Model discrimination in water distribution systems

G. Darvini

*Department IEA, Università Politecnica delle Marche, via Brecce Bianche, I-60131 Ancona, Italy

Abstract

Model discrimination procedures are useful tools for selection of the best models to be used to represent a specific process. In the paper a sequential discrimination procedure is adopted in the water distribution system analysis for selection of the most suitable model to perform hydraulic simulations and for estimation of precise model parameters, that are usually treated as independent techniques. The results indicate that the procedure allows to discriminate among different models, based on the collected experimental data. This results is achieved by selecting the best additional experiments in the design of the sensor location for pressure and/or flow measurements.

© 2013 The Authors. Published by Elsevier Ltd. Open access under CC BY-NC-ND license. Selection and peer-review under responsibility of the CCWI2013 Committee.

Keywords: Water distribution system; Model discrimination; Experimental design; Parameter estimation

1. Introduction

Calibration of computer models for water distribution system (WDS) analysis is an essential performed step in the model building process. Calibration of pipe network systems consists of determining the physical and operational characteristics of an existing system. This is achieved by determining various parameters that, when input into a hydraulic simulation model, will yield a reasonable match between measured and predicted pressures and flows in the network.

Calibration is considered predominantly as the process by which the parametric model uncertainty is being reduced, whilst model building is concerned with reducing the structural model uncertainty. Although the initial
stages of model assembling, which involve collecting basic WDS data (network topology, asset data, etc.), demand allocation and validation or correction of gross input data errors, are important precursors for calibration, often they are considered outside the scope of calibration (Savic et al., 2009).

Actually, very frequently the analyst has to define which model is the most suitable one for a particular application, as the final results and conclusions usually depend on the model used to perform the simulations and provide the interpretation of the analyzed problem. Sometimes, it is not possible to make a proper choice based solely on the available information, because the collected experimental data can be explained adequately by more than one model. In this case, one should perform new experiments for selection of one of the rival models. This procedure, called model discrimination, constitutes an important tool for scientists and engineers. As interpretation and use of available experimental data depend on the model structure, techniques for design of experiments for selection of the best model are of fundamental importance. Besides, experiments must often be designed for estimation of model parameters and reduction of variances of model predictions (or parameter estimates). These two classes of experimental design techniques generally lead to different experimental designs, although model discrimination and reduction of variances of parameter estimates are closely related to each other (Schwaab et al., 2008).

In this paper a technique largely used to discriminate rival models in the field of chemical engineering is applied to model discrimination in water distribution system analysis (Schwaab et al., 2006, Schwaab et al., 2008, Alberton et al., 2011, 2012). This technique makes use of model probabilities, determined based on simple statistical arguments. A numerical example illustrates the application of the proposed discrimination procedure. The obtained results show that this technique is able to discriminate models and simultaneously to improve parameter estimates.

2. Discrimination of rival models

Several criteria for experimental design for model discrimination have been proposed in the literature. A very simple design criterion for sequential model discrimination was proposed by Hunter and Reiner (1965). After performing \( N \) preliminary experiments, a new experimental condition \( x \) should be selected in order to maximize the model discrimination function, defined as

\[
D(x) = [y_1 - y_2]^2
\]  

where \( y_1 \) and \( y_2 \) are the expected model responses for Models 1 and 2 with model parameters \( \theta_1 \) and \( \theta_2 \) estimated from the available experiments. Eq. (1) assumes that model discrimination is improved when the difference between model responses increases.

Buzzi-Ferraris and Forzatti (1983) proposed a criterion for sequential model discrimination where the new experimental condition should maximize the ratio between the variance of model deviations and the mean value of model prediction variances. For models with multiple responses the discriminant between two rival models is given by

\[
D_{m,n}(x) = (y_m - y_n)^T V_{m,n}^{-1} (y_m - y_n)
\]  

where \( y_m \) is a vector of responses of model \( m \) and \( V_{m,n} \) is defined as

\[
V_{m,n} = 2V + V_m + V_n
\]  

where \( V \) is the covariance matrix of experimental deviations and \( V_m \) is the covariance matrix of model prediction deviations calculated from model \( m \). One should maximize the maximum value of \( D_{m,n} \) during the design phase and eliminate bad models during the data analysis phase, in order to avoid the repetitive design of experiments for elimination of bad model candidates. The sequences of designed experiments obtained with the criteria developed by Buzzi-Ferraris and Forzatti (1983), Buzzi-Ferraris et al. (1984) and Buzzi-Ferraris et al.
(1990) do not depend on the ordering of the experimental observations. The optimum value of discriminant $D_{m,n}$ must be larger than the number of model responses; otherwise, discrimination is not possible and the experimental procedure must be halted. The adequacy of model evaluations must be performed with the classical $\chi^2$-test for model adequacy (or any other statistical test developed for analysis of model adequacy).

Recently, Schwaab et al. (2008) proposed the use of model probabilities to formulate a design criterion for model discrimination defined in Eq. (2). In order to concentrate the experimental efforts on discriminating the more plausible models, the discrimination is based on the performance of pairs of models as

$$D_{m,n}(x) = \left(P_m P_n\right)^2 (y_m - y_n)^T V^{-1}_{m,n} (y_m - y_n)$$

where $Z$ is a parameter used to modulate the relative importance of the rival models: if $Z$ is greater than 1, model prediction differences are magnified; if $Z$ is smaller than 1, model prediction differences are minimized. $P_m$ is the relative probability for model $m$ to be the best model, calculated from the absolute model probabilities as

$$P_m = \frac{\phi_m}{\sum_{n=1}^{M} \phi_m}$$

where $\phi_m$ is the absolute probability for model $m$ to be the best model, as defined in Eq. (6). Assuming that model $m$ is perfect, that experiments are well done and that experimental deviations follow the normal distribution, the objective function presented in Eq. (7) approximately follows a $\chi^2$-distribution with $\nu$ degrees of freedom ($\nu = NE-NP_m$, where $NE$ is the total number of experimental points and $NP_m$ is the number of estimated parameters in model $m$). Consequently, a value can be assigned to each model probability in the form:

$$\phi_m = 1 - p\left[\chi^2 \leq F_m\right]$$

where $F_m$ is the minimum value of the objective function obtained in the parameter estimation of model $m$, defined as

$$F_m = \sum_{i=1}^{N} \left( y_{i,m} - y_{i}^e \right)^T V^{-1}_i \left( y_{i,m} - y_{i}^e \right)$$

where $y_{i,m}$ is the vector of model responses, $y_{i}^e$ is the vector of experimental responses and $V_i$ is the covariance matrix of experimental deviations for experiment $i$. As the objective function does not depend on ordering of the experimental observations, the absolute probabilities calculated from Eq. (5) do not depend on ordering of experimentation either. Schwaab et al. (2006) claimed that the absolute model probabilities, calculated from the $\chi^2$-distribution, can be readily used as a classical $\chi^2$-test for model adequacy, allowing for evaluation of absolute model performances and analysis of the relative performances.

3. Design of experiments

After selecting the best model, one may also need to design additional experiments to increase the precision of model parameters and model predictions. This task is closely related with the posterior covariance matrix of parameter estimates (Bard, 1974); that is, the expected covariance matrix of parameter estimates after the addition of new $k$ experiments to the set of available experimental data and re-estimation of model parameters. The first criterion defined in the literature for minimization of variances of parameter estimates consists in the minimization of the determinant of the posterior covariance matrix of the parameter estimates, since this determinant is proportional to the volume of the hyper-ellipsoid that defines the parameter confidence region. Consequently, minimizing the determinant is equivalent to increasing the confidence in the parameter estimates.
One can propose the execution of two blocks of experimental design: the first one for model discrimination and the second one for precise parameter estimation, after selection of the best model. As sometimes this may lead to long and expensive experimental plans, some authors suggested the formulation of a single design criterion for simultaneous model discrimination and precise parameter estimation.

It must be clear that good model discrimination is closely related with the estimation of precise parameters. The more precise are the model parameters (and, consequently, the model predictions), the easier is the discrimination between model responses. When model predictions are very uncertain, no model discrimination is possible. From Eq. (4) one can see that discrimination is not possible when model prediction variances are high.

In typical model discrimination problems, because the true model is not known, decision-making usually leads to formulation of multiobjective problems. Different objective functions can be proposed and, consequently, distinct optimal solutions can be obtained, requiring the computation and analysis of the non-dominated Pareto sets. Design criteria would propose selection of experiment based on larger discriminant values. Nevertheless, one can certainly argue that the absolute value of the discriminant is not so important when model discrimination is also expected to occur at alternative experimental conditions. In this sense, the number of discriminated models should be more important than the absolute value of the discriminant, although published criteria for model discrimination usually concentrate on the discriminant values and not on the number of discriminated models.

A discriminant for experimental design was proposed by Alberton et al. (2012) as

\[
\psi_m(x) = \left( \sum_{n=1,n \neq m}^{M} \phi_m(x, n) \right) + \lambda \frac{D_m(x)}{D_{m,\text{max}}} \tag{8}
\]

where \( \phi_m(x, n) \) is a binary function that can be equal either to 0 or to 1. \( \phi_m(x, n) \) is equal to zero if the results obtained after performing experiments in \( x \) do not allow for discrimination of model \( n \); otherwise, is equal to 1.

The first term on the right-hand side of Eq. (8) is evaluated by considering that the responses of experiments performed in \( x \) are equal to predictions provided by model \( m \) and after re-estimation of the model parameters for the remaining \((m-1)\) rival models. The discriminant \( D_m(x) \) represents any model discrimination design criterion used to characterize the divergence of the model performances, \( D_{m,\text{max}} \) is the maximum value of the conventional discriminant function in the analyzed experimental region. As the ratio between \( D_m(x) \) and \( D_{m,\text{max}} \) is always smaller than 1, Eq. (8) explicitly indicates that the best candidate trials are the ones that lead to the elimination of the maximum number of rival models. When different experimental conditions allow for elimination of equal number of rival models, the best candidate trial is the one that leads to the maximum \( D_m(x) \) value. The factor \( \lambda \) is an arbitrary parameter (\( \lambda < 1 \)) that can be used to control the importance of the usual discriminant function.

Finally, in order to select the best candidate trial, it is necessary to design a decision-making criterion. Among others, Maxmin, Bayesian or Equal Model Weights optimization criteria can be adopted to select the best experimental conditions in the Pareto set.

4. Methodology

4.1. Sequential model discrimination

According to the usual approach, the calculations begin with the estimation of parameters for \( M \) rival models, using \( N \) available experiments.

The covariance matrix of parameter estimates, based on the \( N \) available experiments can be calculated as follows (Bard, 1974):

\[
V_{b,m} = \left( \sum_{i=1}^{N} \mathbf{B}_m^T(x_i) \mathbf{N}^{-1}(x_i) \mathbf{B}_m(x_i) \right)^{-1} \tag{9}
\]
where $V_{\theta,m}$, is the covariance matrix of parameter estimates for model $m$, $V$ is the covariance matrix of experimental deviations and $B_m$ is the sensitivity matrix that contains the first derivatives of model $m$ responses with respect to its parameters, as

$$B_m = [b_{m,r,s}] = \frac{\partial y_{r,m}}{\partial \theta_s}. \quad (10)$$

Tests of model adequacy are then performed for all proposed models. Adequate models are then used for design of the new experimental condition. By following Schwaab et al. (2008), it is proposed here that the posterior covariance matrix of parameter estimates be used for model discrimination. The use of the posterior covariance matrix of parameter estimates can be justified by the fact that model performances are compared after execution of the designed experiment and re-estimation of model parameters for all models. The posterior covariance matrix of model parameters can be defined as (Bard, 1974)

$$V_{\theta,m}(x_{N+1}) = \left[ B_m^T (x_{N+1}) V^{-1}(x_{N+1}) B_m (x_{N+1}) + V_{\theta,m}^{-1} \right]^{-1} \quad (11)$$

and the calculation of the posterior covariance matrix of model predictions should be performed as

$$V_m(x_{N+1}) = B_m(x_{N+1}) V_{\theta,m}(x_{N+1}) B_m^T(x_{N+1}) \quad (12)$$

In Eq. (11) all terms are calculated as functions of the new selected experimental condition $x_{N+1}$. Finally, the posterior covariance matrix of the differences between model predictions becomes

$$V_{m,n}(x_{N+1}) = 2V(x_{N+1}) + V_m(x_{N+1}) + V_n(x_{N+1}) \quad (13)$$

and the discriminant becomes

$$D_{m,n}(x_{N+1}) = \left( P_m P_n \right)^T \left[ y_m(x_{N+1}) - y_n(x_{N+1}) \right] \left[ V_{m,n}(x_{N+1}) \right]^{-1} \left[ y_m(x_{N+1}) - y_n(x_{N+1}) \right]. \quad (14)$$

One must observe that the design criteria depend on the values of the parameter estimates, which has been omitted from the equations for the sake of simplicity.

By considering all pairs of models $m$ and $n$, the discriminant can be computed as

$$D(x_{N+1}) = \sum_{m=1}^{M-1} \sum_{n=m+1}^{M} D_{m,n}(x_{N+1}). \quad (15)$$

The procedure consists of 4 steps. For each model $m$, $m=1,M$: 1) calculate $D_{m,max}$ in the whole experimental grid. 2) Simulate the response variables for $x_{N+1}$ as experimental points. 3) For each model $n \neq m$, $n=1,M$ re-estimate the parameters and calculate model probability. Then, calculate $D_{m,n}(x_{N+1})$ and update $D_m(x_{N+1})$ and $D_{m,max}$. If the model $n$ has been discarded, update $\phi_m(x_{N+1},n) = \phi_m(x_{N+1},n) + 1$. 4) Calculate $\psi_m(x_{N+1})$.

Among $m$ models, by means the decision criterion the best additional experiment is selected. In the example, the Bayesian and the equal weights criteria are adopted in the form (Alberton et al., 2011):

$$x_{N+1}^{best} = \arg \max_m \sum_{m=1}^{M} P_m \psi_m(x_{N+1}) \quad (16)$$
\[
x_{i+1}^{\text{best}} = \arg \max \frac{1}{M} \sum_{m=1}^{M} \psi_m(x_{i+1}).
\] (17)

If after the additional experiment the model discrimination is not attained, the procedure is repeated until achievement of one of the two stopping criteria: successful model discrimination or conclusion that model discrimination is not possible within the proposed experimental grid (Schwaab et al., 2006).

### 4.2. Calibration procedure

For calibrating the network model the approach proposed by Greco and Del Giudice (1999) can be adopted. The objective of this method is to minimize the sum of the squares of the differences between the unknown decision variable \( \theta \) and the initially assumed parameter \( \theta^* \), under the constraint that measured and computed response of the model differ by less than a specific tolerance. By this approach the knowledge embodied in the engineering rules concerning an initial parameter assignment plays a substantial role in the calibration process.

The objective function is

\[
\min f(\theta) = \sum_{s=1}^{S} \left( \theta_s - \theta_s^* \right)^2
\] (18)

where \( S \) is the number of parameter, and the constrains are

\[
\left| y_{r,m} - y_{r}^c \right| < K_r, \quad r = 1, \ldots, R
\] (19)

being \( R \) the number of responses of the model and \( K_r \) the imposed tolerance for response \( r \).

Having assumed, as a first approximation, a linear behavior of the network, the set of constrains can be rewritten as

\[
\left| y_{r,m} - y_{r}^c - \sum_{s=1}^{S} \left[ b_{m,s} \left( \theta_s - \theta_s^* \right) \right] \right| < k_r, \quad r = 1, \ldots, R
\] (20)

where the tolerance \( k_r \) must be taken slightly less than the corresponding \( K_r \).

The algorithm consists in four steps. 1) Compute the model response \( y_{m} \) with the parameter initially assumed. 2) If the constraint (19) is satisfied the calibration process is completed, otherwise a new set of constraints (20) is derived by computing the sensitivity matrix \( B_m \). 3) Solve the problem that minimizes the objective function (18) under the constrains (20). 4) Compute the model response with the new estimates of parameters. Steps 2-4 are repeated until convergence of the procedure. The iteration is required by the nonlinearity of the problem.

### 5. Application to test case

The above procedure is applied to water distribution systems for discriminating rival models and improving the parameter estimate. Two models of the same network are considered for the comparison. The model response is the pressure at nodes, whereas the model parameter is the pipe roughness.

A hypothetical network of 13 pipes, 9 junction nodes and 5 loops was used to test the procedure. The network is illustrated in Fig. 1 and the pipe characteristics are shown in Table 1. The arbitrary set of roughness values \( \theta^{**} \) has been assigned to represent the actual network.
Initially, nodes 4, 7, 8, and 9 are equipped with a pressure sensor. Given the actual nodal demands of Table 2, the simulation model uses the actual roughness values to compute the pressure at nodes where sensors are installed. These pressure values represent the measured pressure.

Water distribution network simulation model EPANET (Rossman, 2000) was used to calculate nodal pressures, but other software can be employed.

The pipes are divided in three groups with the same roughness coefficient, and the initial estimates of pipe roughness $\theta^*$ reported in Table 1 are arbitrarily chosen. The reduction of the parameter dimension is a common engineering practice for reducing the number of unknown parameters in the WDS calibration. The influence of pipe groupings on the model error and the model prediction error was thoroughly analyzed by Mallick et al. (2002) and it is out of the scope of this work. Anyway, a small parameter dimension can lead to large model error and small model prediction error.

Table 1. Pipe characteristics.

<table>
<thead>
<tr>
<th>Pipe</th>
<th>Diameter (mm)</th>
<th>Length (m)</th>
<th>Actual roughness $\theta^{**}$ (mm)</th>
<th>Initial estimated roughness $\theta^*$ (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>250</td>
<td>500</td>
<td>1.50</td>
<td>1.20</td>
</tr>
<tr>
<td>2</td>
<td>200</td>
<td>500</td>
<td>1.30</td>
<td>1.20</td>
</tr>
<tr>
<td>3</td>
<td>250</td>
<td>500</td>
<td>0.90</td>
<td>1.20</td>
</tr>
<tr>
<td>4</td>
<td>200</td>
<td>500</td>
<td>1.20</td>
<td>1.20</td>
</tr>
<tr>
<td>5</td>
<td>150</td>
<td>500</td>
<td>1.10</td>
<td>1.00</td>
</tr>
<tr>
<td>6</td>
<td>200</td>
<td>500</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>7</td>
<td>150</td>
<td>500</td>
<td>1.40</td>
<td>1.00</td>
</tr>
<tr>
<td>8</td>
<td>200</td>
<td>500</td>
<td>0.80</td>
<td>1.00</td>
</tr>
<tr>
<td>9</td>
<td>150</td>
<td>500</td>
<td>0.70</td>
<td>0.50</td>
</tr>
<tr>
<td>10</td>
<td>150</td>
<td>500</td>
<td>0.60</td>
<td>0.50</td>
</tr>
<tr>
<td>11</td>
<td>150</td>
<td>500</td>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td>12</td>
<td>150</td>
<td>500</td>
<td>1.05</td>
<td>0.50</td>
</tr>
<tr>
<td>13</td>
<td>150</td>
<td>707</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Table 2. Nodal demand and pressure values.

<table>
<thead>
<tr>
<th>Node</th>
<th>Actual demand (l/s)</th>
<th>Actual pressure (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The actual network, hereafter named model 1, was compared with model 2 that differs for the diameter of the pipe 13, that is 175 mm instead of 150 mm. When one model reaches a relative probability higher than 98%, the discrimination procedure is successfully finished with one model selected as the best one. Experimental variances are set to 0.01 for each \( y_r \). The errors of experimental responses are assumed to be independent from each other.

Using the measured pressure at sensor nodes and from the initial estimates of pipe roughness, the parameter of both models were estimated, as presented in Table 3. \( P_m \) is the relative probability of model \( m \) to be the best model and \( \text{det} V_{\theta,m} \) is the determinant of the covariance matrix of parameter estimates for model \( m \), that is proportional to the volume of the confidence region of model parameters (Bard, 1974). After execution of the preliminary experiment, the model discrimination is not possible and new experiments must be designed in order to discriminate between them. The discriminant value computed by Eq. (14) is 13.68, indicating that model discrimination is possible. Since some authors argue strongly against the use of the model probabilities in discrimination procedure (Buzzi-Ferraris and Manenti, 2009), in the example is assumed \( Z = 0 \). Table 3 shows that the model 1 has a larger probability to be the best model and the determinant of the covariance matrix of parameter estimates for model 1 is smaller than value obtained for model 2.

### Table 3. Preliminary experiment.

<table>
<thead>
<tr>
<th>Model</th>
<th>( \theta_1 ) (mm)</th>
<th>( \theta_2 ) (mm)</th>
<th>( \theta_3 ) (mm)</th>
<th>( P_m ) (%)</th>
<th>( \text{det} V_{\theta,m} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.111</td>
<td>0.980</td>
<td>0.489</td>
<td>97.7</td>
<td>1.5 (-10^{25})</td>
</tr>
<tr>
<td>2</td>
<td>1.108</td>
<td>0.992</td>
<td>0.529</td>
<td>2.30</td>
<td>6.4 (-10^{25})</td>
</tr>
</tbody>
</table>

In order to design a new experimental condition, the selection was performed by adding a sensor to the previous set. The cases a, b and c refer to nodes 5, 6 and 3, respectively. The values of the discriminant computed by Eq. (8) are reported in Table 4 for both models and three cases. By adopting the equal weights decision-making criterion expressed by Eq. (17) the best experimental condition is the case c. The same additional experiment is chosen when the Bayesian approach is adopted.

### Table 4. Design experiments.

<table>
<thead>
<tr>
<th>case</th>
<th>( \psi_1 )</th>
<th>( \psi_2 )</th>
<th>((\psi_1^2+\psi_2^2)/2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>0.73</td>
<td>0.95</td>
<td>0.84</td>
</tr>
<tr>
<td>b</td>
<td>1.99</td>
<td>0.92</td>
<td>1.45</td>
</tr>
<tr>
<td>c</td>
<td>1.95</td>
<td>0.99</td>
<td>1.47</td>
</tr>
</tbody>
</table>

Table 5 shows the results obtained after execution of the new experimental condition in all the cases considered. In the cases a and b model 1 does not reach relative probability that is higher than 98% and the discrimination between the two rival models was not attained. On the contrary, locating the new sensor at node 3 (case c) the model 2 may be eliminated and the model 1 is chosen as the best model.
Table 5. Additional experiments.

<table>
<thead>
<tr>
<th>case</th>
<th>$D_{1,2}$</th>
<th>$P_1$</th>
<th>$P_2$</th>
<th>det$V_{\theta,1}$</th>
<th>det$V_{\theta,2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>19.93</td>
<td>96.3</td>
<td>3.7</td>
<td>$5.1 \times 10^{-28}$</td>
<td>$1.2 \times 10^{-27}$</td>
</tr>
<tr>
<td>b</td>
<td>19.67</td>
<td>91.3</td>
<td>8.7</td>
<td>$1.4 \times 10^{-28}$</td>
<td>$3.1 \times 10^{-28}$</td>
</tr>
<tr>
<td>c</td>
<td>14.61</td>
<td>99.7</td>
<td>0.3</td>
<td>$2.0 \times 10^{-28}$</td>
<td>$4.3 \times 10^{-28}$</td>
</tr>
</tbody>
</table>

From comparison among the determinants of the covariance matrix of parameter estimates reported in Tables 3 and 5, one can observe that the addition of a sensor leads to a noticeable reduction of the uncertainty for both the models. Moreover, in every cases, the model 1 gives a parameter error smaller than obtained for model 2. Nevertheless, the absolute minimum value is reached in the case b, when the additional sensor is placed at node 6.

The parameter estimation and the pressure at nodes equipped with sensors obtained after the chosen additional experiment for models 1 and 2 are reported in Table 6.

Table 6. Results after the best additional experiment.

<table>
<thead>
<tr>
<th>Model</th>
<th>$\theta_1$ (mm)</th>
<th>$\theta_2$ (mm)</th>
<th>$\theta_3$ (mm)</th>
<th>$p_1$ (m)</th>
<th>$p_1$ (m)</th>
<th>$p_2$ (m)</th>
<th>$p_2$ (m)</th>
<th>$p_3$ (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.173</td>
<td>0.774</td>
<td>0.381</td>
<td>59.595</td>
<td>55.228</td>
<td>50.190</td>
<td>50.135</td>
<td>50.393</td>
</tr>
<tr>
<td>2</td>
<td>1.144</td>
<td>0.795</td>
<td>0.604</td>
<td>59.608</td>
<td>52.912</td>
<td>50.557</td>
<td>50.509</td>
<td>50.583</td>
</tr>
</tbody>
</table>

From comparison with results reported in Table 3, one can see that, after the best additional experiment, the pipe roughness of the first group varies less than 6% from the value obtained after preliminary experiment, whereas the variation of the roughness of the other two groups is over 20%.

6. Conclusion

This work presented an approach for discriminating rival models in water distribution system analysis. The proposed procedure was largely used in the field of the chemical engineering and it was here applied to define the most suitable model among those available during the model building. This aspect is usually neglected or considered outside the scope of calibration, but it is a very important step in the model assembling to perform hydraulic simulations and to provide an accurate interpretation of the final results.

As shown in the example, the procedure allowed for discrimination of the correct model, by selecting the best additional experiments in the whole experimental field. Besides, the approach seems to be able to reduce the uncertainty in the parameter estimates.

Nevertheless, the procedure has been tested by adopting a very simple network scheme and, furthermore, only two rival models have been compared. Additional investigations are, therefore, required to check the capability of the method in the case of more complex problems.

References