMSE*) should be less than 3% for the predicted tank storage levels and network pressures.

A number of different ANN structures were tried before adopting the one shown in Figure 3. This has 29 input values (input neurons), 80 neurons in the hidden layer and 15 output values (output neurons). The input values comprise 13 pumps' status (on/off), 9 tank storage levels at the current time (t), 6 current demands (one for each DMA) and 1 valve setting (10–50 m pressure head). The output values are 5 energy-consumption amounts (one for each pumping station), 9 tank storage levels at the next time-step ($t+1$) and 1 pressure node (at the critical point). The ANN was trained with 12,000 input/output vector pairs and tested with an additional 5000 vector pairs. These vector pairs were generated randomly, using the EPANET hydraulic simulation model of the network to determine the consequences of different combinations of starting conditions, control settings and demands. The range of tank storage levels used exceeded the physical dimensions of each tank by 1 m both above and below the actual size, all data being normalized over the range. Normalization and the deliberate possibility of introducing non-feasible solutions proved effective in enhancing the robustness of the GA-ANN process. The *RMSE*s of the normalized data were 0.449 and 0.481% for the training and testing sets respectively, which translates into a discrepancy of some 5 cm, averaged over all storage tanks. An example of a 24-h comparison between the levels given by EPANET and those predicted by the trained ANN for all nine storage tanks is shown in Figure 4.

Development of the GA optimizer

During the past decade, GAs have become increasingly popular in water-distribution management, primarily for the design of networks but more recently in pump scheduling (see Rao & Salomons 2007). A typical GA would normally have three distinct steps viz.:

(i) *initial population generation* in which the GA generates a population of chromosomes (strings), usually at random, each string being a representation of all the decision variables;

(ii) *computation of each string's fitness* where the value of the objective function is evaluated, any infeasible solutions being penalized;

(iii) *generation of a new population* by means of selection, cross-over and mutation, where selection involves choosing strings from the current population according to their fitness values, cross-over is based on combining random portions of two current strings to

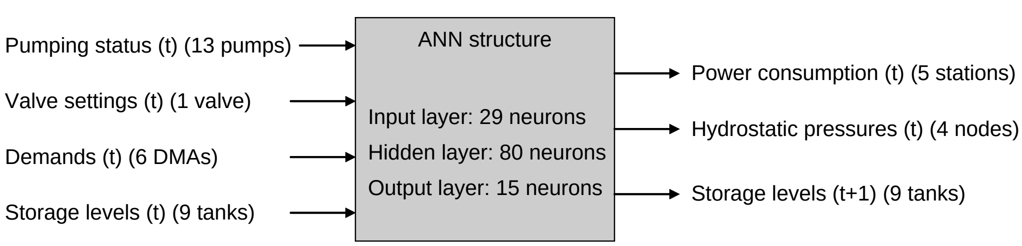

Figure 3 | Structure of the Haifa-A ANN predictor.

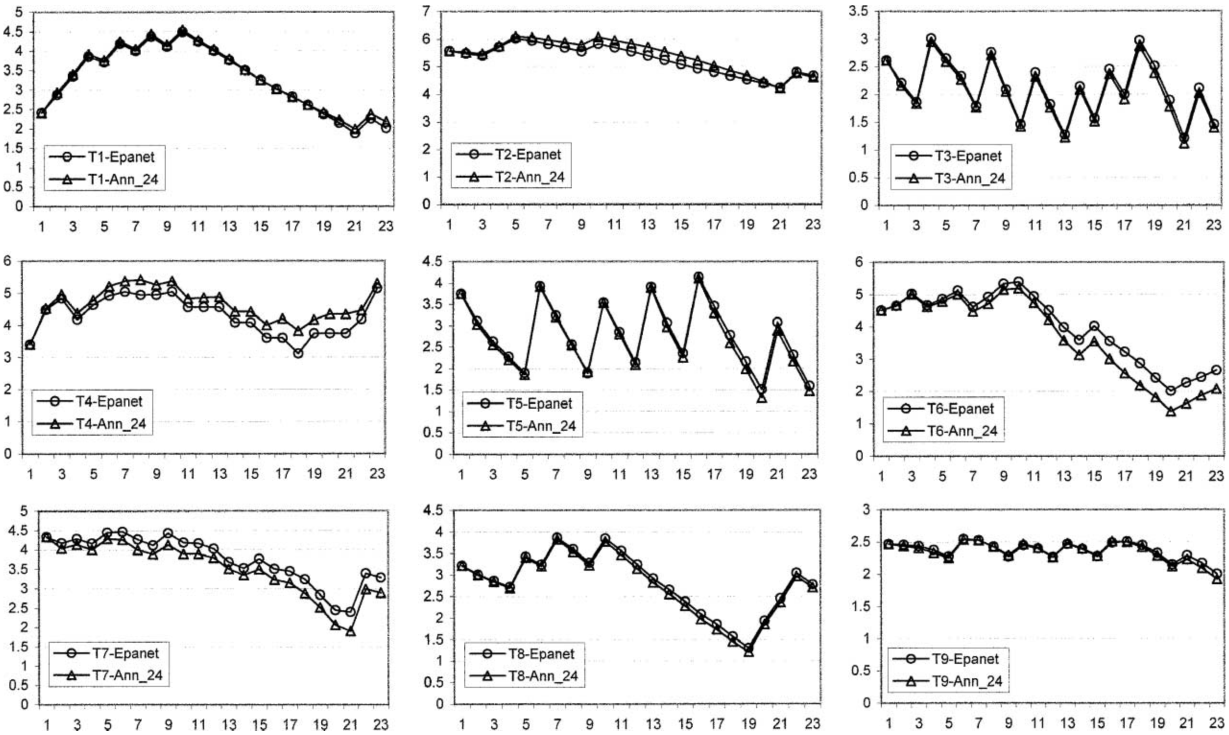

Figure 4 | Comparison of the water levels using EPANET and the ANN predictor.

form two new strings and mutation is the random change of a value in one of the new string's decision variables.

Steps (ii) and (iii) are repeated as the search for the best solution progresses until a stop condition is encountered, which could take the form of a convergence criterion, a maximum number of generations or a maximum run time.

In the case of the GA that has been used in the DRAGA-ANN software package, a number of additional features have been included to improve the computational efficiency and the consistency of convergence, which are major considerations for real-time control. These include using the contemporaneous portion of the previous near-optimal operating strategy as the initial starting conditions for the next update instead of a random set of values and the use of the elitist principle in which the best solution found so far is automatically chosen for the next generation rather than just having a high probability of selection. Other features relate to the modification of the fitness scaling which is applied to adjust the range of fitness values during the search procedure, in order to avoid extreme values wielding too much influence. Another was the introduction of an adaptive penalty function which ensures the penalty coefficients are neither too large or too small at each stage of the optimization process. A more detailed explanation of these modifications can be found in the third paper of this series (Rao & Salomons 2007).

In the case of the Haifa-A network, each string comprised 408 bits, with 1 bit for each pump at each time-step up to the operating horizon ($1 \times 13 \times 24 = 312$) and 4 bits for the PRV at each time-step ($4 \times 1 \times 24 = 96$). Use of a binary code enabled the 4 bits relating to the PRV to represent 16 discrete values within the range of 10–50 m of water pressure. For this particular application, the GA operated with a population size of 50, a cross-over probability of 0.76 and a mutation probability of 0.002. The tournament size for selection was 4 and the total number of generations was 1000.

Combining the GA optimizer with the ANN predictor

Initially, the combined GA-ANN for the Haifa-A network was applied to a series of separate 24-h simulations, using a number of different demand profiles and initial tank storage levels. For the purposes of this exercise, the EPANET model was used as a surrogate for the real network so it was possible to compare the results from GA-ANN with those from GA-EPANET under the same conditions, to ascertain what impact the *RSME* of the ANN had on error accumulation. In particular, it was noticed that these discrepancies were not uniform: for some tanks the deviations were small, for others they were significantly larger. The question then arose how best to address this problem of error accumulation which, to a greater or lesser extent, would always be present. The initial solution that was implemented was to include a 30 cm tolerance zone above the top and below the bottom of the normal operating range for each tank. If the tank storage levels at the end of each 24-h run were within all of the tolerance zones, then the operating strategy was deemed acceptable. However, this was not always the case. Therefore, when it came to developing the dynamic version of the control process (about which more will be said later), the additional practice of using GA-EPANET to confirm the operating strategy derived by GA-ANN was originally adopted. Subsequently, it was found that if the accuracy of the ANN could be improved, as was the case, this practice was unnecessary for normal operations but may well have added value in abnormal situations, as explained later.

The dynamic version of the GA-ANN control process is necessitated by the fact that demands are changing continuously and therefore pumps and valves need to be adjusted at regular intervals (in this case hourly), if optimal control is to be realized or at least approximated. The way in which this is achieved requires the use of the SCADA facilities to establish the existing state of the network by ascertaining tank storage levels, network pressures, status of pumps and valves, etc. Thereafter, any discrepancies between the measured values and those predicted at the previous time-step for the current time are grounded to eliminate error accumulation, especially in the tank storage levels. Using the contemporaneous portion of the previous near-optimal control strategy with a revised end point as starting conditions, the GA-ANN searches for the optimal control settings relating to the following 24 h. In doing so, it takes account of the short-term demand forecasts for each DMA, the electricity tariff structure and the operating constraints. Having determined the new control strategy relating to the following 24 h, the control settings for the current time-step are implemented. Then the control system waits for the next update of the SCADA facilities, before repeating the whole cycle. A simplified schematic representation of this process is shown in Figure 5.

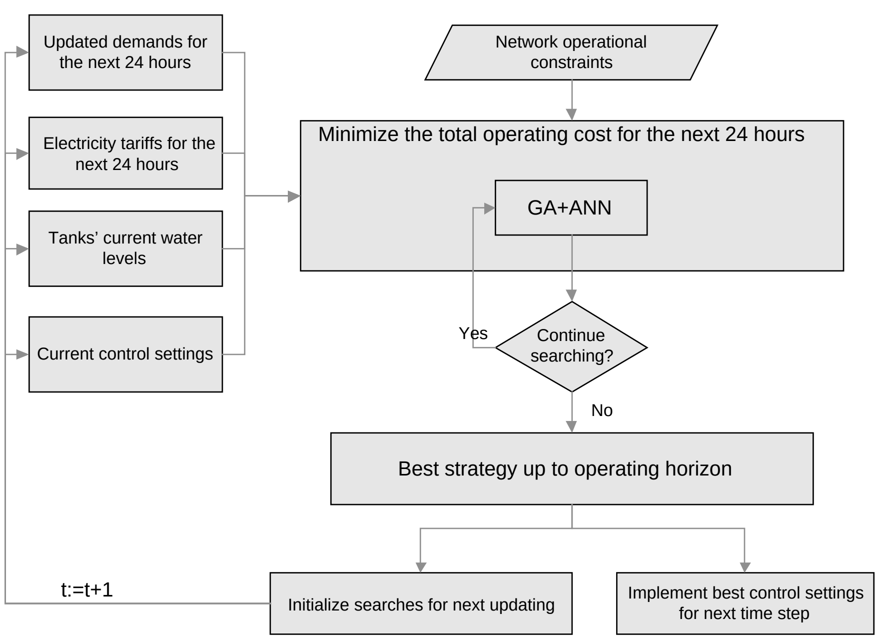

Figure 5 | The DRAGA-ANN control system.

Short-term demand forecasting

Since the near-optimal control strategy covers a period of 24 h ahead, demand forecasts are required of the expected amounts of water needed for consumption, in addition to leakage. It is well known from past experience that demands are extremely variable and therefore these forecasts have a high degree of uncertainty, which is a source of concern in developing a control system for water distribution. Notwithstanding that demand forecasts can be revised with each update of the operating strategy, they have to be realistic in order to exploit the electricity tariff structure in full, without infringing the operational constraints. Therefore, demand forecasts have to be as accurate as possible and the control system robust enough to absorb any unexpected deviations.

As hourly demand information for the Haifa-A network was not available, data from a similar-sized area with the same urban characteristics have been used. These data, which relate to the year 2000, have been scaled so as to produce surrogate hourly demands for the case study. Analysis of these data highlights a marked periodic behaviour in the demand for water. As with many other cities, seasonal, weekly and daily demand patterns can be seen (Figure 6). These patterns form the basis of the demand-forecasting model that has been used, which is fully described by Alvisi et al. (2007), in the fourth paper of this special edition. It suffices here to say that the model comprises two modules, a daily module and an hourly module, each of which has a periodic component reflecting the longer-term effects and a persistence component representing the shorter-term memory of the process. The daily water demand is first forecast on the basis of a seasonal cycle modelled by a Fourier series, which is modified according to the day of the week, thereby taking account of the weekly cycle (the daily periodic components). A daily persistence component, which is modelled using time-series analysis, is then added to account for the residuals. Thereafter, an hourly cycle (the hourly periodic component), depending on the type of day and season, is

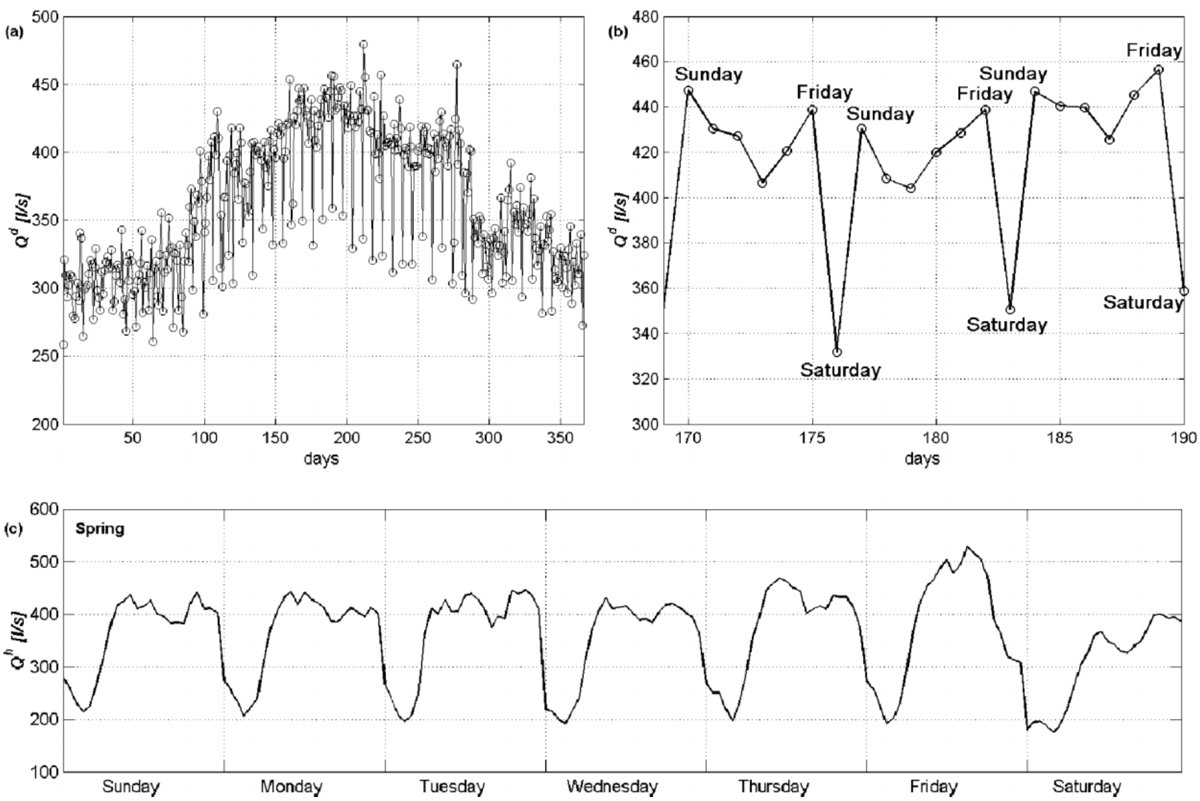

Figure 6 | Daily demands for Haifa-A's DMA1.

superimposed on the daily water-demand forecast. Finally, the hourly persistence component, based on regression, is included to provide hourly demand forecasts which are updated each hour for up to 32 h ahead. Whilst a 32-h horizon is used as the general case, a suitable operating horizon can be selected for the particular application. In the specific case of DMA1, the largest of Haifa-A's district metering areas, the *RMSE* for the 1-h ahead forecast was 36.4 l/s and 43.2 l/s for longer time horizons up to 24 h, with corresponding mean absolute errors of 8.6 and 10.3% (Figure 7).

EVALUATION OF THE CONTROL SYSTEM DEVELOPED

Comparison of energy costs

The optimization package, as applied to the Haifa-A network, was run for the entire year 2000. At each hour, a demand forecast was made, the optimization routine called and an operating strategy derived for the following 24 h. Thereafter, the control settings for the current time-step were implemented on the EPANET hydraulic simulation model (acting as the real Haifa-A network) and the consequences calculated using the observed water demands for that time in order to obtain the 'actual' energy cost incurred and the tank storage levels at the end of the hour. At the next update of the control process, the GA-ANN predicted values of the storage levels for the current time-step were compared with those from the EPANET model (acting as the SCADA facilities) and any discrepancies grounded as they would be in practice, before repeating the whole process. The energy costs that would have been incurred had the control system been in place were then aggregated for each month so that they could be compared with the computed costs for the existing operating regime.

Given that, at the present time, no consideration is given to the energy tariff structure, it is perhaps not surprising that hourly energy costs were not available. Therefore, an EPANET model of the network with the current operating rules embedded was formulated and run for the year 2000, with the same demands and energy tariff structure used in the optimized version. It can be seen from Figure 8 that operating costs vary with the seasons: in

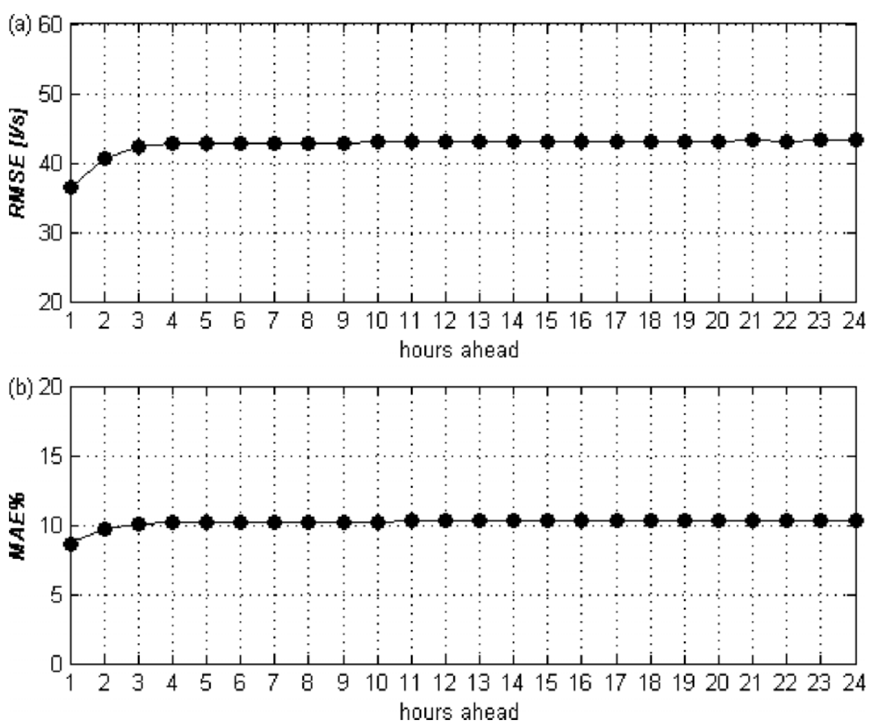

Figure 7 | Errors associated with the 1-h ahead demand forecast.

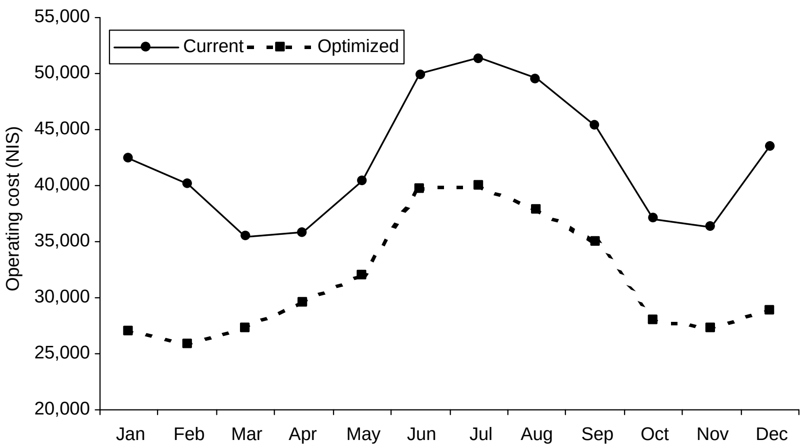

Figure 8 | Comparison of monthly operating costs for current practice and near-optimal control.

summer, when demands are larger, pumps have to be operated during the less attractive energy tariff periods whilst in winter it is possible to restrict pumping to the cheaper tariff rates. As a result, the cost savings that can be achieved during the winter months tend to be higher than the rest of the year. Table 1 indicates that the annual optimized energy cost is 68,265 euros, in comparison with the estimated current-practice energy cost of 91,573 euros, a potential saving of 23,308 euros or 25.4%.

Table 1 | Comparison of monthly operating costs for current practice and near-optimal control

Month	Operating cost for current practice (euro)	Optimized operating cost (euro)
January	7662	4862
February	7234	4655
March	6386	4919
April	6467	5342
May	7308	5763
June	9031	7178
July	9286	7214
August	8937	6822
September	8180	6308
October	6670	5062

Operational performance

Each 24-h operating strategy for every hour during the year 2000 has been checked for any violations of the operational constraints imposed on the network. At no time throughout the year did any DMA fail to receive the required amount of supply. Nor did any storage tank empty or over-top. Similarly, the control system ensured that each tank recovered to its prescribed storage level at the fixed time each morning. Indeed, apart from a few insignificant infringements of the maximum and minimum limits defining the normal operating range, which would probably have been within the measurement error of the SCADA facilities, there were no violations of any kind throughout the entire simulated year of operations. As an example, the water level in the Ramat-Shaul (RS) tank is shown in Figure 9, for the first week in August, together with the minimum/maximum operating limits and the prescribed storage level at 8:00 am.

In addition, the control system developed seems to be remarkably robust inasmuch that it can cope with significant differences between the tank storage levels predicted using the GA-ANN and the subsequent 'observed' levels at the next update of the SCADA facilities. This is probably due to the fact that using the 'observed' levels as initial conditions provides a degree of feedback control. The effect was evident when using earlier versions of the ANN predictor, which had sporadic but comparatively large discrepancies in relation to the EPANET model. These

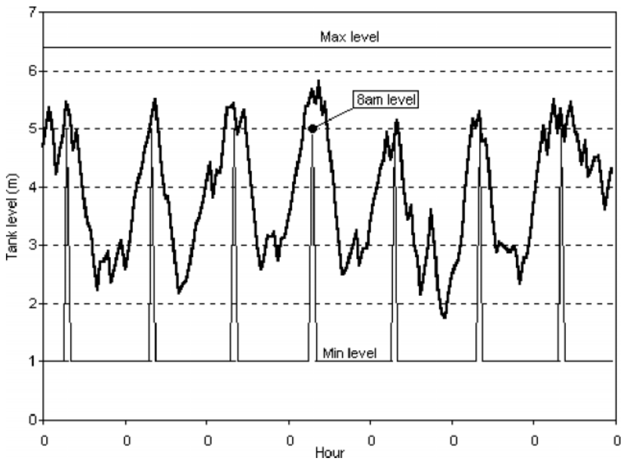

Figure 9 | Compliance with the prescribed 8:00 am water-level constraint in the Ramat-Shaul storage tank.

only caused relatively small violations of the operating limits if the 'actual' storage levels were at or close to the operating margins. Therefore, providing there are minimum and maximum operating limits imposed by the control staff for other reasons, it would seem that the buffer storage they provide would also give adequate protection against the inevitable imperfect modelling of the real network using EPANET, or for that matter, measurement errors in the SCADA facilities.

It is perhaps worth repeating that only the control settings for the current time-step are implemented, as a new operating strategy is generated at the next update of the SCADA facilities. Nevertheless, there is still value in having an optimized, feasible operating strategy for the entire 24-h period, in the event that there is a failure in the SCADA facilities. In order to evaluate this feature, the entire length of each 24-h operating strategy, at each hour throughout the year, was checked against the EPANET model, using the same control settings that were implemented in the GA-ANN process. It was found that in 1234 out of the 8784 hourly optimizations for the year (14%), the water level exceeded the error tolerance zone in at least one of the storage tanks, usually towards the end of the 24-h operating period, as a result of error accumulation. Whilst this may seem high, in practice the frequency would have been less had the SCADA services resumed within the 24-h operating period. Moreover, had the *RSME* of the ANN predictor been larger, GA-EPANET could be used in place of the GA-ANN as the restriction on computing time would no longer apply, thereby eliminating one potential source of error.

SCADA facilities

In order to implement the control system developed, SCADA facilities would be needed for (i) monitoring the current state of the network (pump status, tank storage levels, valve settings, pressure heads, etc.); (ii) relaying these data to the control centre (remote terminal units, communications equipment, repeater stations, licences, etc.); (iii) computing the near-optimal control strategy (computing facilities, etc.) and (iv) remote operation of the control apparatus (switch-gear, mechanized valves, etc.). For the Haifa-A network, the budget price of the SCADA equipment required, including measurement sensors, site commissioning, documentation and training, is in the region of 250,000 euros. Whilst it would be difficult to cost-justify the installation of SCADA facilities for Haifa-A from the energy cost savings alone, there are other benefits arising from remote surveillance including the reduction in manpower costs. Moreover, the Haifa Water, Sewage and Drainage Department has already embarked on upgrading its existing SCADA facilities for water supply. Therefore, the marginal cost of further upgrading the facilities to meet the requirements of the control system developed would be substantially less than the cost quoted.

CONCLUSION

Achievement of aims

The Haifa-A case study has provided a rigorous testing of the water-distribution control system developed under the auspices of the POWADIMA research project. For a variety of reasons, the Haifa-A network cannot be described as typical but, having demonstrated the control system on a somewhat more complicated example than was originally intended only serves to bolster confidence in applying it to more conventional networks. Evaluating the benefits has shown that the near-optimal control operating costs compare favourably with those relating to current operating practice, indicating a potential reduction of about 25% in energy costs. Again, this should not be regarded as typical as the opportunities for cost saving are probably greater in the Haifa area than elsewhere. Nevertheless, it is reasonable to assume that applying the control system to other networks would realize worthwhile cost savings.

The evaluation has also shown improved performance in terms of service to customers and the ability to observe any practical operating constraints that might be imposed. The methodology is both flexible and robust, incorporating a high degree of realism which is imparted by the hydraulic simulation model underpinning the decision mechanism. Scale and complexity do not appear to be insurmountable problems as the domain knowledge of the simulation model can be captured in a far more computationally efficient form. In the case of the Haifa-A network, the efficiency gain is approximately 25 times faster than using the 112-node EPANET model. This, of course, does not mean that the GA-ANN is 25 times faster than GA-EPANET since the GA itself takes a substantial portion of the computing time. However, even for small networks, it does make a significant difference to the run time for calculating each 24-h operating strategy, which in this instance averaged about 4 min on a modern Pentium 4 computer.

Future prospects

The Haifa Water, Sewage and Drainage Department's staff have participated in many aspects of the project, including the provision and verification of the network details, data capture and validation, practical advice and guidance on technical issues, etc. It is expected that this collaboration will continue into the future, with the implementation of the DRAGA-ANN control system for the whole of Haifa, taking advantage of the impending upgrade of the Department's SCADA facilities. At the same time, the opportunity will be taken to introduce further refinements and improvements in the computational procedures, adjusting and augmenting them where necessary in the light of experience gained during the implementation phase.

ACKNOWLEDGEMENTS

This case study forms part of the POWADIMA research project, which was funded by the European Commission under its Vth Framework thematic programme on Energy, Environment and Sustainable Development (Contract Number EVK1-CT-2000-00084). In particular, it comes under Key Action 1 (Sustainable Management and Quality of Water), priority 1.3.1 (Management of Water in the City). The members of the POWADIMA consortium wish to express their thanks to the Commission and its project officers for its support and their encouragement respectively. We are also grateful to the Haifa Water, Sewage and Drainage Department and especially its Director, Yaron Ben-Ari, for the close co-operation. Finally, we wish to acknowledge the valuable technical assistance provided by the following students at The Technion during various phases of the project, namely, Yigal Loyevsky, Sivan Klas and Shirra Resnik.

REFERENCES

Alvisi, S., Franchini, M. & Marinelli, A. 2007 A short-term, pattern-based water demand-forecasting model. *J. Hydroinformatics* **9** (1), 39–50.

Jamieson, D. G., Shamir, U., Martinez, F. & Franchini, M. 2007 Conceptual design of a generic, real-time, near-optimal control system for water distribution networks. *J. Hydroinformatics* **9** (1), 3–14.

Rao, Z. & Alvarruiz, F. 2007 Use of an artificial neural network to capture the domain knowledge of a conventional hydraulic simulation model. *J. Hydroinformatics* **9** (1), 15–24.

Rao, Z. & Salomons, E. 2007 Development of a real-time, near-optimal control process for water-distribution networks. *J. Hydroinformatics* **9** (1), 25–38.

Rossman, L. A. 2000 *EPANET User's Manual*. US EPA, Cincinnati, OH.