Journal of Hydroinformatics | 09.1 | 2007

Optimizing the operation of the Valencia water-distribution network

Fernando Martínez, Vicente Hernández, José Miguel Alonso, Zhengfu Rao and Stefano Alvisi

ABSTRACT

The second of the two case studies in the POWADIMA research project, the Valencia water-distribution network, serves a population of approximately 1.2 million and is supplied by surface water via two treatment plants which have significantly different production costs. The only storage available is located at the treatment plants, each of which has its own pumping station. The management of the network is a complex operation involving 4 pressure zones and 49 operating valves, 10 of which are routinely adjusted. The electricity tariff structure varies with the hour of the day and month of the year. The EPANET hydraulic simulation model of the network has 725 nodes, 10 operating valves, 2 storage tanks and 17 pumps grouped at the two pumping stations. The control system that has been implemented comprises an artificial neural network predictor in place of the EPANET model and a dynamic genetic algorithm to optimize the control settings of pumps and valves up to a 24 h rolling operating horizon, in response to a highly variable demand. The results indicate a potential operational-cost saving of 17.6% over a complete (simulated) year relative to current practice, which easily justifies the cost of implementing the control system developed.

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

Fernando Martínez (corresponding author)
Grupo REDHISP, Departamento de Ingeniería Hidráulica y Medio Ambiente,
Universidad Politécnica de Valencia,
Valencia 46022,
Spain
Tel.: +34 96 387 9610
Fax: +34 96 387 9619
E-mail: *fmartine@hma.upv.es*

Vicente Hernández
José Miguel Alonso
Departamento de Sistemas Informáticos y Computación,
Universidad Politécnica de Valencia,
Valencia 46022,
Spain

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

Stefano Alvisi
Dipartimento di Ingegnería,
Università degli Studi di Ferrara,
Ferrara 44100,
Italy

INTRODUCTION

Second case study

As part of the POWADIMA (Potable Water Distribution Management) research project, the near-optimal control system described earlier in this special edition (Jamieson *et al.* 2007) has been applied to two case studies of different sizes. The Valencia water-distribution network is the larger of the two, the other being the Haifa-A network in northern Israel (Salomons *et al.* 2007). The selection of Valencia was a natural choice inasmuch that the Universidad Politécnica de Valencia (UPVLC) had already modelled the Valencia network at various levels of aggregation, during the past ten years or so (Martínez *et al.* 1999). Moreover, UPVLC already had an established working relationship with Aguas de Valencia, the local water-services company. However, on this occasion, the roles were reversed, with Aguas de Valencia acting as the subcontractor to UPVLC. Amongst other things, this relationship provided access to the required information on the physical details relating to the network, as well as the actual operating costs.

The ostensible reason for having one case study larger than the other was to assess the impact of scale on the degree of difficulty encountered in applying the control system developed. It would also serve to provide a comparison of the computing time required to calculate the near-optimal control strategy at each update, which is a critical issue in determining whether real-time control is a practical proposition. In the event, the complexity of the

doi: 10.2166/hydro.2006.018

Haifa-A network tended to mask the scale effect. Nevertheless, the Valencia network comprises some 1200 km of pipes and supports a population of approximately 1.2 million in Valencia and the surrounding areas, compared with the Haifa-A network's 60,000 population.

Purpose of exercise

The principal reason for applying the control system to the Valencia network was to gain practical experience in its implementation on a real, city-wide basis. In doing so, there would be other more tangible benefits including an estimate of the potential cost-savings arising from implementing the control system and an evaluation of its performance. To achieve these ends, it would be necessary to simulate the near-optimal control process over an extended period of time to assess its performance in relation to the size and variability of demands during different seasons. This implies there would be a need to run the simulation over a period of at least one year in order to compare the optimized operating costs with those actually incurred during the same period, using current operational practices. Evaluating the performance of the control system would be undertaken in terms of the service provided to customers and compliance with the operational constraints imposed.

THE VALENCIA WATER-DISTRIBUTION NETWORK

Physical description

Valencia is located on the eastern seaboard of Spain. The city itself has a population of some 800,000 but the network serves a total population of approximately 1.2 million when the surrounding towns and villages are included. In the past, the whole area was predominantly served from boreholes but the rising concentration of pollutants in groundwater, including nitrates, have forced the authorities to enlarge the existing water-distribution network for Valencia in order to supply treated surface water to the whole of the metropolitan area. As a result, the entire population is now supplied by two large water-treatment plants namely La Presa which is located north-west of the city centre and Picassent to the south-west. Figure 1 shows the general layout of the network.

La Presa, which is the older of the two water-treatment plants, currently has a treatment capacity of 2.2 m^3/s and a storage capacity of 90,000 m^3 at an elevation of 112 m above sea level. In all, it has a total of 10 fixed-rate pumps in 4 distinct groups. Picassent has a treatment capacity of 3.0 m^3/s which, in combination with La Presa, provide sufficient capacity to meet the projected demands for the next 15 years. Picassent's storage capacity is 100,000 m^3, at an altitude of 92 m. The total number of fixed-rate pumps installed is 10, arranged in 2 groups, of which 3 pumps are on standby. The factors that make the Valencia network different from most are, firstly, there is no storage within the overall system apart from the tanks at the two water-treatment plants and, secondly, there are no booster pumps within the network itself. As a consequence, the whole water-distribution operation is largely dependant upon gravity and valves, using the pumps at the treatment plants to pressurize the network as and when necessary.

Current operational practice

The delivery pressures in the network have to be kept within a 25–60 m range, depending upon the particular pressure zone of which there are 4, the lower ones being closer to the coast where the topography is near sea level. Therefore, remote controlled valves have to be used manually to maintain the appropriate pressures, particularly at night when demands are low. Valves are also used to control the flow rates from each water-treatment plant. In general, the treated water coming from Picassent is the cheaper and is therefore preferred whenever possible. Moreover, Picassent's storage capacity acts as balancing tanks in such a way that they are filled during the night when energy is cheaper, with a view to supplying the stored water during the day when energy is more expensive. In order to fill Picassent's storage tanks during the night without increasing the delivery pressure at the point of entry to the city, a valve is closed to isolate the storage tanks from the network, whilst up to 3 pumps supply water directly to the city at low pressure. In this way, most of the network is fed from La Presa during the night and from Picassent in the morning, whilst in the afternoon both plants deliver approximately the same amount of water.

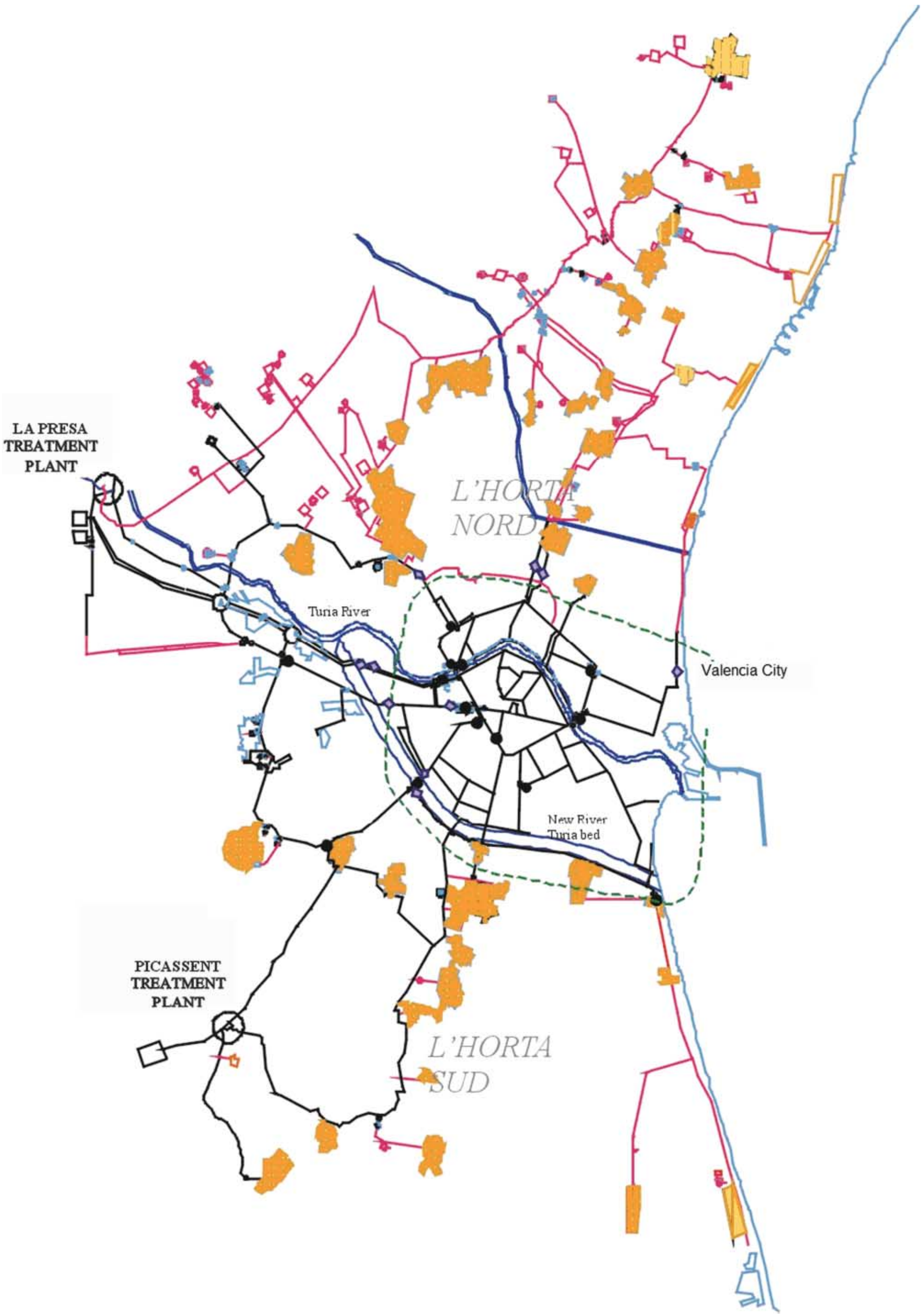

Figure 1 | The Valencia water-distribution network.

Hydraulic simulation model

The Valencia water-distribution network has been previously modelled on several different levels of detail, ranging from 128 nodes to 26,000 nodes. For the purposes of this exercise, the 725-node network has been chosen, which is about six times the size of the Haifa-A model. This corresponds to a 772-pipe network, which is more than adequate for operational control purposes. In keeping with this reduced model, all of the storage tanks at each of the two treatment plants have been aggregated to form one composite storage tank at each plant, again without any significant loss of realism. Similarly, the number of valves has been reduced from 49 to 10, as the majority are set at fixed positions which affects the network hydraulics but not the operational control. The 10 operating valves selected are the most important in terms of controlling flows and pressures at critical points throughout the network. However, all 17 operational pumps have been modelled individually since they have differing characteristics. Moreover, 6 separate district metering areas (DMA) have been included, which coincide with the different demand boundaries.

FORMULATING THE OPTIMIZATION PROBLEM

Objective function

As with most, if not all, water-distribution networks, the operational-control objective function is to satisfy the demands at minimal operating cost, without violating any of the operational constraints. Normally, the operating cost equates to the energy cost incurred in pumping. However, in the case of Valencia, there is an additional complication in the form of two sources of supply with significantly different production costs. Therefore, in this particular instance, the relevant operating costs are the combined energy costs and production costs incurred at each time-step from the present time up to the operating horizon. The operational constraints comprise the standards of customer service, such as, for example, the minimum statutory delivery pressure, in addition to the physical constraints such as the maximum and minimum water levels in storage tanks to prevent overtopping and emptying, respectively.

Operating horizon

Whilst it is possible to envisage a rolling operating horizon of greater than 24 h where the storage available is exceptionally large, the majority of water-distribution networks operate on a 24 h cycle in which the storage tanks are refilled overnight when the charge for electricity is low and then drawn down during the daytime hours when the demands are high. This not only reduces the operating costs but also ensures a turnover of the water in storage, thereby avoiding stagnation. Therefore, a 24 h operating horizon has been adopted, which is the same as current practice.

Time-step

Bearing in mind that the control process has to not only calculate the near-optimal control settings for the current time-step but also each time-step up to the operating horizon, it is immediately apparent that the choice of time-step to be used is of crucial importance to the computational time required. Without knowing in advance what the computational load would be, a fairly conservative time-step of 1 h was adopted as a compromise between what was desirable for real-time control and the ability to complete the computation before the next update.

End-point determination

Since the objective function is to minimize the operating cost and drawing on storage is always cheaper than pumping, there is a danger with optimization that the control strategy may continue to draw on storage until a point is reached where it is impossible to refill the storage tanks before the next daily cycle. This is not normally a problem with a 24 h operating horizon but if, for any reason, that were to be reduced to, say, 12 h, then the possibility of jeopardizing the following day's supplies for a short-term gain in reducing operating costs may occur. Either way, to avoid this possibility, an operational constraint has been imposed on each of the two storage tanks, requiring them to be at or above a prescribed water level at a given time each morning, as currently happens in practice. This time is fixed to coincide with the end of the low, night-time energy-tariff period. In this way, provided there is a refill operating constraint before the operating horizon, the problem is avoided. However, apart from when the operating horizon and the fixed time coincide, there is still a 'loose end' to the operating strategy. Therefore, to eliminate the problem entirely, the loose end is temporarily anchored to the values of the storage tank water levels 24 h earlier, thereby preventing the control strategy from over-emptying the tanks (see Salomons *et al.* (2007) for a more detailed explanation).

Operational constraints

As mentioned previously, the operational constraints comprise a mixture of the physical constraints relating to the distribution network and the standards of service provided to the customers. The former consist of:

(i) limits imposed on the minimum and maximum water levels in each storage tank, defining the normal operating range;
(ii) the prescribed water level in each storage tank whose value has to be achieved or exceeded at a fixed time every morning;
(iii) a maximum limit on power consumption at each of the two pumping stations, equating to the installed capacity.

The normal operating range, which is less than the capacity of the storage tank at each of the two water-treatment plants, was set by the operators to minimize the risk of the tanks emptying or overtopping as a result of communication delays, errors in SCADA (Supervisory Control And Data Acquisition) measurements, etc. Similarly, in order to minimize the risk of supply shortages for whatever reason, the operators required the storage tanks to be almost full in the early morning, which conveniently coincides with the need to prevent the control process from over-emptying the tanks (see the section on end-point determination). As for the limit on power capacity, this was to ensure that, if all the pumps were switched on, their combined consumption did not exceed the installed power capacity. If it did, then the control strategy would select the best combination of pumps within the maximum power constraint imposed.

In this instance, the standards of service to customers consist of:

(i) compliance with the 25–60 m delivery pressure for demands in each of the 6 DMA, depending on the pressure zone in which the DMA is located;
(ii) minimum flow rates of between 300–400 l/s, depending on location.

Whereas the critical points for pressure tend to be sited on the boundaries between pressure zones and towards the peripheries of the network, those relating to flows are generally found near to the pumping stations. The locations of these critical points, together with those for the 10 operating valves, are depicted in Figure 2.

Energy tariff structure

The energy tariff structure that has been negotiated with the local power company follows a fairly typical daily pattern comprising 6 discrete tariff blocks, each having a different rate, which in turn depend on the plant (see Figure 3). Whilst the duration of these blocks generally remains the same throughout the year, the charges incurred change with the month of the year. The only exception is the month of August when there is a flat-rate tariff as a result of the low demand for electricity during the vacation period. Rather than attempting to define the tariff structure as an algorithm, the hourly cost of electricity has been incorporated into the optimization process as a two-dimensional look-up table.

APPLICATION OF CONTROL SYSTEM TO VALENCIA NETWORK

Outline of methodology used

In order to optimize the operational performance of the water-distribution network at the same time as minimizing the energy costs, it is necessary to have some means of predicting the consequences of different pump and valve settings on the behaviour of the network. Hitherto, the only way of achieving that end has been to use a hydraulic simulation model of the network. However, for real-time control, this would be somewhat impractical for a network the size of Valencia, owing to the computational burden simulation imposes, in combination with optimization. Nevertheless, some form of simulation is necessary, since experimenting directly with the real network would not be allowed. Therefore, the approach adopted has been to capture the knowledge base of a conventional hydraulic simulation model of the network, using an artificial neural network (ANN). Thereafter, the ANN predictor is used as part of the control process in preference to the hydraulic simulation model since computationally, it is far more efficient.

Whereas the ANN is employed to predict the impact of different control settings, a genetic algorithm (GA) is used

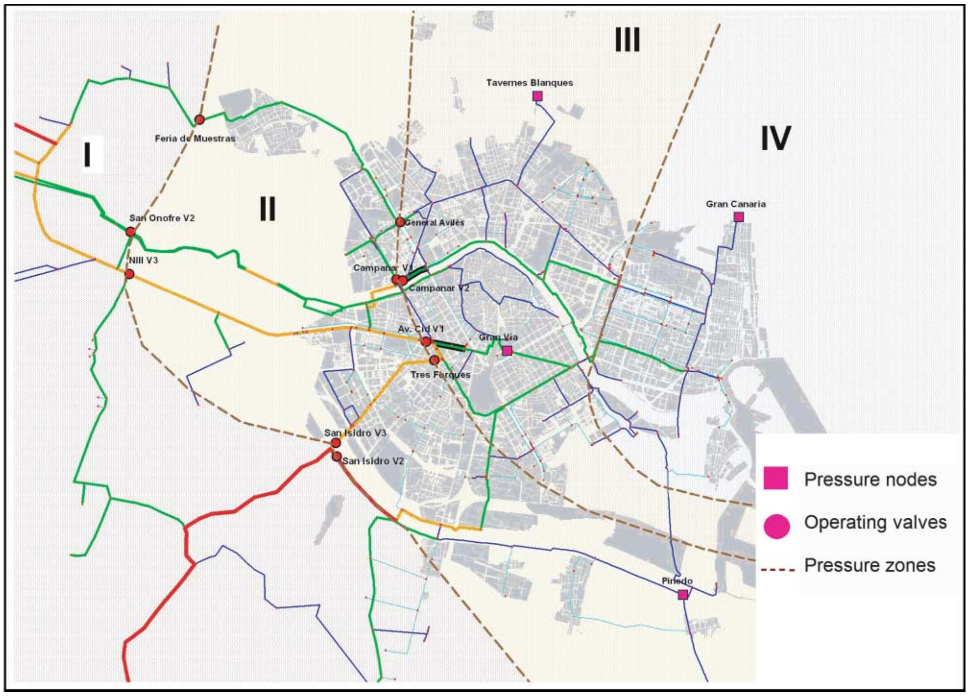

Figure 2 | Location of critical points and operating valves.

to select the 'best' combination not only for the current time but also for each time-step up to the operating horizon, the latter being required to defer pumping until the lower-cost tariff periods, wherever possible. However, only the control settings for the current time-step are implemented. Following the next scan of the SCADA facilities at the new time-step, any discrepancy between the observed water level in each of the two storage tanks and the corresponding value forecast at the previous time-step is 'grounded' so as to minimize error accumulation, prior to calculating a new operating strategy. This whole process is continually repeated on a rolling basis after each new update of the SCADA data, using a software package called DRAGA-ANN (Dynamic Real-time Adaptive Genetic Algorithm – Artificial Neural Network). A much fuller description of the methodology used can be found in the paper by Rao & Salomons (2007).

Developing the ANN predictor

An ANN can be regarded as a universal mapping function, which relates one multivariate space (the input values) to another (the output values). In this particular application, a three-layer, feed-forward ANN has been used to replicate the 725-node EPANET (Rossman 2000) hydraulic simulation model of the Valencia network, in order to provide a means of quickly estimating the consequences of different control settings. This can be achieved by 'training' the ANN with a large number of input/output vector pairs generated off-line using the EPANET model in steady-state mode. The input values comprise different combinations of pump and valve settings, the initial water level in each of the storage tanks, and the full range of demands for the various DMAs. The output values consist of the hydrostatic pressures and flow rates at the critical points in the network, the water level in each storage tank at the end of the next time-step and the power consumption incurred at each pumping station. Subsequently, the performance of the trained ANN is verified using a series of 'testing' sets generated in the same way as the training sets.

Whereas the numbers of neurons in the input and output layers are determined by the numbers of input and output values respectively, the number of neurons in the hidden layer is somewhat subjective and usually based on experience, coupled with a degree of experimentation. Similarly, the acceptable error in replicating the hydraulic

	P1	P2	P3	P4	P5	P6
Euro/kWh						
La Presa	0.073	0.065	0.060	0.054	0.051	0.033
Picassent	0.075	0.067	0.062	0.056	0.052	0.034

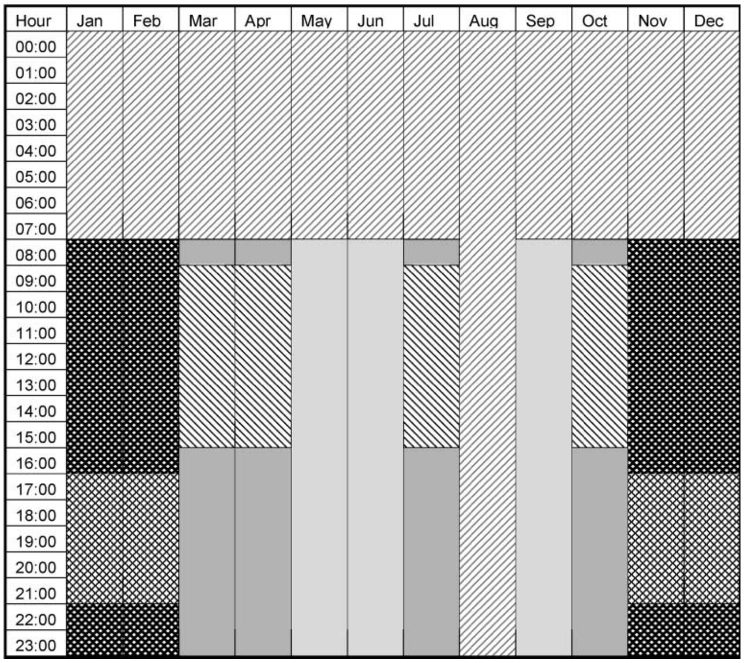

Figure 3 | Energy tariff structure for Valencia.

simulation model can only be determined in the context of the purpose for which it was designed, which again requires some experimentation. However, it is fair to say that, in the case of Valencia, determining the appropriate architecture of the ANN proved to be a fairly protracted and sometimes frustrating procedure based on trial-and-error. Nevertheless, after several attempts, a satisfactory structure was found for the ANN, which is shown in Figure 4.

The ANN depicted has 24 neurons in the input layer, 100 in the hidden layer and 15 in the output layer. The input values comprise the operational status of the 6 groups of pumps (off/on), the settings of the 10 selected remote-controlled valves, the demands of the 6 DMAs and the initial water levels of the 2 storage tanks. The output values are the power consumption of the 6 groups of pumps, the pressures at the 4 critical points, the flow rates at the 3 critical points and the resulting water levels at the 2 storage tanks. The number of vector pairs used for the training stage amounted to some 2500, which were generated randomly using the EPANET model.

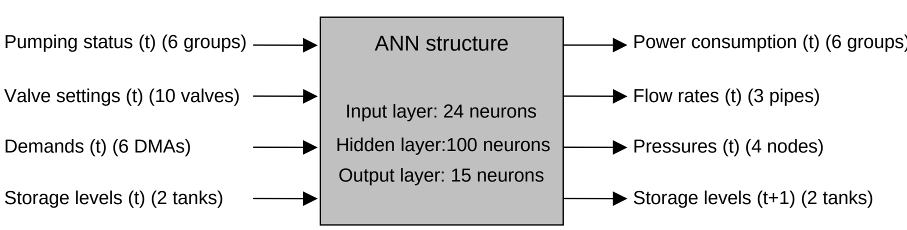

Figure 4 | Structure of the Valencia ANN predictor.

A further 800 were used in the testing process. The range of water levels used in training and testing deliberately exceeded the physical dimensions of the storage tanks by 2 m in both directions in order to introduce some infeasible solutions. Additionally, the full range of possible valve openings was used rather than just the usual operating range. Similarly, the values of the demands incorporated ranged from zero to more than the current requirements, thereby providing an allowance for growth. The resulting root mean squared error (*RMSE*) of the normalized data was 1.20 and 1.30% for the training and testing sets, respectively. A comparison of the water-level forecasts made using the EPANET model and the trained ANN is given in Figure 5.

Developing the GA optimizer

GAs form a class of stochastic optimization techniques analogous to Darwin's theory of evolution in which the better (fitter) solutions at any one time not only survive but lead to further improvements. The mechanics by which this is achieved are loosely based on the biological (genetic) processes of selection, cross-over and mutation. Since there is no parent population at the outset, the original chromosomes (strings), which represent the decision variables, are usually generated at random. The fitness value of each string is then calculated in terms of the objective function, with any infeasible solutions being penalized. Thereafter, strings are selected from the parent

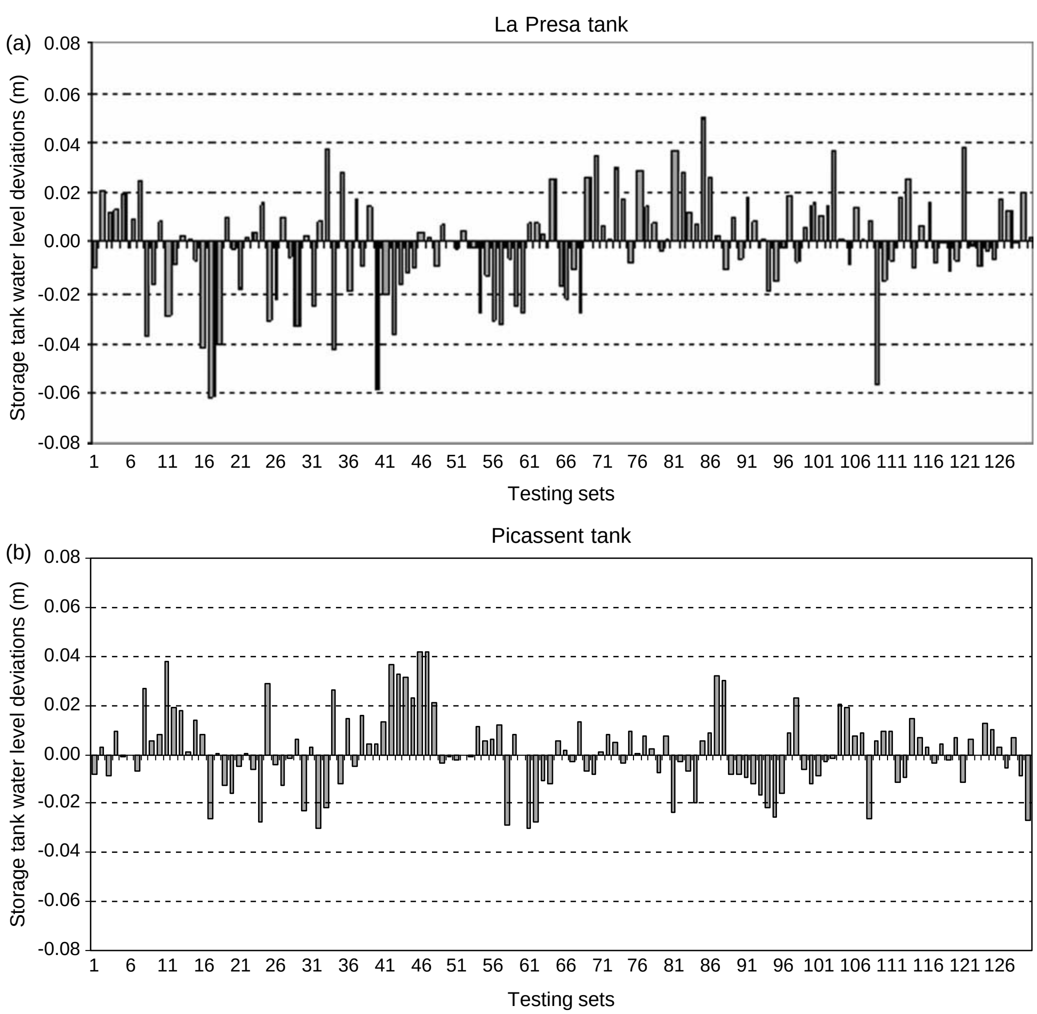

Figure 5 | Comparison of water-level forecasts, using EPANET and the ANN predictor.

population according to their fitness before being paired: the higher the fitness, the more likely they are to be selected. Subsequently, each paired strings share their genes (bits) on a probabilistic basis during the cross-over phase to create two new offspring. This new generation is then included in the population pool and their fitness values evaluated before the selection and cross-over processes are repeated. Occasionally, a mutation is introduced for good measure by altering a bit in the string randomly, to see if any improvement in fitness value results. After a large number of generations, the fittest solution found is taken to be the global optimum. However, bearing in mind the number of generations possible within the computational time available, the global optimum cannot be guaranteed. For this reason, the fittest solution found after a finite number of generations is referred to as the 'near-optimal' solution.

The GA used in DRAGA-ANN has been tailored specifically for real-time control to improve the computational efficiency and consistency of convergence. These enhancements include the use of the elitist principle and modifications to the initialization procedure. The elitist principle that has been incorporated ensures that the fittest solution that has been found so far is automatically chosen for the next generation, rather than having a high probability of selection. As for the initialization procedure, rather than randomly setting the decision variables at each update of the control process, the contemporaneous portion of the previous near-optimal control strategy with the current control settings tagged on the end for the 24th time-step is used on the basis that the control settings do not change radically from one update to the next. In this way, the new near-optimal control strategy for the current time-step can be derived using fewer GA generations than would otherwise be necessary. Other improvements relate to the fitness scaling and the introduction of an adaptive penalty function. Details of these modifications can be found in the paper by Rao & Salomon (2007).

In this instance where the operating horizon is 24 h and the time step is 1 h, each string consisted of 1368 bits, with 1 bit for each of the 17 pumps(on/off) at each time-step up to the operating horizon ($1 \times 17 \times 24 = 408$) and 4 bits for each of the 10 operating valves at each time-step up to the operating horizon ($4 \times 10 \times 24 = 960$). The adoption of 4 binary bits gives a choice of 16 different settings for each valve. Based on more than a hundred test runs, a robust combination of GA parameters was found to be a population size of 50, a cross-over probability of 0.765 and a mutation probability of 0.002. The tournament size for selection was 4, with the number of generations for each new operating strategy being 2000.

Combining the GA optimizer with the ANN predictor

As with the Haifa-A case study, initially the combined GA-ANN control process for the Valencia network was applied to a series of separate 24 h simulations, using different demand profiles and starting conditions. The purpose of this exercise was not only to gain experience in applying the control process but also to examine the impact of error accumulation. Since it was not possible to use the real network to determine the outcome of the GA-ANN control settings, the comparison had to be with GA-EPANET. This of course assumes the EPANET model provides a perfect match to reality, which obviously is not always the case. However, the comparison would provide some indication of the impact the ANN predictor's *RSME* had on the control process.

Early attempts of using the dynamic version of the control process for longer-term runs were not particularly successful. This manifested itself with the ANN-based control strategy, taking additional water from the more expensive source than was necessary in practice, thereby increasing the operating costs rather than reducing them. At the time, the reasons for this were not known but they were thought to be a consequence of the implicit hydraulic constraints within the EPANET model, which were preventing the optimization process from taking more water from the cheaper source. Suspicion subsequently fell on the complex system of valves, which controlled most of the flows and pressures in the network. Therefore, another ANN predictor was trained using extra training sets that covered a broader combination of pump and valve settings. Although the results of re-running the GA optimizer with the revised ANN predictor were better than previously, they were still not entirely acceptable. Eventually, a concerted effort was made to understand the complex interaction between the various valves and include their detailed behaviour within the EPANET model, prior to training what was subsequently adopted as a satisfactory ANN for control purposes. The latter is the one described in the earlier section on developing the ANN predictor.

Having resolved the difficulties encountered, it was possible to finalize the dynamic version of the control system for the Valencia network, using the DRAGA-ANN software package. For real-time control, it is necessary to adjust the control settings frequently to accommodate the short-term fluctuations in demands. In this case, the assumption is that these updates would be hourly, following a scan of the SCADA facilities to define the current state of the network. Any discrepancies between the observed water levels in the storage tanks and those predicted at the previous time-step for the present time are grounded by setting the latter to the same values as the former. However, it is not necessary to ground any discrepancies between the corresponding predicted and observed values of the demands since they are automatically compensated for in the demand forecasting procedure. Rather than randomly initializing the decision variables at each update, the contemporaneous portion of the previous near-optimal control strategy is used for the first 23 time-steps, with the current control settings being appended for the 24th time-step. In this way, the control process is rolled forward by one time step. Subsequently, the GA-ANN searches for the near-optimal control settings for the following 24 h, taking account of the energy tariff structure, operational constraints as well as the revised demand forecast for each DMA. Having determined the near-optimal control strategy for the following 24 h, the control settings for the present time are implemented via the SCADA facilities, prior to waiting for the next update when the whole process is repeated. Figure 6 shows a simple schematic of the DRAGA-ANN control system.

Short-term demand forecasting

DRAGA-ANN is a feed-forward control system inasmuch that the control settings implemented for the current time-step not only take into account the current demands but also anticipate the future demands over the next 24 h. Therefore, it is necessary to incorporate some means of forecasting future demands. To that end, the demand-forecasting process developed by Alvisi *et al.* (2007) has been applied to the Valencia network. In this model, water demand is predicted on the basis of seasonal, weekly and daily periodic components, to which are added the persistence components, which account for the short-term memory. More specifically, the average daily water demand

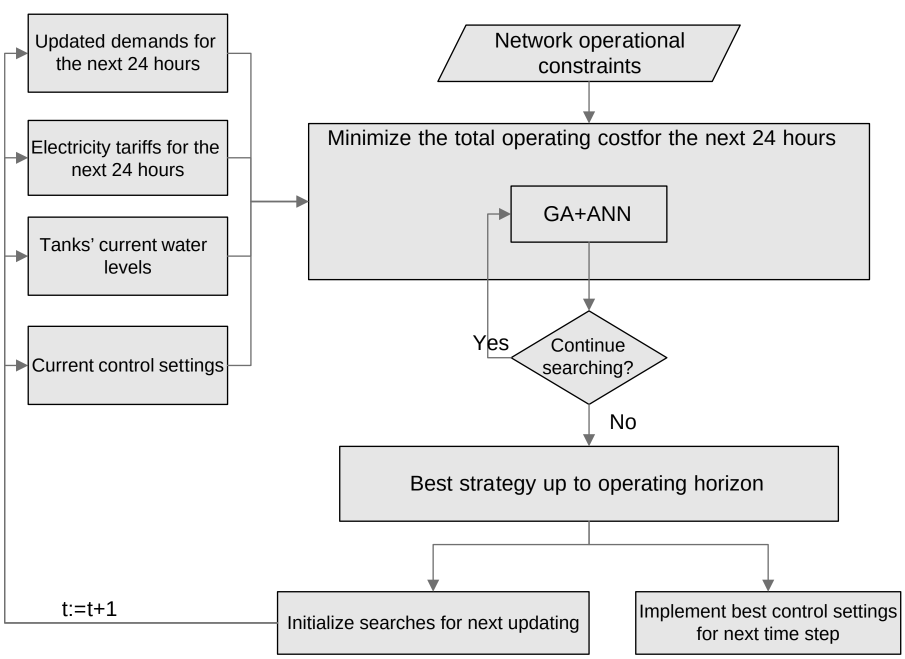

Figure 6 | The DRAGA-ANN control system.

on the day or days covered by the rolling 24 h forecasting window can be predicted by Fourier series for seasonal and weekly patterns, which are subsequently modified according to the day of the week, in addition to a daily persistence component modelled by means of time-series analysis. These daily forecasts are then combined with the appropriate hourly demand pattern depending on the day of the week and season, to which is added a short-term persistence component represented by linear regression, to provide hourly forecasts over the following 24 h period.

As mentioned previously, the Valencia water-distribution network has been divided into some 6 DMAs, which act as the definition of the boundaries for the different demand areas. The largest of these by far is DMA1, which covers the city centre. Others cover the surrounding areas, such as DMA6, which is a nearby seaside resort. The 2001 demand data for all 6 DMAs have been derived by calculating the differences between the inflows and corresponding outflows for each DMA, using the SCADA information provided by Aguas de Valencia. Cursory examination of the resulting demand patterns indicates a wide variation, with those for DMA1 showing a marked reduction in demand during August, whereas those for DMA6 are higher throughout the summer and certain public holidays. Therefore, the demands for all 6 DMAs have been forecast separately. However, for the Valencia distribution-network as a whole, the *RMSE* is 218 l/s for the 1 h ahead forecast and 239 l/s for up to 24 h ahead, with corresponding mean absolute errors (*MAE*) of 4.7% and 5.1%, respectively.

EVALUATING THE PERFORMANCE OF THE CONTROL SYSTEM

Comparison of operating costs

Since there is no way of knowing how the network would have reacted to the near-optimal control settings other than by means of simulation, any comparison with the actual operating costs incurred has of necessity to be on that basis. Therefore, the DRAGA-ANN control process, as applied to the Valencia network, has been run for the full year of 2001. As part of that exercise, the control settings for the current time-step were implemented on the EPANET hydraulic simulation model (acting as a surrogate for the real network) and the consequences calculated using the observed water demands for that time, in order to estimate the operating costs incurred and the resulting water levels in the storage tanks at the end of the current time-step. In updating the operating strategy, the one-step ahead predicted values of the storage-tank water levels were compared with those from the EPANET model (as if they were the SCADA measurements) and any discrepancies grounded as they would be in practice, prior to repeating the whole process. In this way, the operating costs that would have been incurred had the control system been in place could be estimated. As shown in Figure 7, these have been aggregated on a monthly basis and compared with the combined monthly production and energy costs actually incurred for that year, which have been provided by Aguas de Valencia.

It is evident from Figure 7 and the cost comparisons in Table 1 that the potential saving in operating costs vary with the month of the year. The smallest difference between the near-optimal control costs and the current-practice costs occurs during the month of August when a significant portion of the population is on vacation and a flat-rate electricity tariff is in place. Here, the potential savings in percentage terms are only two-thirds of those for the other months, suggesting that improved pump scheduling accounts for about a third of the cost savings whilst the other two-thirds are a result of improved operational performance. The latter includes both pressure control and utilisation of the two sources of supply. Overall, the comparison of the near-optimal control operating costs with those for current practice indicates a possible reduction of 242,000 euros per year, which equates to a saving of some 17.6%.

The potential cost savings due to improved operational performance would have been achieved without compromising the standards of service to the customers. Detailed examination of the simulated output showed that each DMA received the required amount of supply at the required delivery pressure throughout the year. Similarly, at no time during the year did the storage tanks empty or overtop. Moreover, the control system ensured that both storage tanks recovered to their prescribed water levels at the fixed time each morning. What few minor infringements that did arise were confined to the maximum and minimum limits defining the normal operating range, which in all probability would have been within the error of the SCADA measurements.

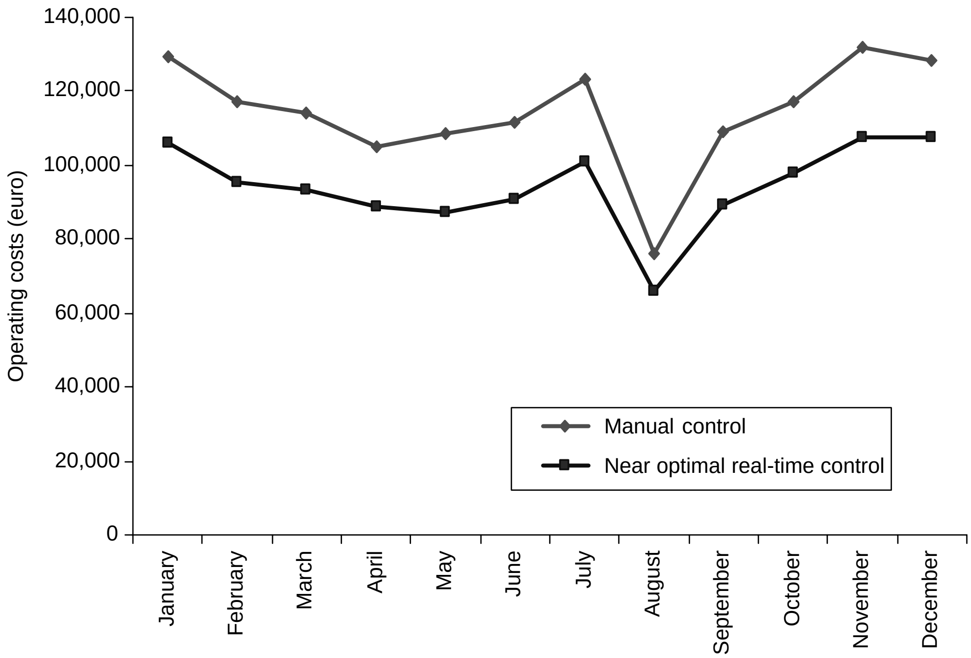

Figure 7 | Comparison of monthly operating costs for current practice and near-optimal control.

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

	Monthly operating costs	
Month	**Current practice**	**Near-optimal real-time control**
January	129,313	105,999
February	116,997	95,404
March	114,277	93,148
April	104,788	88,588
May	108,556	87,111
June	111,774	90,992
July	123,249	100,715
August	75,971	65,820
September	108,998	89,087
October	117,285	97,695
November	131,818	107,661
December	128,356	107,625
Totals (euro)	1371,382	1129,845
Saving (euro)		241,537
Saving %		17.61%

Although an assessment of the impact of errors in the SCADA measurements, representation of the network by the EPANET model, etc., was not part of the formal evaluation, the problems encountered with the early versions of the ANN predictor indicated that the control system seemed surprisingly robust. With regular updating and grounding of errors, it was apparent that the impact of discrepancies was comparatively small and usually confined to infringements of the normal operating range in the storage tanks, which only occurred when the water levels were close to the limits. Obviously, the normal operating ranges could be reduced, thereby increasing the buffer storage for contingencies. However, restricting the operating range also reduces the operational flexibility, which in turn increases the operating costs. Therefore, it becomes a trade-off between reducing the risks and increasing the operating costs, which can only be resolved by the operational staff themselves in their particular circumstances.

SCADA facilities required

SCADA facilities are an obvious pre-requisite to the implementation of the DRAGA-ANN control system: without these facilities, any approximation to real-time control would be impossible. Therefore, an estimate has been made as to the cost of defining the current state of the network and relaying the data to a control centre. These costs include sensors, remote terminal units, telecommunications equipment, repeater stations, licences, computing equipment, switch gear and mechanized valves. For the Valencia network, the estimated cost of the basic SCADA facilities is approximately 325,000 euros. Comparing this cost with the operational savings suggests that it would only take 16 months to recoup the outlay. Moreover, since Valencia already has SCADA facilities installed, the cost of upgrading the existing instrumentation to meet DRAGA-ANN's requirements would be considerably less than the cost quoted.

CONCLUSION

Overall capability of control system

As the second of the two case studies, the Valencia water-distribution network has ably demonstrated that the DRAGA-ANN control system is fully capable of operating on a city-wide basis. Whilst it could be argued that this is an extravagant claim as the network only comprised some 725 nodes, in fact the number of nodes is not particularly important. Indeed, it would have been possible to have used the 26,000-node model of the network but then it becomes a question of whether the marginal improvement in accuracy was worth the increased time and effort required in generating the training/-testing sets. Either way, the structure of the resulting ANN would have been similar. However, the computational gain of using an ANN in place of the very detailed EPANET model would have been even more dramatic. As it was, for the Haifa-A 112-node network, the ANN predictor was 25 times faster than using EPANET, whereas for the Valencia 725-node network, that had increased to about 94.

In practice, the main limitation on the size and complexity of the network that can be accommodated is the time between successive updates of the near-optimal operating strategy since that defines the computational time available to compute the strategy. In this instance, a fairly conservative time-step of 1 h was used. However, with the benefit of hindsight, that could have been reduced as each new operating strategy took less than 10 minutes on average to compute using a modern Pentium 4 computer. Bearing in mind that the GA takes most of the computational time, the extent to which that would be possible depends on the number and type of decision variables, represented by the length of the binary string. However, even with the present computing capability, it would seem entirely feasible to have a 15-min time-step with a 1500-bit string length.

Future intentions

Having been involved with the POWADIMA research project from the outset, Aguas de Valencia is fully aware of the progress to date and are now preparing to implement the DRAGA-ANN control system. With that in mind, the water-services company has appointed UPVLC to advise on the SCADA facilities required to support the implementation phase. Rather than simply accepting what has already been achieved, it is likely that steps will be taken to refine the control system developed to meet the specific requirements of the operating staff, in the knowledge of what is actually possible. This is expected to include the unbundling of the storage capacity at each of the two water-treatment plants and further consideration of the current 1 h time-step, with a view to reducing it if at all possible.

ACKNOWLEDGEMENTS

The Valencia case study formed an integral part of the POWADIMA research project, which was funded by the European Commission under its Vth Framework Research Programme, Key Action 1 (Sustainable Management and Quality of Water), priority 1.3.1 (Management of Water in the City). The contract number was EVK1-CT-2000-00084. The members of the consortium would like to express their thanks to the Commission and its project officers for their support and encouragement. The members are also grateful to Aguas de Valencia for its close co-operation, without which this case study would not have been possible. Finally, we wish to acknowledge the important contributions made by Vincent Bou during various stages of this project.

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.

Martínez, F., Signes, M., Savall, R., Andrés, M., Ponz, R. & Conejos, P. 1999 Construction and use of a dynamic simulation model for the Valencia metropolitan water supply and distribution network. In *Water Industry Systems: Modelling and Optimisation* (ed. D. A. Savic & G. A. Walters), Research Study Press Ltd, Baldock, UK, pp. 155–174.

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.

Salomons, E., Goryashko, A., Shamir, U., Rao, Z. & Alvisi, S. 2007 Optimizing the operation of the Haifa-A water-distribution network. *J. Hydroinformatics* **9** (1), 51–64.