Theoretical study on pipe friction parameters identification in water distribution systems

<table>
<thead>
<tr>
<th>Journal:</th>
<th>Canadian Journal of Civil Engineering</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>cjce-2018-0021.R2</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Article</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>31-Dec-2018</td>
</tr>
</tbody>
</table>
| Complete List of Authors: | Liu, Yongxin; Harbin Institute of Technology, Key Laboratory of Cold Region Urban and Rural Human Settlement Environment Science and Technology, Ministry of Industry and Information Technology; School of Architecture  
Song, Li; University of Oklahoma  
Luo, Peng; Harbin Institute of Technology  
Jin, Hong; Harbin Institute of Technology |
| Keyword: | theoretical analysis < type of paper to review, water supply-irrig; drain < Hydrotechnical Eng., municipal < Environ & Sanitary Eng. |
| Is the invited manuscript for consideration in a Special Issue? | Not applicable (regular submission) |

https://mc06.manuscriptcentral.com/cjce-pubs
Title: Theoretical Study on Pipe Friction Parameters Identification in Water Distribution Systems

Yongxin Liu (corresponding author)
Lecturer, Post doctorate, Ph.D
Key Laboratory of Cold Region Urban and Rural Human Settlement Environment Science and Technology, Ministry of Industry and Information Technology; School of Architecture, Harbin Institute of Technology, Harbin, 150001, China. Email: liuyx_hit@163.com.

Li Song
Associate Professor, Ph.D
University of Oklahoma, 865 Asp Ave., Norman, OK, USA.

Peng Luo
Associate Professor, Ph.D
School of Architecture, Harbin Institute of Technology, Harbin, 150001, China.

Hong Jin
Professor, Ph.D
School of Architecture, Harbin Institute of Technology, Harbin, 150001, China.

Word count: About 6,500 words
Abstract: In water distribution systems (WDSs), operational modeling results could be affected by accuracy of pipe friction parameters (PFPs). Under a single hydraulic condition, unique values of PFPs cannot be achieved, even with the availability of pressure and discharge values at every node. This study established a theoretical model of PFP identification in WDSs by decoupling variables. Then, equations for identifying PFPs were expressed through energy conservation equations of a tree and relationships between pressure losses and flows in pipes under different hydraulic conditions. Further, equations for identifying PFPs can be transformed into linear simultaneous equations by substituting variables whose solvability is easy to study. The aim of this study is to develop a theoretical framework for identifying unique values of PFPs and provide a theoretical basis for an actual problem of PFP identification in a WDS. Moreover, a theoretical demonstrative example is presented to illustrate processes of obtaining unique and acceptable values of PFPs.

Key words: Water distribution systems; Pipe friction parameters; Nonlinear equations; Hydraulic calculation; Multiple hydraulic conditions.
Introduction

With rapid growth in population and increase in demand, supplying adequate water with reasonable energy consumption is becoming a global challenge. Water distribution systems (WDSs) are closely related to residents’ daily lives. Pipe friction parameters (PFPs) are fundamental pipe characteristics in WDSs studies, such as hydraulic calculation, optimal operation, operation simulation, and leakage diagnosis (Nishiyama and Filion 2013). As pipe roughness depends on its internal diameter, age, and other uncertain factors, many published papers (Creaco and Pezzinga 2015; Alvarruiz et al. 2015) propose the need for a method to determine pipe roughness under operational hydraulic conditions (flow conditions). Alvisi and Franchini (2010) stated that the quantity and quality of measuring data are important to all the calibration methods of WDSs. The study on WDSs mainly use two models: steady and transient flow models. As the effect of friction in pipes is usually not significant on transient flow models (Gong et al. 2013), this paper will focus on the use of steady flow models, by first summarizing previous relevant literature on these models.

Since the 1980s, scholars have studied identification and calibration problems of WDSs based on least squares principles. Ormsbee (1989) developed an implicit mathematical model for calibrating a WDS; the calibration model uses a nonlinear optimization algorithm along with a general network solver to adjust selected model parameters. This method needs pre-estimation values of model parameters. Nash and Karney (1999) presented a method based on least squares principles to solve models of WDSs, and the calibration of network models could be expressed and solved by minimizing the difference between observed and predictive values. Reddy et al. (1996) proposed a least squares method based on Gauss-Newton method. This method is efficient when it is applied in a small-scale WDS. Shamloo and Haghighi (2010) adopted a genetic algorithm to determine the best flow variation sources and solved this optimization model by using the sequence quadratic programming method. Acceptable results
are obtained using the method in large-scale WDSs. However, these calculation processes consume a considerable amount of time. The optimization methods based on least squares principles have been widely applied in PFP identification. The objective is to determine acceptable results of PFPs to ensure that the entire difference between observed and calculated values is the least value. However, as the number of independent unknowns is greater than that of simultaneous equations, the equations of PFP identification have infinitely many solutions. Any acceptable solution can be considered as the PFP identification results, which however, are not real resistance characteristics. The hydraulic model of WDSs built by considering these results simply guarantee the minimization of the distance between calculated and observed values at sampling nodes and pipes. With respect to computational efficiency, Liu and Zou (2011) applied an optimization method to PFP identification, the time consumption of which in a small example WDS is tens of minutes. In brief, for identifying PFPs based on the least squares principle, computational accuracy and efficiency must be improved (Vassiljev et al. 2015).

Conventional PFP identification methods are based on the least squares principle (optimization method). In addition, explicit methods play an important role in the study of WDS calibration, and solve a set of equations explicitly to identify unknown parameters. The problem of WDS model calibration is usually referred to as an “inverse” problem (Berardi 2017). Ormsbee and Wood (1986) proposed an explicit calibration algorithm, which revises pipe roughnesses to minimize the difference between the measured and calculated data of nodal pressures. Kapelan et al. (2003) proposed a hybrid genetic algorithm combining the Levenberg-Marquard algorithm with the genetic algorithm, thus obtaining relatively reliable results.

To further improve the calibration efficiency, studies proposed WDS calibration based on the mass continuity and energy conservation equations. To facilitate real-time nodal demand
calibration and state estimation of WDSs, Todini and Rossman (2013) applied the global
gradient algorithm to identify nodal demands, which can be calibrated directly without
solving a network repeatedly. By separating the known and unknown variables in the mass
continuity and energy conservation equations of a WDS, a direct inversion algorithm was
developed for nodal demand calibration of WDSs (Du et al. 2015). To identify PFPs in a
WDS, more unknown variables (including PFPs values) must be solved. Du et al. (2018)
presented a direct inversion algorithm for PFP calibration of WDSs. An identical framework
was proposed by using a global gradient algorithm. The method was proved to be effective for
solving networks and was adopted by EPANET. However, a calibration problem must be
determined. Liu et al. (2012) considered a Moore-Penrose pseudo-inverse solution to
identification equations as PFP identification results. Known variables in equations are
obtained from the observed data under a single hydraulic condition. However, a relationship
between the number of hydraulic conditions and accuracy of PFP results has not yet been
discussed.

The current paper suggests a new expression of PFP identification problem; this may increase
the chance of obtaining unique and acceptable values of PFPs.

Mathematical Model

In this study, hydraulic calculation equations (the mass continuity and energy conservation
equations) were used to express relationships between operational data and PFPs of a WDS. If
some operational data, such as pipe flows and nodal pressures, are available, those data are
viewed as known variables during the PFP identification process.

Considering a WDS of \( n + 1 \) nodes (including a reference node) and \( b \) pipes, the mass
continuity and energy conservation equations (absolute value form) can be written as
where \( a = \) an element of a basic incidence matrix, containing information about which pipes are connected to a particular node. The values \(-1\) and \(+1\) indicate that the flow in the pipe is directed toward and away from the node, respectively; otherwise, the value is 0 (Kumar et al. 2008). Further, \( g \) represents the pipe flow (m\(^3\)/h); \( q \) represents the nodal discharge (m\(^3\)/h); \( p \) is the nodal pressure (m); \( \Delta h \) is the pressure loss in a pipe (m); and \( s \) represents a PFP (h\(^2\)/m\(^5\)).

In the PFP identification processes, pipe flows and nodal pressures are regarded as operational data obtained by observing normal operation and testing conditions. Overall, the objective of this study was to obtain unique and acceptable values of PFPs when the observed values of nodal pressures and discharges are available. If all these observed values under a single hydraulic condition are available, Eq. (1) represents \( b + n \) equations, including \( 2b \) independent unknowns (containing pipe flows and PFPs). Commonly, in a WDS, \( b > n \); thus, the number of independent unknowns is greater than the number of equations. Eq. (1) is under-determined and has infinitely many solutions. Thus, to obtain unique results of PFPs, operational data under multiple hydraulic conditions must be provided. Further, the equations are nonlinear, containing dependent variables. To make these equations solvable, further steps are required to eliminate dependent variables and transform the nonlinear equations into their corresponding linear form.

This paper establishes a framework for PFP identification under multiple hydraulic conditions in a WDS. Theoretically, the method can also be applied in WDSs without flow sensors; however, better results can be obtained with flow sensors. To introduce this PFP identification method, the following two assumptions were made: (1) all nodal elevations are the same at 0
m in a WDS and (2) pumps are not present in the studied part of the WDS.

In the WDS analysis, pipes can be divided into a spanning tree (tree) and a corresponding cotree (cotree). Given any connected graph with \( n \) nodes, a tree contains no cycle and has \( n-1 \) edges. “We can choose a cycle and remove any one of its edges, and the resulting graph remains connected. We repeat this procedure with one of the remaining cycles, continuing until there are no cycles left.” (Wilson 1996) The graph that remains is a tree, and the rest of the graph is a cotree. To eliminate dependent unknowns in the identification processes, flows in the tree pipes, satisfying the mass continuity equation, are written with respect to the flows assigned to cotree pipes (Rahal 1995).

\[
G_t = -A_t^{-1} A_t G_t + A_t^{-1} Q
\]

where \( A_t \) is the tree incidence matrix composed of columns of \( A \) corresponding to tree pipes; \( A_l \) is the cotree incidence matrix composed of columns of \( A \) corresponding to cotree pipes; \( G_t \) is the subvector of \( G \) composed of tree pipe flows (m³/h); \( G_l \) is the subvector of \( G \) composed of cotree pipe flows (m³/h); and \( Q \) is the nodal discharge column vector (m³/h).

Eq. (2) can be transformed into algebraic equations. The element of the \( i \)th row and \( j \)th column of matrix \( -A_t^{-1} A_t \) of dimension \( n \times (b-n) \) can be expressed as \( c_{ij} \). Similarly, the element of the \( i \)th row and \( j \)th column of matrix \( A_t^{-1} \) of dimension \( n \times n \) can be expressed as \( d_{ij} \). The flow in the \( i \)th tree pipe can then be expressed as

\[
g_i = \sum_{j=1}^{b} c_{ij} g_j + \sum_{j=1}^{n} d_{ij} q_j \quad (i = 1, 2, \ldots, n)
\]

where \( g_i \) represents the flow in the tree pipe (m³/h); \( g_t \) represents the flow in a cotree pipe (m³/h); and \( q \) is a nodal discharge (m³/h).

According to the graph theory and hydraulic calculation method, for a WDS with \( n+1 \) nodes (including a reference node) and \( b \) pipes, there are \( (b-n) \) cotree pipes and \( n \) tree pipes (Wilson 1996). Eq. (1, b) can distinguish whether a pipe belongs to a selected tree or not.
Pressure losses in tree and cotree pipes can be written as

\[
\sum_{j=1}^{n} a_{ij} \cdot p_j = |\Delta h_i| = s_i \left( \sum_{j=1}^{n} c_{ij} \cdot g_j + \sum_{j=1}^{n} d_{ij} \cdot q_j \right)^2 \quad (i = 1, 2, \cdots, n)
\]

where \( a_{ij} \) and \( a_{ij} \) are elements of the tree and cotree incidence matrices, respectively; \( \Delta h_i \) and \( \Delta h_i \) are the pressure losses in the tree and cotree pipes (m), respectively; and \( s_i \) and \( s_i \) are the PFPs of the tree and cotree (h\(^2\)/m\(^5\)), respectively. \( \Delta h_i \) and \( \Delta h_i \) can be calculated using Eqs. (4) and (5), respectively. Once pressure losses in pipes are available by measuring devices and sensors arranged on WDSs, the PFPs of the tree and cotree can be expressed individually.

To analyze PFP identification problems, the energy conservation equations of WDSs must be focused on. The energy conservation equations of tree pipes under the \( k \)th hydraulic condition can be expressed as follows:

\[
\sum_{j=1}^{n} a_{ij} \cdot p_j = |\Delta h_i^{(k)}| = s_i \left( \sum_{j=1}^{n} c_{ij} \cdot g_j^{(k)} + \sum_{j=1}^{n} d_{ij} \cdot q_j^{(k)} \right)^2 \quad (i = 1, 2, \cdots, n; \quad k = 1, 2, \cdots, m)
\]

By submitting different values of \( k \), Eq. (6) can express energy conservation equations of tree pipes under all different hydraulic conditions. A PFP identification problem can be analyzed by first solving Eq. (6).

To simplify the expression of PFP identification, an equation was introduced that represents relationships between the pressure losses and cotree pipe flows. The equation between different hydraulic conditions can be written after obtaining the observed value of every nodal pressure. As \( \Delta h_{i0} \) appears in the denominator, every cotree pipe flow in the first hydraulic condition (selected as a reference condition) cannot be equal to zero. In a real-world scenario, flows exist in all the pipes under normal operational hydraulic conditions. If the normal
condition is satisfied, the reference condition can be selected randomly; this does not impact this study. Then, the ratio of cotree pipe flow \( j \) is written with respect to the ratio of pressure losses in the corresponding pipe between the \( k \)th and first conditions. Combined with the pipe flows of the cotree under the first hydraulic condition, the ratio can be expressed as

\[
\Delta h_i(k) = \frac{g_{i}^{(k)}}{g_{i}^{(1)}} = \text{sign}(\Delta h_i^{(k)}/\Delta h_i^{(1)}) \cdot \sqrt{\frac{\Delta h_i^{(k)}}{\Delta h_i^{(1)}}} \quad (i = 1, 2, \ldots, b-n; \quad k = 1, 2, \ldots, m)
\]

where \( \Delta h_i(k) \) is the ratio of the \( i \)th cotree pipe flows between the \( k \)th and first hydraulic conditions. In Eq. (7), \( \text{sign} \cdot (\cdot) \) is a sign function; if the variable is larger and smaller than 0, the function value is 1 and \(-1\), respectively; otherwise the function value is 0.

Under the \( k \)th hydraulic condition, Eq. (4) can be written, in view of Eq. (7), as follows:

\[
\left( 1 \right) \left( 1 \right) (1) (1) (1) (1)
\sum_{j=1}^{b} a_{ij} \cdot p_i^{(k)} = \Delta h_i^{(k)} = s_{ij} \left( \sum_{j=1}^{b} c_{ij} \cdot \Delta h_j^{(k)} \cdot g_{ij}^{(1)} + \sum_{j=1}^{b} d_{ij} \cdot q_j^{(1)} \right)^2 \quad (i = 1, 2, \ldots, n; \quad k = 1, 2, \ldots, m)
\]

As dependent variables, pipe flows of a cotree will not appear except for the first hydraulic condition. Owing to the difficulty in analyzing and solving nonlinear equations, variable substitution was introduced to transform Eq. (8) into a linear equation of PFP identification:

\[
\text{sign}(\Delta h_i^{(1)}) \cdot \sqrt{\Delta h_i^{(1)}} \cdot s_{ip} - \sum_{i=1}^{b-n} c_{ij} \cdot \Delta h_j^{(1)} \cdot g_{ij}^{(1)} = \sum_{j=1}^{b} d_{ij} \cdot q_j^{(1)} \quad (i = 1, 2, \ldots, n; \quad k = 1, 2, \ldots, m)
\]

where \( s_{ip} = 1/s_i^{1/2} \). The coefficient of Eq. (9) can be represented by the observed values of nodal pressures and discharges. The unknowns contain \((b-n)\) cotree pipe flows under the first hydraulic condition and \( n \) variables corresponding to PFPs of a tree. Eq. (9) can substitute Eq. (6), containing no dependent unknowns.

Can unique values of PFPs be obtained under a single hydraulic condition? Given the observed values of nodal pressures and discharges under a certain condition, Eq. (9) contains \( n \) subequations and \( b \) independent unknowns. As \( b \) (the number of pipes) is greater than \( n \) (the number of nodes) in a WDS, Eq. (9) represents a set of under-determined equations and has

\[
https://mc06.manuscriptcentral.com/cjce-pubs
\]
infinitely many solutions. However, unique solutions cannot be obtained. If unique values of PFPs cannot be obtained under one hydraulic condition, multiple hydraulic conditions must be provided to obtain unique or more accurate results for PFP identification.

As Eq. (9) is the first equation to be solved, it is a key part of PFP identification in a WDS. Under \( m \) hydraulic conditions, \( m \times n \) can be greater than \( b \), making unique solutions obtainable. If unique solutions to Eq. (9) can be obtained, unknowns, including PFPs of a tree and pipe flows of a cotree, could be solved. Once pipe flows of a cotree are found, PFPs of the cotree can be computed using Eq. (5) under the first hydraulic condition.

The parameter inversion process is written as

\[
\begin{align*}
  s_i &= (s_{\text{ori}})^2 \\
  s_j &= \Delta h_j^{(i)} (g_j^{(i)})^2
\end{align*}
\]

(10)  

\( i = 1, \ 2, \ \cdots, \ n; \ j = 1, \ 2, \ \cdots, \ b-n \)

Based on the parameter inversion process, PFPs can be expressed uniquely, allowing the solution analysis of Eq. (9).

The PFP identification problem under \( m \) hydraulic conditions can be transformed to find solutions to Eq. (9), and the linear simultaneous equations. Eq. (9) can also be used to judge the existence of unique values of PFPs. If unique solutions to Eq. (9) are found, the achievement of unique PFP identification results will be guaranteed. As Eq. (9) constitutes a set of linear simultaneous equations, the relationship among numbers of equations, independent unknowns, and solutions to equations can be studied using the theory of linear equation. If \( m \) hydraulic conditions are available, Eq. (9) contains \( m \times n \) subequations and \( b \) independent unknowns, with \((b-n)\) cotree pipe flows under the first hydraulic condition and \( n \) tree PFPs. To obtain unique solutions to an equation, the number of subequations must be greater than or equal to the number of independent unknowns. If hydraulic conditions of a WDS are sufficient, unique results of PFPs may be obtained; else, accurate results of PFPs cannot be obtained through theoretical calculations.
Under multiple hydraulic conditions, Eq. (9) constitutes a set of linear simultaneous equations, expressing a PFP identification problem. A coefficient matrix of Eq. (9) can be expressed by the observed values of nodal pressures and discharges under the corresponding hydraulic condition. The coefficient matrix is used as a tool to judge the existence of unique values of PFPs. According to the linear equation theory, a coefficient matrix of a set of linear simultaneous equations with a full-row rank is a sufficient and necessary condition for the existing unique solutions. The row rank of a matrix is the number of independent rows (Strang 2009). The row rank of the coefficient matrix will increase when adding hydraulic conditions, thus increasing the probability of forming a full-row-rank coefficient matrix. If the coefficient matrix is a full-row rank, unique solutions can be solved. In addition, if the number of hydraulic conditions can guarantee that the coefficient matrix is a full-row-rank matrix, the further addition of hydraulic conditions will not be necessary. Thus, the criteria mentioned earlier can be used to determine if unique values of PFPs can be obtained, thus gaining the advantage of not requiring additional hydraulic conditions.

Regardless of whether the amount of hydraulic conditions is adequate, unique results must be expressed. The current literature comprises two cases according to the rank of the coefficient matrix. First, if the coefficient matrix has a full-row rank, unique solutions can be calculated directly; else Moore-Penrose pseudo-inverse solutions are introduced to compute the “best fit” solutions (Penrose 1955). Second, when the coefficient matrix has a full-row rank, special Moore-Penrose pseudo-inverse solutions can express unique solutions to Eq. (9). Values of PFPs can then be calculated through the unique solutions.

Theoretical example

In this section, the authors show the verification of the method utilizing operational data under multiple hydraulic conditions to identify unique values of PFPs through an example WDS. In
case studies, unique values of PFPs can be calculated by utilizing observed values of all nodal pressures and discharges under multiple hydraulic conditions. The topology of the example WDS is shown in Fig. 1. (Cotree pipes consist of pipes 6, 7, 11, and 12)

Insert Fig. 1

The example WDS contains twelve pipes and nine nodes, wherein the nodes represent one water source and eight water users, as listed in Table 1.

Insert Table 1

By using the value of every PFP, the nodal pressures and discharges under four hydraulic conditions used in the identification processes were computed according to hydraulic calculation equations. In the following sections, the nodal pressures and discharges under different hydraulic conditions are assumed as “known values,” (Table 2) and the PFPs and pipe flows are unknowns.

Insert Table 2

For assessing the above-mentioned PFP identification method, four identification modes were combined with different hydraulic conditions to describe different identification circumstances of the example WDS. In this study, the selection of the identification modes (combination of hydraulic conditions) was designed to explain different relationships among the number of equations, number of variables, and row rank of the coefficient matrix. The combinations of the four identification modes are listed as follows. Identification mode 1
includes the first hydraulic condition; identification mode 2 includes the first and second hydraulic conditions; identification mode 3 includes the first and fourth hydraulic conditions; and identification mode 4 includes the first, second, and third hydraulic conditions. In every identification mode, operational data are available under certain hydraulic conditions. For instance, in identification mode 1, nodal pressures and discharges under the first hydraulic condition can be considered as known variables to identify PFPs in the example WDS.

The identification results of every identification mode can be solved using calculation tools (such as ABS, INV, DIAG, and PINV) included in MATLAB 2011. Consider the calculation process of identification mode 3 as an example. According to the definition, a basic incidence matrix \( A \) related to the example WDS can be written as follows. The nodal discharges and pressures under different hydraulic conditions are listed in Table 2.

\[
A = \begin{bmatrix}
-1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -1 & 0 & -1 & 1 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & 0 & -1 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & -1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & -1 \\
\end{bmatrix}
\]

Then, variables \( s_{pi} \) and \( e_{p} \) can be obtained by solving Eq. (9). To illustrate the calculation processes, variable symbols (including a coefficient matrix, nonhomogeneous term, and solution) and values corresponding to identification mode 3 are shown in Fig. 2 (Screenshots are captured from Matlab 2011).

Then, PFPs of the entire WDS are calculated using Eq. (10). Table 3 lists the identification values and deviations of PFPs under every identification mode.
Compared with the setting values of PFPs, the accuracy of identification results evaluated by the deviations of PFPs can be summarized as follows.

- In identification mode 1, operational data under the first hydraulic condition are available for PFPs identification, and the number of simultaneous equations is less than the number of unknowns. According to the linear equation theory, there are no unique solutions. Thus, the “identification results” are unsatisfied here.

- In identification mode 2, operational data exist under two hydraulic conditions. Although the number of simultaneous equations is more than the number of unknowns, the coefficient matrix of simultaneous equations is not column full rank. Compared with identification mode 1, increase of the coefficient-matrix rank from eight to nine results in an improvement in the identification results. However, acceptable results of PFPs identification were not achieved.

- In identification mode 3, the coefficient-matrix rank is 12. Then, according to the linear equation theory, obtaining unique results of PFPs can be guaranteed. The maximum identification deviation of PFPs is less than 0.95% and the average identification deviation of PFPs is 0.43%; these values are accurate enough for the engineering field. As the identification results are obtained by increasing the amount of hydraulic conditions, they can reveal the unique and acceptable values of PFPs in a WDS. This identification mode can illustrate the probability to obtain unique values of PFPs under multiple hydraulic conditions.

- In identification mode 4, operational data are obtained under three hydraulic
conditions. As the rank of the coefficient matrix is also 12, unique results of PFPs are guaranteed. The maximum identification deviation of PFPs is less than 0.78% and the average identification deviation of PFPs is 0.23%. The comparison of the results of modes 3 and 4 shows that when the rank of the coefficient matrix equals to the column full rank, the addition of hydraulic conditions only slightly increases the chance of finding the unique results of PFPs or improving precision.

Conclusions

In this paper, the mass continuity and energy conservation equations were used to express relationships between operational data and PFPs of a WDS. Brief forms of equations can be obtained by eliminating dependent unknowns. The study also proposed methods for solving these equations and inverting the parameters. The theoretical analysis showed that the number of hydraulic conditions is a key factor in addition to operational data. Furthermore, quantitative relationships exist among hydraulic conditions, equations, variables, and rank of a coefficient matrix. Moreover, by analyzing the quantitative relationships, the condition of obtaining accurate PFPs results was illustrated. The following major conclusions can be drawn from this study.

1) Given the observed data of nodal pressures under multiple hydraulic conditions, the theoretical example mentioned in this paper shows that unique values of PFPs can be obtained.

2) Acceptable results of PFP identification can be achieved through theoretical and empirical analyses. The results show that when a coefficient matrix corresponding to the simultaneous equations of PFP identification is column full rank, unique and acceptable results can be obtained.

3) The proposed method has higher computational efficiency than an optimization method.
Moreover, the time consumption in the four identification modes is 15.0, 15.8, 20.6, and 92.5 ms, respectively.

4) Identification modes 3 and 4 show that when a column-full-rank coefficient matrix is achieved, the addition of hydraulic conditions only slightly increases the chance of finding the unique results of PFPs or improving precision. However, time consumption will increase with the number of hydraulic conditions.

**Discussions**

1) The theoretical analysis showed that the proposed method can be applied to complex networks, and the solution processes do not differ.

2) In the identification processes, the use of the least number of hydraulic conditions to identify unique values of PFPs can alleviate the difficulty of obtaining more multiple hydraulic conditions.

3) In this paper, we only discussed PFP identification by using observed data of nodal pressures. In reality, the observed points have many arrangements; this can guarantee a unique PFP identification result. A quantitative relationship exists among nodal pressures, pipe flows, and hydraulic conditions. For example, if all the nodal pressures and cotree pipe flows can be observed, a unique identification result of PFPs can be obtained theoretically using the observed data under a single hydraulic condition.

4) Through this study, we believe that the proposed method can be further applied to solve similar problems in other fields such as heat and gas-supply networks.

**Acknowledgments**

This study was supported by the National Natural Science Foundation of China (51678177) and the Heilongjiang Postdoctoral Financial assistance (LBH-Z15083).
References


**Figure Legends**

**Fig. 1. Schematic diagram of an example WDS**

**Fig. 2 A sample of initial matrix and identification results**
### Table 1 Details of the example WDS

<table>
<thead>
<tr>
<th>Pipe No.</th>
<th>From node</th>
<th>To node</th>
<th>Diameter (mm)</th>
<th>Length (m)</th>
<th>PFP (h²/m²)</th>
<th>Flow (m³/h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>515</td>
<td>150</td>
<td>1.0×10⁻⁴</td>
<td>800</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>3</td>
<td>311</td>
<td>150</td>
<td>1.0×10⁻⁴</td>
<td>320</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>4</td>
<td>515</td>
<td>150</td>
<td>1.0×10⁻⁴</td>
<td>800</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>5</td>
<td>363</td>
<td>150</td>
<td>1.0×10⁻⁴</td>
<td>400</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>6</td>
<td>311</td>
<td>150</td>
<td>1.0×10⁻⁴</td>
<td>300</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>5</td>
<td>311</td>
<td>150</td>
<td>1.0×10⁻⁴</td>
<td>300</td>
</tr>
<tr>
<td>7</td>
<td>5</td>
<td>6</td>
<td>261</td>
<td>150</td>
<td>1.0×10⁻⁴</td>
<td>200</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>7</td>
<td>311</td>
<td>150</td>
<td>1.0×10⁻⁴</td>
<td>300</td>
</tr>
<tr>
<td>9</td>
<td>5</td>
<td>8</td>
<td>261</td>
<td>150</td>
<td>1.0×10⁻⁴</td>
<td>240</td>
</tr>
<tr>
<td>10</td>
<td>6</td>
<td>9</td>
<td>261</td>
<td>150</td>
<td>1.0×10⁻⁴</td>
<td>200</td>
</tr>
<tr>
<td>11</td>
<td>7</td>
<td>8</td>
<td>261</td>
<td>150</td>
<td>1.0×10⁻⁴</td>
<td>240</td>
</tr>
<tr>
<td>12</td>
<td>8</td>
<td>9</td>
<td>261</td>
<td>150</td>
<td>1.0×10⁻⁴</td>
<td>150</td>
</tr>
</tbody>
</table>
Table 2. The known operational data (nodal pressures and discharges) under different hydraulic conditions

<table>
<thead>
<tr>
<th>Node Number</th>
<th>Condition One Pressures (m)</th>
<th>Discharge s (m³/h)</th>
<th>Condition Two Pressures (m)</th>
<th>Discharge s (m³/h)</th>
<th>Condition Three Pressures (m)</th>
<th>Discharge s (m³/h)</th>
<th>Condition Four Pressures (m)</th>
<th>Discharge s (m³/h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>110</td>
<td>1580.90</td>
<td>110</td>
<td>1679.10</td>
<td>110</td>
<td>1618.10</td>
<td>110</td>
<td>1535.70</td>
</tr>
<tr>
<td>2</td>
<td>50.00</td>
<td>-84.10</td>
<td>42.00</td>
<td>-134.10</td>
<td>48.00</td>
<td>-43.40</td>
<td>54.00</td>
<td>-30.60</td>
</tr>
<tr>
<td>3</td>
<td>40.00</td>
<td>-51.70</td>
<td>32.00</td>
<td>-51.70</td>
<td>37.00</td>
<td>-15.50</td>
<td>43.50</td>
<td>-32.50</td>
</tr>
<tr>
<td>4</td>
<td>45.00</td>
<td>-223.50</td>
<td>37.00</td>
<td>-271.60</td>
<td>41.00</td>
<td>-182.80</td>
<td>48.00</td>
<td>-171.00</td>
</tr>
<tr>
<td>5</td>
<td>36.00</td>
<td>-301.00</td>
<td>28.00</td>
<td>-301.00</td>
<td>31.00</td>
<td>-263.90</td>
<td>38.50</td>
<td>-280.30</td>
</tr>
<tr>
<td>6</td>
<td>33.00</td>
<td>-264.60</td>
<td>25.00</td>
<td>-264.60</td>
<td>27.00</td>
<td>-271.30</td>
<td>35.01</td>
<td>-266.50</td>
</tr>
<tr>
<td>7</td>
<td>37.00</td>
<td>-59.20</td>
<td>29.00</td>
<td>-59.20</td>
<td>30.00</td>
<td>-86.80</td>
<td>38.50</td>
<td>-73.70</td>
</tr>
<tr>
<td>8</td>
<td>32.00</td>
<td>-282.20</td>
<td>24.00</td>
<td>-282.20</td>
<td>24.00</td>
<td>-336.30</td>
<td>33.01</td>
<td>-310.90</td>
</tr>
<tr>
<td>9</td>
<td>30.00</td>
<td>-314.60</td>
<td>22.00</td>
<td>-314.60</td>
<td>21.00</td>
<td>-418.10</td>
<td>30.51</td>
<td>-370.20</td>
</tr>
</tbody>
</table>

*The first node is considered as a reference node. All nodal elevations are assumed as 0 m in the WDS. The relative pressure at the reference node is 110 m.
Table 3. The identification results of PFPs under different identification modes

<table>
<thead>
<tr>
<th>Pipe No.</th>
<th>Identification Values of PFPs (h²/m³)×10⁻⁴</th>
<th>Identification Deviation of PFPs (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mode 1</td>
<td>Mode 2</td>
</tr>
<tr>
<td>1</td>
<td>0.458</td>
<td>1.000</td>
</tr>
<tr>
<td>2</td>
<td>0.431</td>
<td>0.524</td>
</tr>
<tr>
<td>3</td>
<td>3.419</td>
<td>1.000</td>
</tr>
<tr>
<td>4</td>
<td>0.418</td>
<td>2.176</td>
</tr>
<tr>
<td>5</td>
<td>0.378</td>
<td>0.472</td>
</tr>
<tr>
<td>6</td>
<td>31.202</td>
<td>1.285</td>
</tr>
<tr>
<td>7</td>
<td>6.427</td>
<td>4.621</td>
</tr>
<tr>
<td>8</td>
<td>3.170</td>
<td>0.790</td>
</tr>
<tr>
<td>9</td>
<td>0.578</td>
<td>2.141</td>
</tr>
<tr>
<td>10</td>
<td>0.548</td>
<td>0.741</td>
</tr>
<tr>
<td>11</td>
<td>5.038</td>
<td>0.746</td>
</tr>
<tr>
<td>12</td>
<td>3.079</td>
<td>1.553</td>
</tr>
</tbody>
</table>

Average Deviation 0.43 0.23
Fig. 1. Schematic diagram of an example WDS
Fig. 2 A sample of initial matrix and identification results

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7.7460</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>3.1626</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
<td>8.0626</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>0</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3.7421</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2.6459</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2.8261</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.7318</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Coefficient matrix

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7.4621</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.0274</td>
<td>0</td>
<td>1.0485</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>3.2402</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.0800</td>
<td>0</td>
<td>1.1181</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
<td>7.8739</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-1.0274</td>
<td>0</td>
<td>-1.0485</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3.9368</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.0274</td>
<td>-1.0800</td>
<td>1.0485</td>
<td>-1.1181</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2.9153</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.0800</td>
<td>0</td>
<td>1.1181</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3.0820</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-1.0485</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2.3450</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.0485</td>
<td>-1.1181</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2.1211</td>
<td>0</td>
<td>0</td>
<td>1.1181</td>
</tr>
</tbody>
</table>

Non-homogeneous term

\[
\sum_{j=1}^{n} q_j \cdot q_j^*(k) = \sum_{j=1}^{n} d_j \cdot q_j^*(k)
\]

Solution

\[
\text{Eq. (9)}
\]

\[
(k = 1 \text{ and } 2)
\]