16th Conference on Water Distribution System Analysis, WDSA 2014

Real-Time Model of a Large-Scale Water Distribution System

W. P. Cheng\textsuperscript{a*}, T. C. Yu\textsuperscript{a}, G. Xu\textsuperscript{b}

\textsuperscript{a}College of Civil Engineering and Architecture, Zhejiang University, Hangzhou, Zhejiang, China
\textsuperscript{b}China Guangzhou water supply Corporation Guangzhou, Guangzhou, China

Abstract

This paper presents a real-time WDS hydraulic model combining with field measurements provided by supervisory control and data acquisition systems (SCADA). The system is composed of three parts, namely SCADA, state estimation server and client terminal. For the real-time demand estimation, a weighted least-squares scheme based recursive state estimator and local linear matrix transform algorithm are applied. The WDS model system is applied in Guangzhou city, which is one of the largest cities in China, and proves that the proposed nodal demand correction algorithms are effective for real-time WDS model.

© 2014 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/3.0/). Peer-review under responsibility of the Organizing Committee of WDSA 2014

Keywords: water distribution network; state estimation; real-time; nodal demand.

1. Introduction

A lot of computer software currently is available for the offline WDS hydraulic and water quality analysis, but not specifically for the task of real-time water distribution system management. And most of the water distribution network model analysis is applied for adequate hydraulics under extreme loading cases, such as combinations of peak demands and fire flows, or alternatively tested for adequate water quality parameters, such as water age or the concentration of chlorine and other constituents, under normal operating loads. An off-line model consists of a set of loads or demands on the system, namely an extended period simulation (EPS). The demands in the off-line model are not the result of real-time measurements, but instead consist of educated guesses that can be derived from a number of sources such as typical usage by consumers, customer billing records, and required firefighting loads.

* Corresponding author. Tel.: 086-0571-87951346; fax: 086-0571-87951346.
E-mail address: chengweiping@zju.edu.cn

1877-7058 © 2014 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/3.0/). Peer-review under responsibility of the Organizing Committee of WDSA 2014
doi:10.1016/j.proeng.2014.11.212
Now, SCADA is widely used to provide on-line monitoring data, becoming eyes of the water distribution system. However, the instrumentation employed in real-world networks is highly sparse, often consisting of pressure and flow measurements at pump stations, reservoirs and key nodes only, less than 1/10 of node points. State estimation (SE) is defined as a process of combining the field measurements and mathematical network model so as to gain a global system view and calculate the state variables of interest that are not directly measured; sometime it is named as data assimilation in meteorology. SE has been widely studied since the 1960’s for power systems, and various techniques for SE have been developed and widely used in that power industry. Since the late 1970’s, researchers have tried similar SE algorithms in WDS analysis. The state variables of water distribution system (WDS) are nodal demands, nodal pressures, water quality, and pipe (valve, pump) flow rates. State estimation models of WDS are used to provide more information to fill in the large gaps in data provided by the SCADA system, which can lead to savings in time and money and provide better customer service. The current approach to realize water distribution networks state estimation is a combination of measurement via SCADA system and WDS numerical modeling.

SE is often similar to procedures as those of parameter estimation (PE) that is the process of fitting the model output to the field data so that the model adequately represents the real system. Nodal demands and the pipe roughness are looked as the parameter in PE; sometime, the nodal demand also is defined as the state variable in SE. Pipe roughness is time invariant or varying very slowly. Consumers demands occur along pipes at discrete locations are normally allocated to a demand node at a pipe junction, named as nodal demands. It is not typically directly measurable, regarded as stochastic variables which fluctuate about an estimated mean value. Usually, average daily water consumptions are estimated from population densities or customer billing records for off-line models. Then temporal variations are generated using a diurnal curve for typical consumer types in an extended period simulation (EPS). Although these demand estimates may be appropriate for design purposes, they are not appropriate for real-time use. Nodal demands, among the various input parameters in simulation model, are the most uncertain and variable dynamic state of a WDS. Therefore, demand estimation plays the most important role in the real-time WDS simulation. In some words, accurate knowledge of temporally and spatially variable nodal demand is a prerequisite for pressure and water quality prediction and consequently improves the system operations.

Unlike the power systems, water distribution networks often have a low degree of instrumentation. The application of SE algorithms to WDS is still an ongoing research because of low measurement redundancy, especially for the real-time estimation. In this paper, a real-time WDS state estimation modeling approach is developed by linking field data (from SCADA) with a computer simulation model of the largest water distribution system in China. The monitoring data and water distribution model are assimilated synchronously to give a real-time state estimation and decision support for the system dispatching office.

2. WDS Model and SE Philosophy

Hardy Cross [1] proposed the use of mathematical methods for calculating flows in complex networks. This manual, iterative procedure was an epoch-making advance in the water distribution system calculation and used throughout the water industry for almost 40 years, even extended to the early stage of computer age. The machine implementations of Hardy Cross methodology were developed and were in widespread use by the 1980s [2]. The usability of these models was greatly improved in the 1990s with the introduction of the public domain EPANET model [3] and other Windows-based commercial water distribution system models. Hydraulic models simulated flow and pressures in a distribution system under steady-state conditions where all demands and operations remained constant. However, in practice, the pressure and flow vary over time since nodal demands change over the course of a day or a week. EPS models were developed to simulate distribution system behavior under time-varying demand and operational conditions. It has been applied widely within the water industry and is an integral part of most water system design, master planning, and fire flow analyses.

Parameter calibration studies in WDS were based on trial and error procedures [4, 5]. Later, explicit (analytical or direct) calibration approaches, where the unknown parameters are solved from the same number of equations given from field measurements, were developed [6]. The implicit (automatic or indirect) methods were developed including optimization [7,8] and Gauss-Newton based weighted least-squares (WLS) approaches [9-11]. More recently, some efforts have been attempted to quantify uncertainties associated with estimated parameters and model
predictions. First-order second-moment (FOSM) approximation [12,13] and Bayesian recursive optimization approach [14] have been applied to estimate the uncertainty instead of random sampling approach.

Powell et al. [15] presented three state estimators and compared their ability for the real-time SE in terms of bad data rejection, noise rejection, stability, and elapsed time for convergence to estimate nodal pressures using field measurements of nodal heads and pipe flows. Bargiela and Hainsworth [16] estimated nodal pressure heads and quantified their uncertainty in terms of confidence bounds considering errors in field measurements of nodal pressure and pipe flow. They found that a linearization of the mathematical model of the system provided a good approximation of uncertainty bounds with less computational effort compared to the Monte Carlo simulation (MCS). Carpentier and Cohen [17] presented an optimization technique for demand estimation and leak detection using pipe flow measurements under steady condition. Nagar and Powell [18] proposed an uncertainty approximate method based on linear fractional transformations and semi definite programming approach. They estimated pressure heads and their confidence bounds considering measurement noise as well as parametric uncertainties (in their case, uncertain pipe roughness coefficients). They assumed a set of nodal demands and pressure heads were measured. Work has also progressed in modifying the standard WLS scheme. Sterling and Bargiela [19] proposed an alternative formulation, known as the weighted least absolute values method for nodal head estimation using nodal head and demand measurements. Andersen and Powell [20] presented an implicit SE technique for leak detection for an idealized grid network under steady conditions. Andersen et al. [21] proposed a constrained WLS SE scheme to investigate the effect of introducing measurement bounds. All of these methods assumed that demands were measured or could be estimated as pseudo measurements from population and users demographics. More recently, Davidson and Bouchart [22] presented algorithms for adjusting estimated demands by matching the model solutions with SCADA data by combining heuristics with WLS. Their effort focused on underdetermined systems so no uncertainty analyses were performed. Kumar et al. [23] proposed a SE method using graph-theoretic approach for well instrumented networks. They applied the method to two urban water networks assuming sufficient measurements, such as pipe flow rates, nodal pressures, and demands are available. Kang and Lansey [24] presented A two-step sequential method for dual estimation of demand and roughness coefficient based on a weighted least-squares scheme, and it is applied to two network systems including a midsized real WDS.

3. Real-time State Estimation of Water Distribution System

3.1. WDS hydraulics equation

Steady-state hydraulic relationships in WDS can be described by the nodal flow continuity and pipe head-loss equations. The unknowns in these equations are flows in each pipe, Q, and total energy heads at each node, H. Flow continuity must be satisfied at each junction node

\[ \sum_{j \in f_i} \pm q_{ij} + Q_i = 0 \quad (i = 1, 2, \ldots, M) \]  
\[ q_{ij} = R_{ij} \left( H_i - H_j \right) \]  
\[ R_{ij} = \left( 10.667 c_{ij}^{-1.852} d_{ij}^{-4.87} l_{ij} \right)^{-\alpha} \]

where \( q_{ij} \) denotes pipe flows and \( Q_i \) is nodal demand, and a pipe flow entering a node is given a negative sign. \( f_i \) is the set of pipes supplying to and carrying flow from node \( i \), respectively, \( d_{ij} \) denotes the pipe diameter; \( H_i \) is the nodal head; \( l_{ij} \) is pipe length, \( \alpha = 1/1.852 \); is the Hazen-Williams parameter; \( M \) is total number of nodes.
3.2. Real-time Nodal Demand Estimation Model

In most conditions, pipe roughness is time invariant or varying very slowly. Therefore, the ability of the precise spatial distribution of demands estimation is of primary importance for a real-time WDS model. The WLS method has been widely used for parameter estimation. In the present paper, the algorithm minimizes the sum of square differences between measured and computed values of pipe flows, nodal pressure heads, and the source supplement. Mathematically, the objective function can be written into:

$$ J_{\text{min}} = \sum_{i=1}^{n_h} (W_{hi}^h - h_o^i)^2 + \sum_{j=1}^{n_q} (W_{qj}^q - q_o^j)^2 + \sum_{k=1}^{n_k} (W_{kt}^s - q_o^k)^2 $$

(4)

Subject to

$$ G(H, Q) = 0 $$

(5)

$$ Q_L \leq Q \leq Q_u $$

(6)

where $J_{\text{min}}$ is the objective function to be minimized; $h_o^i$ and $h_o^j$ are measured and predicted pressures at node $i$, respectively. $q_o^j$ and $q_o^j$ are measured and predicted pipe flows at pipe $j$, respectively. $q_o^k$ and $q_o^k$ are measured and predicted source supplement flows at water station $k$, respectively; $W_{hi}^h$, $W_{qj}^q$, and $W_{kt}^s$ are weighted factors applied to the different terms to ensure that they have similar magnitudes and units; $n_h$, $n_q$, and $n_k$ are numbers of observed nodal pressure, pipe flow, and source supplement; $H$ is the vector of nodal head; $Q$ is the vector of nodal demand; $Q_L$ and $Q_u$ are lower and upper nodal demand bounds; $G$ is a nonlinear equation describing the flow and pressure in the water distributions network. The key to solve Equation (4) is to get the optimal search vector. One dimension search method, i.e., golden section search, is applied along with this vector in the optimization process. Based on Cheng’s work (2011) on development of the solution methodology using the local linearization matrix transformation. Fig 1 shows the whole process of real-time water distribution state estimation. The program can run automatically, and once it is initiated, it will continue without any manual operation.

4. Practicing Application in Water distribution of Guangzhou (China)

4.1. Model of Guangzhou Water distribution system

We applied this method to the real water distribution system of Guangzhou city, which is one of largest cities in China. The system supplies 4,000,000 m$^3$/d drinking water for 16,000,000 people, covering 2000 sq. km. It has 7 water sources and 21 pump stations. The system is skeletonized to a hydraulic model, which have 16,889 nodes and 22,451 links (Fig. 2). All pipes with diameters greater than 250 mm are modelled, and total length of pipes is about 2,500 km. 121 pressure meters and 69 flow meters are used to monitor the system state. Fig. 2 shows the model, in which J1~J7 are the 7 water sources, and N1~N8 are critical pressure control points, and F1~F4 are four flow meter points.

4.2. Pump characteristics curve identification

The water distribution system in Guangzhou city has 154 pumps. Some pumps have been in service for many years. The flow-head characteristics curve will vary with the service time. The field measurement is with high cost and time-consuming. At least three month is needed for all pumps test, while the proposed SCADA can provide the pump states, total flow rate and head with much less time. The flow-head curve can be identified as the following model.
\[
\min \sum_{j=1}^{jn} \left( Q_i^j - \sum_{j=1}^{jn} \pi_i^j q(j) \right)^2
\]  

(6)

where \( i \) denotes the time number, \( jn \) is the total number of pumps. \( j \) is the number of pumps. \( \pi_i^j \) is the state of pump \( j \) at time number \( i \). \( \pi_i^j \) is 1 when the pump is open, otherwise, it is zero. \( Q_i^j \) is the measured pump station flow rate. Usually, the pump head-flow characteristic curve can be expressed as \( h = h_0 - sq^2 \). All parameters are solved by the nonlinear conjugate gradient algorithm (NLCG). Two examples are shown in Fig. 3 and Fig. 4. Fig. 3 is the result of source Xintan (J7), which has 11 pumps. Fig. 4 is the result of source Nanzhou (J6), which has 9 pumps.

Fig. 1 Flow chart of real-time WDS simulation

4.3. Platform of Real-time WDS SE in Guangzhou City

The system is composed of three parts. The first part is SCADA, which handles all information from the sensors, including the flow rate, pressure, water level of tank, valve state, pump state, and pump rotate speed. There is more than 2,000 information every 3 minutes. Every fifteen minutes, the information is stored to the MSSQL database. The second part is real-time WDS SE solver server. The hardware of this server is a Dell 710 server, which have two Intel Xeon E5620 CPU and 16G memory. An interface program is developed to get information from the SCADA database. 120 pressure meters and 69 flow meter are used for the data assimilation. At first, all data is validated, and the invalid data out of reasonable range is picked out.
Fig. 2 WDS model of Guangzhou city

Fig. 3 Measured and fitted pumps’ flow of Xintan source (J7)

Fig. 4 Measured and fitted pumps’ flow of Nanzhou source (J6)
Then, the valid information (flow rate, pressure, water level of tank, valve state, pump state, and pump rotate speed) is assimilated using the above optimization procedure every 15 minutes. All nodal pressure and pipe flow are solved and stored into the database. The system can store a week’s result, more than 400 million records. The third part is the client terminal. System controllers can browse the pressure distribution, flow rate, etc. They also can simulate system state by opening or closing pump, analyse the water distribution system response, help system controller to make a decision. It is a real-time decision supporter.

![Platform of Real-time WDS SE](image)

**Fig.5 Platform of Real-time WDS SE**

### 4.4. Result of Real-time WDS SE

The SCADA system has 120 water head meter and 69 flow meters (including the meter in the water station). The real-time solver adjusts the nodal demand according to this information every fifteen minutes. The nodal demand calibration needs about 4 minutes each time (including obtaining information from SCADA). It spends about 1 minute to update the simulated result (nodal head, nodal demand and pipe flow rate) to the database. The flowchart can be found in Fig. 1. Usually, the off-line state estimator is set up with about 75% of the signals and checking to see if it can match the other 25% of signals. The method is applied to test how it works in areas that were not used. The effectiveness of this algorithm has been discussed in details in [25].

The WDS controller may pay more attention on the accuracy of source supplement, for they must determinate production of water plant. The performance of the real-time SE is compared by using the measured nodal pressure, pipe flow and the supplement of each source. Fig.6 presents the estimated and measured flow rates of seven sources. 90% relative errors of source supplements is less than 10% except the source J1. The 80% relative error of J1 is less than 15%. It is because that supplements proportion of source J1 is too small, which is 3%~6% of the total demand. The 90% estimated water head error is less than 1.5m, and 95% estimated water head error is less than 2.0 m. It agrees with the national criterion of China (the standard water head requirement of the civil water device is 1.5~2m in China). Fig. 7 presents the error distribution of nodal head.
Fig. 6 Estimated and measured source supplements.

Fig. 7. Error distribution of nodal head
5. Discussion

In most condition, the accuracy is primary topic of the off-line system. However, the stability is the most important thing for the real-time system. This system has run more than one year since 2011. There are two sentinel to help the system run away from the damage of the missing data, lost signals and wrong data. At first, the data out of the reasonable range is abandoned.

The secondly, the corrected algorithm including SVD uses Wiggins method to remove the high frequency error, the details can be found in [25]. The second importance thing is the efficiency of model. Usually, the calibration is time-cost work. The model is calibrated before it used. So it is good in most condition. In addition, it corrects the nodal demand very 15 min. It improves a small step each time. Its accuracy may be low at the beginning. After a few days, its precision is good enough to support the controllers to make their decision. This strategy will is little time-cost for every time, and at the same time it can keep the model fresh and effective. The last thing is the accuracy. It is the primary substance in most papers, and it is same in this paper. However, it is not right to the operator. First thing is that the system can work. The second thing is that the system can provide the result in the
limited time. The third thing is the accuracy. All sensors have observational error 0.1~1 m. Sometimes, it is up to 2~4 m. the flow meter precision is 1% in the laboratory, but it may exceed 10% on site. It is very difficult to determine accuracy of the simulated result or observed values. The total head loss in Guangzhou is more than 25 m. The experience in Guangzhou city shows that 2 m discrepancy between them is acceptable. Some papers use different mathematics methods to evaluate the accuracy. It does not work in the real-life work. From the viewpoint of engineers, the absolute value is more important. Moreover, this criterion is not absolute. The error is probability event. No one can promise that his simulate error is less than 2 m at any time. When the error is exceed this criterion, the operator can utilize the sensibility analysis to make them decision. It always gives right direction.

Acknowledgements

The author would like to acknowledge the financial support from the key Special Program on the S&T of China for the Pollution Control and Treatment of Water Bodies (22012ZX07408-002-003) and the National HighTechnology Research and Development Program("863"Program) of China(2012AA062608).

References