Analysis of uncertainties in water systems using neural networks.

Bogdan Gabrys and Andrzej Bargiela
Department of Computing, The Nottingham Trent University

Introduction

There are two main sources of uncertainty and errors in water distribution system simulations. The first one is associated with the modelling of physical elements and represents static (or slowly changing) inaccuracy of the network model. The second source of uncertainty has a dynamic nature and represents inaccurate predictions of consumptions and measurement noise. While, for instance, the pipes' C factors change gradually over years or decades, consumptions and flows in the network change from minute to minute and are not well characterised as statistical processes (the difficulty of deriving and validating a probability density function for water use at some location at some particular moment).

Model inaccuracy

A considerable amount of research focused on improving the network model accuracy has been accomplished and reported by many researchers in the last two decades. The research has identified that one of the main sources of inaccuracy of network model is the simplified representation of various hydraulic elements. Consequently, the topics of network model calibration and simplification (skeletonization) and their effect on the quality of the results produced with these models have been studied at great length[10,16,17,21]. The conclusion appears to be that if enough effort is put into refining the network models, any practical degree of accuracy can be attained.

Consumption predictions and measurement uncertainty

The uncertainty associated with both the consumption predictions and the measurements has a more dynamic nature and because of that it is considerably more difficult to quantify. Often, when considering water network operations, researchers had tried to avoid tackling the problem by assuming that this data is known exactly. However, it is now recognised that ignoring these uncertainties leads to unrealistic simulation results.

Although the predictions of total water consumptions in large systems can be accurate to several percent of the load[4,9,20], the equivalent predictions at a nodal level, for a small number of consumers, are frequently 50- 100- or more% in error. Attempts to model nodal consumptions by categorising the types of consumptions and combining them to represent the overall nodal load[21,24] have been successful in that they enhanced the understanding of the origins of the uncertainty, but the assessment of the impact of this uncertainty on the performance of water system models has received only little attention in the literature[2,3].

State estimation

The calculation of all flows and pressures in a water distribution system can be accomplished by formulating and solving the mass and/or energy conservation equations as implied by the measurements in the system. Since the flows and pressures are related through the equations describing the hydraulic elements, it is necessary to calculate either flows or pressures but not both. In either case however it is necessary that there are as many independent equations as there are variables that are to be calculated.

Unfortunately, using a minimal set of measurements, while being correct, is not a very practical proposition since the measurements can, and do, become corrupted or unavailable thus preventing the analytical solution to the equations. A state estimation technique[1,18,19] overcomes this difficulty by allowing the processing of additional equations, which are added to the minimal set so as to provide a degree of immunity to meter failures. However, the computational effort associated with the state estimation procedure is considerably higher than the solution of the minimum measurement set. Consequently, a lot of research has been directed towards improving the efficiency and the numerical characteristics of the state estimation procedures through the use of sparsity exploiting techniques and numerically stable factorisation [1,3,18]. An alternative route to the improvement of computational performance has been the use of parallel and distributed computation[14,15]. This approach has enabled state estimation of large-scale systems but has also highlighted the potential problems associated with the coordination of state estimates calculated for individual sub-systems[14,22].

More recently the state estimation problem has been formulated in terms of analog neural networks[6,7,8,11,12,13] which, by exploiting their highly parallel structure, delivers high computational efficiency while optimizing the global state estimation criterion. The neural network approach is thought to combine the efficiency of hierarchical algorithms, implemented on parallel or distributed systems, with known robustness and optimality of the global state estimators.

Water network model

A mathematical model of a water distribution network combines the physical laws governing the system with the pressure-flow equations for each element in the system. The model relates either, the network's nodal pressures or the network's flows to measurement and pseudomeasurement values and is expressed by the following vector equation:

$$\mathbf{z}=g(\mathbf{x})+\omega \qquad (1)$$

where **z** is a measurement vector; g(**x**) are nonlinear functions describing system; **x** is a state vector; ω is an unknown vector that accounts for measurement noise, model errors and disturbances.

Since ω is unknown and non-negligible, finding a solution to equation (1) involves minimization of the discrepancies between the actual measurements and the values calculated from the mathematical model. This can be expressed as:

$$\min_{\boldsymbol{x}} \quad E(\hat{x})=\frac{1}{2}(\boldsymbol{z}-\boldsymbol{g}(\hat{x}))^T\boldsymbol{W}(\boldsymbol{z}-\boldsymbol{g}(\hat{x})) \qquad (2)$$

where $\hat{\boldsymbol{x}}$ is an estimate of the state vector; **E**() is a cost function to be minimized; $\boldsymbol{W} = diag[w_1, w_2, \ldots, w_m]$ is a measurement weight matrix.

The state estimate vector $\hat{\boldsymbol{x}}$, calculated in this way, is found for a specific set of measurements represented by **z**. The question therefore arises: How confident one can be about this state estimate, when it is known that the measurement vector, **z**, is not single valued but can take a whole range of values from the region [**z**-δ**z**, **z**+δ**z**], which indeed, reflects the reality of the (uncertain) knowledge about a large proportion of consumptions in a water system.

In order to account for the input data uncertainty, the following model with unknown-but-bounded errors has been adopted:

$$\mathbf{z}=g(\mathbf{x})+\omega,\ |\omega_i| \le |e_i|,\ i=1,...,m \qquad (3)$$

where $\boldsymbol{e}$ is the vector representing the maximum expected measurement errors.

The solution of (3) can no longer be expressed as a single valued state estimate but as a set of confidence limits[3] (i.e. the upper and lower limits on every variable of the state estimate).

The best known and mathematically the most reliable method of quantifying the uncertainty of the solution to non-linear systems with uncertainties is the Monte Carlo method. The basic idea behind this method is to use the deterministic state estimator repeatedly for a large number of measurement vectors chosen from within the range [**z**-δ**z**, **z**+δ**z**]. Each calculated state estimate is checked against the maximum and minimum values obtained in earlier simulations and the new limits are set as appropriate. In this way the error bounds for state variables are gradually increased and, after many trials, they asymptotically reach their true values.

The accuracy of this method stems from its full recognition of the non-linearity of the water network model but its application is restricted by its computational inefficiency

By linearising the network models and calculating the sensitivities of state variables with respect of individual measurements, we were able to obtain a good approximation to Monte Carlo results while avoiding repeated solution to network equations[3,11]. The solution described in this paper builds on this result and combines it with a very fast, and efficient solver of an overdetermined set of linear equations.

Neural network approach

A simple recurrent neural network has been used as a basis for constructing a system capable of finding the state estimates with corresponding confidence limits.

The three-layer network shown at Fig. 1 plays a central role as a very fast solver of the linearised water network equations

$$\Delta\boldsymbol{z}^{(k)} = \boldsymbol{J}^{(k)}\Delta\hat{\boldsymbol{x}}^{(k)} + \boldsymbol{r} \qquad (4)$$

and the associated optimization problem

$$\min_{\Delta\hat{x}^{(k)}} \quad E(\Delta\hat{\boldsymbol{x}}^{(k)})=\frac{1}{2}(\Delta\boldsymbol{z}^{(k)} - \boldsymbol{J}^{(k)}\Delta\hat{\boldsymbol{x}}^{(k)})^T \boldsymbol{W}(\Delta\boldsymbol{z}^{(k)} - \boldsymbol{J}^{(k)}\Delta\hat{\boldsymbol{x}}^{(k)}) \qquad (5)$$

where $\boldsymbol{J}^{(k)}$ is a Jacobian matrix evaluated at $\hat{\boldsymbol{x}}^{(k)}$; **r** is a vector of residuals (an estimate of ω); k=0,1,... represents a step of the estimation process; $\Delta\boldsymbol{z}^{(k)} = \boldsymbol{z} - \boldsymbol{g}(\hat{\boldsymbol{x}}^{(k)})$; and $\Delta\hat{\boldsymbol{x}}^{(k)}$ represents a correction vector at the *k*-th step of estimation process.

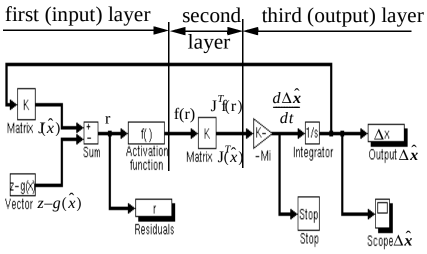

Fig. 1. Recurrent neural network for solving systems of linear equations.

The process of constructing this neural network involves the application of the standard gradient method to (5) and obtaining a set of differential equations on the basis of which the neural network is designed. The details about the construction of the neural network, its VLSI implementation and its performance have been published in[6,7,8,12,13].

Using the neural network as a building block the combined State Estimation / Confidence Limit Analysis system has been constructed as illustrated in Fig. 2. The process of finding a solution by this system consists of two separate stages.

At the first stage of calculations a neural linear equations solver is combined with Newton-Raphson iterations to find a solution to an overdetermined set of nonlinear equations. During this stage, the Jacobian matrix, $\boldsymbol{J}$, is updated and the optimal state estimate vector, $\hat{\boldsymbol{x}}$, is found for a single valued vector of measurements **z**.

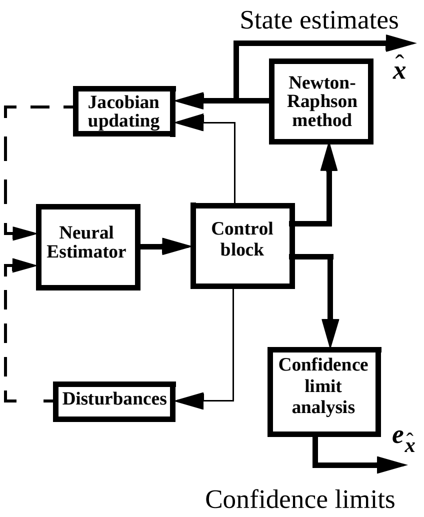

Fig. 2. Neural based system for state estimation and confidence limit analysis.

The second stage of calculations, uses the same NN, but the Jacobian matrix evaluated at $\hat{x}$ remains constant. The confidence limits are evaluated by accumulating the effects of the perturbations of the measurement vector. The result is that, rather than a single deterministic state estimate, a set of all feasible states corresponding to a given level of measurement uncertainty is identified.

The set is presented in the form of upper and lower bounds for the individual variables,

$$\left[\hat{x}^{l},\ \hat{x}^{u}\right] = \left[\hat{x} - e_{\hat{x}},\ \hat{x} + e_{\hat{x}}\right] \qquad (6)$$

and hence provides limits on the potential error of each variable.

An example simulation result for one of the nodes in a 34-node water network is presented in Fig. 1. The iterative nature of the solver is apparent by looking at the lower part of the diagram which indicates the instances of activation of the linear estimator. The upper part of the diagram shows that the selected state variable converges asymptotically during the period 1-2 and, during the period 2-3, the process of calculation of the confidence limits is accomplished. It is clear that the Confidence Limits Analysis technique benefits here from an extreme efficiency of the neural equation solver. The simulated time of the calculations, as indicated in Figure 3., is of the order of microseconds. But, it has to be pointed out that such a fast execution is achievable only if the neural structure was implemented in hardware (e.g. using VLSI or electro-optical technology). The more detailed description of the system from Fig. 2 can be found in[13].

Although the presented results concern water distribution systems the neural solver could be applied to a broad class of nonlinear systems. The only requirement is the differentiability of the nonlinear system model in order to arrive at the linearised model that can be solved by the neural estimator.

Summary

Modelling and simulation of all complex engineering systems is subject to a degree of uncertainty. Every effort should be made to ensure that the mathematical models reflect the physical system operations as accurately as possible. However, it is argued here that some uncertainty is unavoidable and the potential effects of this uncertainty must not be ignored.

Recent research has resulted in a number of methods for the assessment of the influence of uncertainty in mathematical models or measurements. They are referred to in the literature as sensitivity analysis, error propagation, perturbation theory, or confidence limit analysis. A characteristic feature of all these methods is that they are computationally demanding.

This paper makes a contribution to the future development of real-time decision support in water distribution systems by demonstrating that the recurrent neural networks have potential for a dramatic improvement of computational efficiency of state estimation and confidence limit analysis.

It is envisaged that, with the current rate of development in the electronics industry and the emergence of new technologies, an implementation that takes the full advantage of the inherent parallelism of neural networks will be realised in the very near future.

References

1. Bargiela A. (1984). "On-Line Monitoring of Water Distribution Networks", *PhD Thesis*, University of Durham, May 1984.
2. Bargiela A. (1998). "Uncertainty - A Key to Better Understanding of Systems", *Keynote Lecture, Proc. 10th Eur. Simulation Symposium,* ISBN 1-56555-147-8, pp.11-19.
3. Bargiela A. and G.D.Hainsworth. (1989). "Pressure and Flow Uncertainty in Water Systems", *J. of Water Resources Planning and Management*, 115, no 2, pp.212-229.
4. Canu S., R.Sobral, R.Lengelle. (1990). "Formal Neural Networks as an Adaptive Model for Water Demand", *INNC 90 PARIS - INT Neural Network Conference*, vol. 1, pp.131-36.
5. Carpentier, P. and G.Cohen. (1993). "Applied mathematics in water supply network management",

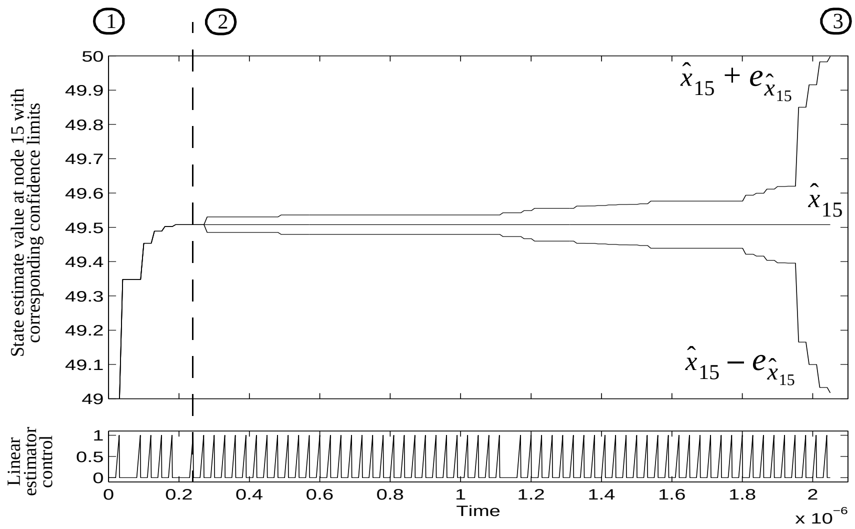

Figure 1: The simulation time diagram for 34-node water network.i) First stage - state estimation;ii) Second stage - calculation of the confidence limits

Automatica, vol.29, no.5, pp.1215-1250.

6. Cichocki, A. and A.Bargiela. (1997). "Neural networks for solving linear inequality systems", *Parallel Computing*, vol.22, no.11, pp.1455-1475
7. Cichocki, A. and R.Unbehauen. (1992). "Neural networks for solving systems of linear equations and related problems", *IEEE Trans. Circuits Syst.*, vol.39 (Feb.), pp.124-138.
8. Cichocki, A. and R.Unbehauen. (1992). "Neural networks for solving systems of linear equations - Part II: Minimax and absolute value problem", *IEEE Trans. Circuits Syst.*, vol. 39 (September), pp.619-633.
9. Cubero R.G., (1991). "Neural Networks for Water Demand Time Series Forecasting", *ANN INT WORKSHOP IWANN 91 PROCS*, pp.453-60.
10. Eggener, C.L. and L.B.Polkowski. (1976). "Network models and the impact of modelling assumptions", *J. of AWWA*, vol.68, no 4, (April).
11. Gabrys B. (1997). *Neural Network Based Decision Support: Modelling and Simulation of Water Distribution Networks*, PhD Thesis, The Nottingham Trent University.
12. Gabrys, B. and A.Bargiela. (1995). "Neural simulation of water systems for efficient state estimation", *Proceedings of Modelling and Simulation Conference ESM'95*, Prague, ISBN1-56555-080-3, pp.775-779.
13. Gabrys, B. and A.Bargiela. (1996). "Integrated neural based system for state estimation and confidence limit analysis in water networks", *Proceedings of European Simulation Symposium*, Genoa, ISBN1-565555-099-4 (Vol.2), pp.398-402.
14. Hartley J.K., Bargiela A., "Parallel simulation of large scale water distribution systems", *Proc. ESM'95*, ISBN1-56555-080-3, pp. 723-727.
15. Hosseinzaman A., "Parallel and distributed simulation of network systems", PhD dissertation, 1995, The Nottingham Trent University.
16. Lansey, K.E. (1988). "A procedure for water distribution network calibration considering multiple loading conditions", *Proc. Int. Symp. on Computer Modelling of Water Distribution Systems*, Kentucky, (May).
17. Ormsbee, L.E. and D.J.Wood. (1986). "Explicit pipe network calibration", *Journal of Water Resources Planning and Management*, ASCE, vol.112, no 2, (April).
18. Schweppe F.C. (1970). "Power system static state estimation", IEEE Transaction PAS, 89, 1, Jan. 19970, pp. 120-135.
19. Sterling, M.J.H. and A.Bargiela. (1984). "Minimum norm state estimation for computer control of water distribution systems", *Proc IEE*, Vol 131, Pt D, No 2, pp.57-63.
20. Sterling, M.J.H. and A.Bargiela. (1985). "Adaptive forecasting of daily water demand", in *Comparative models for electrical load forecasting*, Ed. D.W.Bunn and E.D.Farmer, John Wiley.
21. Suter, J. and C.D.Newsome. (1988). "Network analysis: a user's viewpoint", in *Computer Applications in Water Supply. Volume 1: Systems Analysis and Simulation*, Ed. B.Coulbeck and C.Orr, pp.104-126, Research Studies Press Ltd.
22. Van Cutsem, Th. and M.Ribbens-Pavella. (1983). "Critical survey of hierarchical methods for state estimation of electric power systems", *IEEE Trans. on Power Apparatus and Systems*, vol.PAS-102, no.10, pp.3415-23, (October).
23. Walski, T.M. (1984). *Analysis of water distribution systems*, VanNostrand Reinhold Company.
24. Wright, W.G. and G.J.Cleverly. (1988). "Operational experience of GINAS and WATNET", in *Computer Applications in Water Supply. Volume 1: Systems Analysis and Simulation*, Ed. B.Coulbeck and C.Orr, pp.127-154, Research Studies Press Ltd.