Available online at www.sciencedirect.com

ScienceDirect

Procedia Engineering 186 (2017) 506 – 513

Procedia Engineering

www.elsevier.com/locate/procedia

XVIII International Conference on Water Distribution Systems Analysis, WDSA2016

Data Assimilation in Water Distribution Systems

Cristiana Bragalli[a], Matteo Fortini[a], Ezio Todini[b]*†

[a]*University of Bologna, DICAM - Via del Risorgimento 2, 40136 Bologna, Italy*
[b]*Retired Professor – University of Bologna, BiGeA, Via Zamboni, 67, 40126 Bologna, Italy*

Abstract

Operational management of water distribution networks (WDNs) requires assimilating observations such as pressures at junction nodes, flows in pipes and, whenever available, monitored demand.
Although several data assimilation techniques are today available, ranging from 1-2-3-4 DVAR to Kalman Filters, a problem is posed by the need of preserving the structural relations among state variables in a WDN, such as for instance pressure head, discharge and demand.
Ensemble Kalman Filters certainly can be used to account for non-linearities but, for example, if one tries to assimilate pressure heads and pipe flows at the same time, nothing guarantees that the resulting variables after the data assimilation step, will still obey to the hydraulic structural relations mathematically describing a WDN.
In this work an EnKF based procedure has been implemented, which allows to assimilate three types of observations, namely pressures at junction nodes, flows in pipes and monitored demand.
The procedure allows to assimilating all the observations in three successive steps, while guaranteeing the full satisfaction of the structural relations.
The results, demonstrated over an operational network, show the high performance of the chosen approach.

Peer-review under responsibility of the organizing committee of the XVIII International Conference on Water Distribution Systems

Keywords: Water Distribution Systems; Data Assimilation; Kalman Filters; Ensemble Kalman Filters.

* Corresponding author. Tel.: +39-335-6241723; fax: +39-051-6766736.
E-mail address: ezio.todini@unibo.it

Peer-review under responsibility of the organizing committee of the XVIII International Conference on Water Distribution Systems
doi:10.1016/j.proeng.2017.03.263

1. Introduction

The operational management of a water distribution network requires the availability of a reliable and properly calibrated hydraulic model of the network. It is thus necessary to set up a measurement network that allows to assimilate nodal pressures, pipe flows and, whenever available, demand measures.

Nomenclature

$\mathbf{H}$	nodal pressure heads vector
$\mathbf{H_0}$	known nodal pressure heads vector
$\mathbf{Q}$	pipe flows vector
$\mathbf{d}$	nodal demands vector
$\mathbf{d}_j$	nodal demands ensemble based on prior assumptions
$\mathbf{H}_{\vert\mathbf{d}_j}, \mathbf{Q}_{\vert\mathbf{d}_j}$	the value of state vectors $\mathbf{H}$ and $\mathbf{Q}$ given $\mathbf{d}_j$
$\mathbf{H}_{\vert\mathbf{d}_j,\mathbf{z}_H}, \mathbf{Q}_{\vert\mathbf{d}_j,\mathbf{z}_H}, \mathbf{d}_{\vert\mathbf{d}_j,\mathbf{z}_H}$	the value of state vectors $\mathbf{H}$, $\mathbf{Q}$ and $\mathbf{d}$ given $\mathbf{d}_j$ and measures $\mathbf{z}_H$
$\mathbf{H}_{\vert\mathbf{d}_j,\mathbf{z}_H,\mathbf{z}_Q}, \mathbf{Q}_{\vert\mathbf{d}_j,\mathbf{z}_H,\mathbf{z}_Q}, \mathbf{d}_{\vert\mathbf{d}_j,\mathbf{z}_H,\mathbf{z}_Q}$	the value of state vectors $\mathbf{H}$, $\mathbf{Q}$ and $\mathbf{d}$ given $\mathbf{d}_j$ and measures $\mathbf{z}_H$ and $\mathbf{z}_Q$
$\mathbf{H}_{\vert\mathbf{d}_j,\mathbf{z}_H,\mathbf{z}_Q,\mathbf{z}_d}, \mathbf{Q}_{\vert\mathbf{d}_j,\mathbf{z}_H,\mathbf{z}_Q,\mathbf{z}_d}, \mathbf{d}_{\vert\mathbf{d}_j,\mathbf{z}_H,\mathbf{z}_Q,\mathbf{z}_d}$	the value of state vectors $\mathbf{H}$, $\mathbf{Q}$ and $\mathbf{d}$ given $\mathbf{d}_j$ and measures $\mathbf{z}_H$, $\mathbf{z}_Q$ and $\mathbf{z}_d$
$\mathbf{A}_{11}$	diagonal matrix defined as in [1]
$\mathbf{A}_{12}$, $\mathbf{A}_{21}$, $\mathbf{A}_{10}$	(0,1) topological incidence matrices defined as in [1]
$\boldsymbol{\mu}_H, \boldsymbol{\mu}_Q, \boldsymbol{\mu}'_Q$	vectors of means of $\mathbf{H}_{\vert\mathbf{d}_j}, \mathbf{Q}_{\vert\mathbf{d}_j,\mathbf{z}_H}, \mathbf{Q}_{\vert\mathbf{d}_j,\mathbf{z}_H,\mathbf{z}_Q}$ estimation errors
$\mathbf{P}_H, \mathbf{P}_Q, \mathbf{P}'_Q$	variance-covariance matrices of $\mathbf{H}_{\vert\mathbf{d}_j}, \mathbf{Q}_{\vert\mathbf{d}_j,\mathbf{z}_H}, \mathbf{Q}_{\vert\mathbf{d}_j,\mathbf{z}_H,\mathbf{z}_Q}$ estimation errors
$\mathbf{K}_H, \mathbf{K}_Q, \mathbf{K}'_Q$	Kalman Gain matrices for state vectors $\mathbf{H}_{\vert\mathbf{d}_j}, \mathbf{Q}_{\vert\mathbf{d}_j,\mathbf{z}_H}, \mathbf{Q}_{\vert\mathbf{d}_j,\mathbf{z}_H,\mathbf{z}_Q}$
$\mathbf{z}_H, \mathbf{z}_Q, \mathbf{z}_d$	vectors of measurements relevant to state vectors $\mathbf{H}$, $\mathbf{Q}$ and $\mathbf{d}$
$\mathbf{M}_H$, $\mathbf{M}_Q$, $\mathbf{M}_d$	(0,1) topological matrices relating $\mathbf{z}_H$ to $\mathbf{H}$, $\mathbf{z}_Q$ to $\mathbf{Q}$ and $\mathbf{z}_d$ to $\mathbf{d}$
$\bar{\mathbf{v}}_{\mathbf{z}_H}, \bar{\mathbf{v}}_{\mathbf{z}_Q}, \bar{\mathbf{v}}_{\mathbf{z}_d}$	vectors of means of measurement errors for state vectors $\mathbf{H}$, $\mathbf{Q}$ and $\mathbf{d}$
$\mathbf{R}_{\mathbf{z}_H}, \mathbf{R}_{\mathbf{z}_Q}, \mathbf{R}_{\mathbf{z}_d}$	variance-covariance matrices of measurement errors for state vectors $\mathbf{H}$, $\mathbf{Q}$ and $\mathbf{d}$

Among the established techniques for assimilating observations [2], the Ensemble Kalman Filter (EnKF) [3], allows taking into account the non-linearity of the structural models in state extrapolation, but when using state vectors comprising simultaneously flows and pressures, the fulfillment of the structural hydraulic relations among the state variables cannot be guaranteed at the end of the data assimilation phase. In this work, it is instead presented a technique of successive conditionings, in a cascade of three EnKFs, which allows assimilating the various types of observations (pressures, flow rates and demands) while ensuring the preservation of the structural links among the resulting variables [4].

2. The proposed data assimilation procedure

To start the assimilation process using an EnKF, as in any Bayesian inferential process [5], prior knowledge is described via an appropriate probability distribution. Therefore, our prior knowledge on demands, which are scarcely known in a WDN, is assumed to be described for all the nodes by a log-normal probability distribution with mean $\boldsymbol{\mu}_{d_i}$ and variance $\boldsymbol{\sigma}^2_{d_i}$ where d_i, with $i \in \{1, \dots, n\}$, represents the demand at node i, while n is the number of nodes of the network. Demands at all nodes can be set into a state vector $\mathbf{d}$ of size n.

The assimilation process then begins with the generation an "ensemble" of m members of nodal demands, namely $\mathbf{d}_j$, with $j \in \{1, \dots, m\}$, using the assumed prior distributions. The size of the ensemble must be large enough to ensure that the matrices involved in the assimilation phase, and defined below, are invertible. Each ensemble member is subsequently transformed by means of a hydraulic network model (in this case we used EPANET-2 [6]) in the corresponding ensembles of pressures $\mathbf{H}_{|\mathbf{d}_j}$ and discharges $\mathbf{Q}_{|\mathbf{d}_j}$.

As extensively described in [4], data assimilation proceeds in three successive steps. In the first step, using the nodal pressures $\mathbf{H}$ as the state vector for the Ensemble Kalman Filter (EnKF), each member of the ensemble $\mathbf{H}_{|\mathbf{d}_j}$ is assimilated with the nodal pressure measurements $\mathbf{z}_H$, assumed to be affected by measurement errors characterized by mean $\bar{\mathbf{v}}_{\mathbf{z}_H}$and variance $\mathbf{R}_{\mathbf{z}_H}$. Posterior estimates are then obtained by combining the prior pressure heads estimates $\mathbf{H}_{|\mathbf{d}_j,\mathbf{z}_H}$ with the nodal pressure measurements $\mathbf{z}_H$, using the classical EnKF correction equation:

$$\mathbf{H}_{|\mathbf{d}_j,\mathbf{z}_H} = \mathbf{H}_{|\mathbf{d}_\mathrm{j}} + \mathbf{P}_H\mathbf{M}_H^T\left(\mathbf{M}_H\,\mathbf{P}_H\mathbf{M}_H^T + \mathbf{R}_{\mathbf{z}_H}\right)^{-1}\left(\mathbf{z}_H - \mathbf{M}_H\,\mathbf{H}_{|\mathbf{d}_j} - \bar{\mathbf{v}}_{\mathbf{z}_H}\right) \tag{1}$$

In these equations, $\mathbf{M}_H$ is the measurement topological matrix, relating the state vectors of the ensemble $\mathbf{H}_{|\mathbf{d}_j}$ to the measurements $\mathbf{z}_H$, while $\mathbf{P}_H$ is the prior variance matrix of the pressures, estimated as:

$$\mathbf{P}_H = \frac{1}{m-1}\sum_{j=1}^{m}\left[\left(\mathbf{H}_{|\mathbf{d}_j} - \boldsymbol{\mu}_H\right)\left(\mathbf{H}_{|\mathbf{d}_j} - \boldsymbol{\mu}_H\right)^T\right] \tag{2}$$

with $\boldsymbol{\mu}_H = \frac{1}{m}\sum_{j=1}^{m}\mathbf{H}_{|\mathbf{d}_j}$ the prior ensemble mean.

Conditionally to the resulting posterior estimate $\mathbf{H}_{|\mathbf{d}_\mathrm{j},\mathbf{z}_H}$ one can then update the other state variables. The pipe flows ensemble $\mathbf{Q}_{|\mathbf{d}_j,\mathbf{z}_H}$ can be estimated using one of the many equations, such as the Hazen-Williams equation, relating pressure grade to discharge, while demand estimates $\mathbf{d}_{|\mathbf{d}_j,\mathbf{z}_H}$ are obtained by summing-up confluent pipe flows $\mathbf{Q}_{|\mathbf{d}_j,\mathbf{z}_H}$ at nodes. This conditional approach allows all the updated state variables ($\mathbf{H}_{|\mathbf{d}_j,\mathbf{z}_H}$, $\mathbf{Q}_{|\mathbf{d}_j,\mathbf{z}_H}$ and $\mathbf{d}_{|\mathbf{d}_j,\mathbf{z}_H}$) to match the hydraulics structural equations describing a WDN.

The second step in the proposed data assimilation process, is performed using pipe flows $\mathbf{Q}$ as state vector and the relevant $\mathbf{z}_Q$ measures, assumed to be affected by measurement errors characterized by mean $\bar{\mathbf{v}}_{\mathbf{z}_Q}$and variance $\mathbf{R}_{\mathbf{z}_Q}$, using the ensemble $\mathbf{Q}_{|\mathbf{d}_j,\mathbf{z}_H}$ to describe the prior knowledge on pipe flows. The posterior estimates can be obtained, by combining the prior ensemble with the measurements, as:

$$\mathbf{Q}_{|\mathbf{d}_j,\mathbf{z}_H,\mathbf{z}_Q} = \mathbf{Q}_{|\mathbf{d}_\mathrm{j},\mathbf{z}_H} + \mathbf{P}_Q\mathbf{M}_Q^\mathrm{T}\left(\mathbf{M}_Q\,\mathbf{P}_Q\mathbf{M}_Q^\mathbf{T} + \mathbf{R}_{\mathbf{z}_Q}\right)^{\mathbf{-1}}\left(\mathbf{z}_Q - \mathbf{M}_Q\,\mathbf{Q}_{|\mathbf{d}_j,\mathbf{z}_H} - \bar{\mathbf{v}}_{\mathbf{z}_Q}\right) \tag{3}$$

In these equations, $\mathbf{M}_Q$ is the topological measurement matrix, which relates the state vectors of the ensemble $\mathbf{Q}_{|\mathbf{d}_j,\mathbf{z}_H}$ to the measurements $\mathbf{z}_Q$, while $\mathbf{P}_Q$ is the prior variance matrix of the pressures, estimated as:

$$\mathbf{P}_Q = \frac{1}{m-1}\sum_{j=1}^{m}\left[\left(\mathbf{Q}_{|\mathbf{d}_j,\mathbf{z}_H} - \boldsymbol{\mu}_Q\right)\left(\mathbf{Q}_{|\mathbf{d}_j,\mathbf{z}_H} - \boldsymbol{\mu}_Q\right)^T\right] \tag{4}$$

with $\boldsymbol{\mu}_Q = \frac{1}{m}\sum_{j=1}^{m}\mathbf{Q}_{|\mathbf{d}_j,\mathbf{z}_H}$ the prior ensemble mean.

Once the posterior estimates of the pipe flows $\mathbf{Q}_{|\mathbf{d}_j,\mathbf{z}_H,\mathbf{z}_Q}$ are known, using one of the many available equations relating pressure grade to discharge, such as for instance the Hazen-Williams equation, one can estimate the grades for all the pipes on the basis of which one can update nodal pressures $\mathbf{H}_{|\mathbf{d}_j,\mathbf{z}_H,\mathbf{z}_Q}$ starting from known pressure head $\mathbf{H}_0$ at one (or more) fixed head nodes. Demands $\mathbf{d}_{|\mathbf{d}_j,\mathbf{z}_H,\mathbf{z}_Q}$, are also easily obtained by summing-up confluent pipe flows at nodes. Again, by this procedure all the posterior state variables $\mathbf{H}_{|\mathbf{d}_j,\mathbf{z}_H,\mathbf{z}_Q}$, $\mathbf{Q}_{|\mathbf{d}_j,\mathbf{z}_H,\mathbf{z}_Q}$ and $\mathbf{d}_{|\mathbf{d}_j,\mathbf{z}_H,\mathbf{z}_Q}$ will be constrained to match the hydraulic structural equations describing a WDN.

At this point, whenever demand measurements $\mathbf{z}_d$ are also available one can proceed to their assimilation. Demand measurements $\mathbf{z}_d$ are also assumed to be affected by measurement errors characterized by mean $\bar{\mathbf{v}}_{\mathbf{z}_d}$ and variance $\mathbf{R}_{\mathbf{z}_d}$. To overcome the problems related to the non-Gaussian behavior of the demand state variables, which do not admit negative values, the data assimilation can be carried out using an "inverted" filter, which projects back on pipe flows, the innovation calculated in terms of demand. This was made possible by extending to the EnKF the works of Chou [7] and Basseville et al. [8], who, using a standard KF, adapted the "smoothing" algorithm of [9] to assimilate observations recorded at different scales.

Then, using the linear relationship between the flow and demand $\mathbf{d}_{|\mathbf{d}_j,\mathbf{z}_H,\mathbf{z}_Q} = \mathbf{A}_{21}\mathbf{Q}_{|\mathbf{d}_j,\mathbf{z}_H,\mathbf{z}_Q}$ where $\mathbf{A}_{21}$ indicates the nodes-pipes incidence matrix as defined in [1], one can write the following relation using once again pipe flows as the state vector:

$$\mathbf{Q}_{|\mathbf{d}_j,\mathbf{z}_H,\mathbf{z}_Q,\mathbf{z}_d} = \mathbf{Q}_{|\mathbf{d}_j,\mathbf{z}_H,\mathbf{z}_Q} + \mathbf{P}'_Q\mathbf{A}_{12}\mathbf{M}_d^{\mathrm{T}}\left(\mathbf{M}_d\,\mathbf{A}_{21}\mathbf{P}'_Q\mathbf{A}_{12}\mathbf{M}_d^{\mathrm{T}} + \mathbf{R}_{\mathbf{z}_d}\right)^{-1}\left(\mathbf{z}_d - \mathbf{M}_d\,\mathbf{d}_{|\mathbf{d}_j,\mathbf{z}_H,\mathbf{z}_Q} - \bar{\mathbf{v}}_{\mathbf{z}_d}\right) \tag{5}$$

Equation (5) can also be rewritten in the following form:

$$\mathbf{Q}_{|\mathbf{d}_j,\mathbf{z}_H,\mathbf{z}_Q,\mathbf{z}_d} = \mathbf{Q}_{|\mathbf{d}_j,\mathbf{z}_H,\mathbf{z}_Q} + \mathbf{P}'_Q\mathbf{A}_{12}\mathbf{M}_d^{\mathrm{T}}\left(\mathbf{M}_d\,\mathbf{A}_{21}\mathbf{P}'_Q\mathbf{A}_{12}\mathbf{M}_d^{\mathrm{T}} + \mathbf{R}_{\mathbf{z}_d}\right)^{-1}\left(\mathbf{z}_d - \mathbf{M}_d\,\mathbf{A}_{21}\mathbf{Q}_{|\mathbf{d}_j,\mathbf{z}_H,\mathbf{z}_Q} - \bar{\mathbf{v}}_{\mathbf{z}_d}\right) \tag{6}$$

In Equations (5) and (6) , $\mathbf{M}_d$ is the topological measurement matrix, relating the state vectors of the ensemble $\mathbf{d}_{|\mathbf{d}_j,\mathbf{z}_H,\mathbf{z}_Q}$ with the measurements $\mathbf{z}_d$, while $\mathbf{P}'_Q$ is the prior variance matrix of pipe flows $\mathbf{Q}_{|\mathbf{d}_j,\mathbf{z}_H,\mathbf{z}_Q}$, given by:

$$\mathbf{P}'_Q = \frac{1}{m-1}\sum_{j=1}^{m}\left[\left(\mathbf{Q}_{|\mathbf{d}_j,\mathbf{z}_H,\mathbf{z}_Q} - \boldsymbol{\mu}'_Q\right)\left(\mathbf{Q}_{|\mathbf{d}_j,\mathbf{z}_H,\mathbf{z}_Q} - \boldsymbol{\mu}'_Q\right)^{\mathrm{T}}\right] \tag{7}$$

with $\boldsymbol{\mu}'_Q = \frac{1}{m}\sum_{j=1}^{m}\mathbf{Q}_{|\mathbf{d}_j,\mathbf{z}_H,\mathbf{z}_Q}$ the prior ensemble mean.

Once $\mathbf{Q}_{|\mathbf{d}_j,\mathbf{z}_H,\mathbf{z}_Q,\mathbf{z}_d}$ is known one can update the estimates of $\mathbf{H}_{|\mathbf{d}_j,\mathbf{z}_H,\mathbf{z}_Q,\mathbf{z}_d}$ and $\mathbf{d}_{|\mathbf{d}_j,\mathbf{z}_H,\mathbf{z}_Q,\mathbf{z}_d}$ similarly to what done in the previous step, in order to preserve the structural congruence between the variables and calculate the relevant updated means and variance-covariance matrices.

In essence, using a cascade of three EnKF, each of which with a different state vector, instead of using a single EnKF with a unique state vector incorporating all state variables $\mathbf{H}, \mathbf{Q}$ and $\mathbf{d}$, it is not only possible to assimilate all the alternative types of available measures, but also to preserve the structural links among the different sets of state variables describing a WDN.

3. The case study

3.1 Description of the water distribution network under study

Modena network consists in 317 pipes to design using 13 possible diameters. Required pressure is 20 $[m]$ for all 268 demand nodes. The network is fed by 4 reservoirs (Figure 1), also the network is completely looped with diameters ranging between 100 to 400 $[mm]$ [10]. The Modena WDN can be found as MOD-1 in "http://emps.exeter.ac.uk/engineering/research/cws/resources/benchmarks/design-resiliance-pareto-fronts/large-problems/" where a full description is also available.

3.2 Effectiveness of the data assimilation procedure

The effectiveness of the proposed data assimilation procedure to reduce uncertainty over the estimated pressure heads, flow in pipes and demands, was assessed by randomly selecting a measurement network providing 100 pressure measurements $\mathbf{z}_H$, 100 flow measurements $\mathbf{z}_Q$, and 100 demand measurements $\mathbf{z}_d$

In order to express a rather large prior uncertainty, the prior distribution on demands $\mathbf{d}$ was assumed to be a Log-normal distribution with constant mean and variance at all nodes. The mean $\boldsymbol{\mu}_d = 1.57\ [l\ s^{-1}]$, was estimated as the total released water from the reservoirs divided by the number of delivery nodes, while the standard deviation, expressing a large prior uncertainty on demands was assumed $\boldsymbol{\sigma}_d = 1.0\ [l\ s^{-1}]$. From this Log-normal distribution, an ensemble of $m = 500$ members of demands for all 268 demand nodes were generated.

Table 1. Total Variance and Standard Deviation for initial and posterior estimates

	TV{H}	TSD{H}	TV{Q}	TSD{Q}	TV{d}	TSD{d}
$\|\mathbf{d}$ **(Prior)**	49.807	7.058	13.747	3.708	2.461	1.569
$\|\mathbf{d}, \mathbf{z}_H$	0.179	0.423	1.945	1.395	3.037	1.743
$\|\mathbf{d}, \mathbf{z}_H, \mathbf{z}_Q$	0.049	0.221	0.787	0.887	1,616	1.271
$\|\mathbf{d}, \mathbf{z}_H, \mathbf{z}_Q, \mathbf{z}_d$	0.019	0.138	0.249	0.499	0.356	0.597

Without loss of generality, the measurement sensors where considered to provide unbiased measurements with standard errors of:

$$SE(H) = \pm 0.01\ [m];\quad SE(Q) = \pm 0.03\ [l\ s^{-1}];\ SE(q) = \pm 0.1\ [l\ s^{-1}]$$

and, although accounting for measurement uncertainty in the DA process, the "true" value was used as the measured one, because if measurements are unbiased, their expectation coincides with the true value.

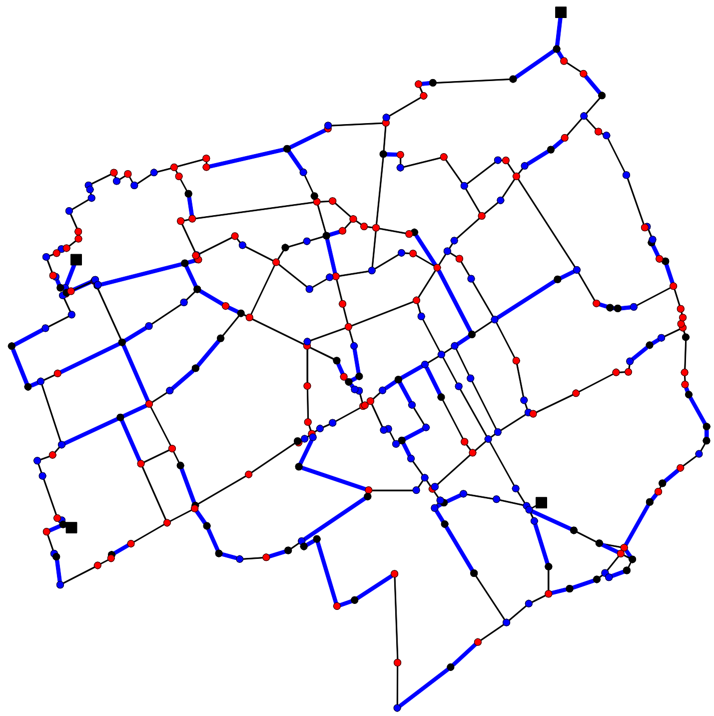

Fig. 1. Schematic representation of the 317 pipes and 268 nodes Modena WDN and positioning of sensors : pressure sensors are represented as blue dots, flow meters as blue segments and nodal demand aggregated measures as red dots. The four reservoirs are shown as black squares.

The results of the data assimilation process are summarized in Table 1 and Figure 2. Table 1 shows the Total Variance (TV) of the prior estimates and of the posterior ones after each of the three successive data assimilation steps. The TV [11], [12] is defined as the sum of two terms: the first one, measures the variance of errors given the estimated ensemble mean, while the second term, measures the average variance of the ensemble mean. We also introduced the Total Standard Deviations (TSD), defined as the square roots of the TV, and used in Table 1 to provide a measure of the standard errors in dimensional form. From the values of TSD it is possible to realize that after assimilating all the observations, the standard error on pressure heads from 7 $[m]$ falls around 14 $[cm]$, while the standard error on pipe flows reduces from 3.7 $[l\ s^{-1}]$ to $\sim$0.5 $[l\ s^{-1}]$ and on demands from $\sim$1.6 $[l\ s^{-1}]$ to $\sim$0.6 $[l\ s^{-1}]$.

As can be additionally seen in Table 1, the reduction of total variance from the prior assumption $(|\mathbf{d})$ is already significant both in terms of pressure head and pipe flow from the assimilation of the head measurements $\mathbf{z}_H$ $(|\mathbf{d}, \mathbf{z}_H)$. This does not happen for the demands which also require, in this case, the assimilation of the additional flow measurements $\mathbf{z}_Q, \mathbf{z}_d$. Nonetheless, after assimilating all the observations $\left(|\mathbf{d}, \mathbf{z}_H, \mathbf{z}_Q, \mathbf{z}_d\right)$, the final result is quite satisfactory with a reduction of over 99% of the TV for pressure heads, 98% reduction for pipe flows and 62% for demands.

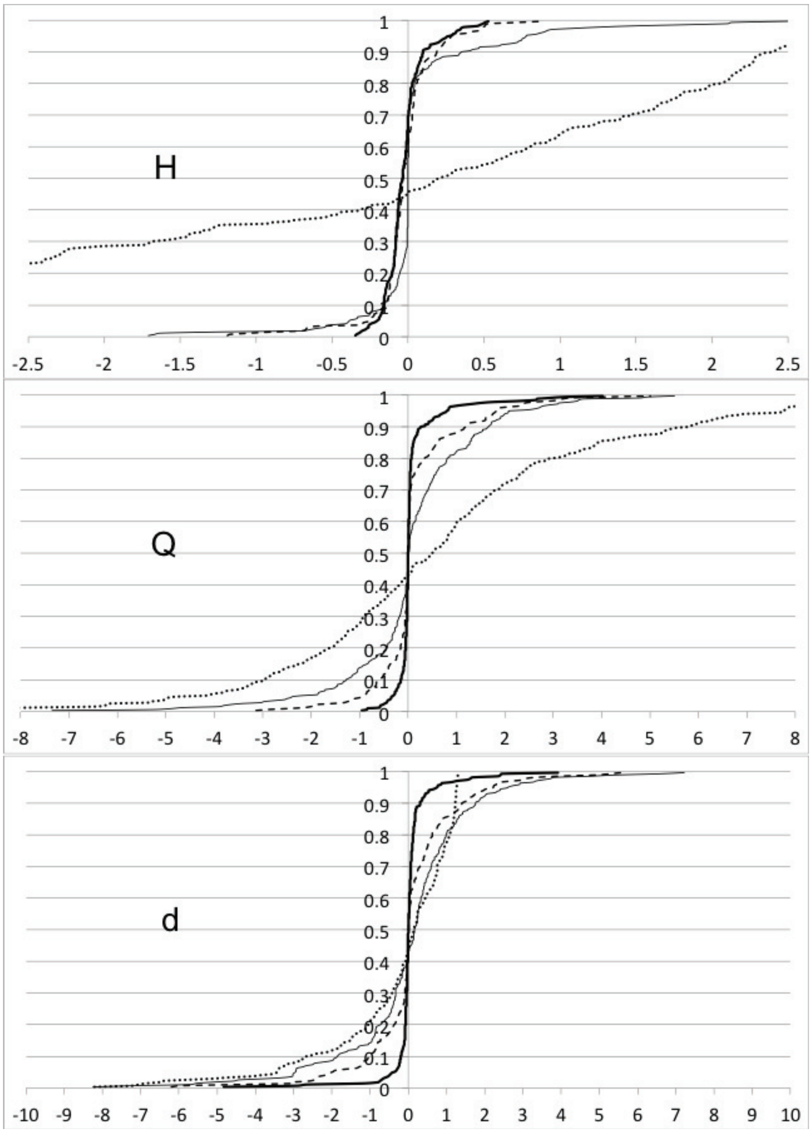

Fig. 2. Prior and posterior empirical distributions of the estimation errors for the three sets of state variables $\mathbf{H}, \mathbf{Q}$ and $\mathbf{d}$ after assimilating in successive steps, as described in Section 2, measurements from 100 pressure sensors $\boldsymbol{z}_H$, 100 flow meters $\boldsymbol{z}_Q$ and 100 demands $\boldsymbol{z}_d$ from the

assumed, randomly generated measurement network. The figure compares the prior distribution although limited to the range of the figures (dotted line); the empirical distribution of errors after assimilating the 100 pressure measurements (thin solid line); the empirical distribution of errors after assimilating both the 100 pressure measurements and the 100 pipe flow measurements (dashed line); the empirical distribution of errors after assimilating all the measurements (thick solid line).

These results can be better appreciated in Figure 2 where the prior and the posterior empirical distribution of the estimation errors, defined as the difference between the ensemble mean and the true value, which is used to assess the performances of the approach, but is assumed unknown in the DA process.

In Figure 2, the dotted line represents the portion of the prior empirical distribution of prediction errors $\left(|\mathbf{d}_j\right)$ falling in the range of the figure. The solid thin line shows the empirical distribution of errors after assimilating the 100 pressure measurements $\left(|\mathbf{d}_j, \mathbf{z}_H\right)$; the dashed line shows the empirical distribution of errors after assimilating both the 100 pressure measurements and the 100 pipe flow measurements $\left(|\mathbf{d}_j, \mathbf{z}_H, \mathbf{z}_Q\right)$ and the solid thick line shows the empirical distribution of errors after assimilating all the measurements including the 100 demand measurements $\left(|\mathbf{d}_j, \mathbf{z}_H, \mathbf{z}_Q, \mathbf{z}_d\right)$.

It can be immediately perceived from Figure 2 that the effect of DA increases the sharpness of the posterior distributions. The successive distributions are more and more cantered on the zero, thus reducing biases, and the percentage of values within a limited range increases. Moreover, it is interesting to underline that the successive conditional data assimilation steps have positive backward effects on the other state variables. In other words, the assimilation of pipe flow observations improves both demands and pressure heads estimates while the assimilation of demand measurements improves both the estimates of flow in pipes and pressure heads.

These results confirm the validity of the new conditional DA approach in which each set of measurements is assimilated in relation to the corresponding state variables instead of operating at once using one EnKF on a single extended state vector. The approach, which benefits are primarily in the preservation of the structural relations among the variables, also allows for faster DA in that one solves three problems of reduced dimensionality instead of a single three times larger problem.

4. Conclusions

In this work, we presented a new approach to assimilate the information originating from a composite monitoring network consisting in pressure head, pipe flow and demand meters into prior uncertain knowledge of demands, expressed in the form ensemble. The DA is performed using a cascade of three successive EnKF, to ensure the preservation of structural hydraulics ties between the state variables all throughout the assimilation process. The approach, which shows great efficiency and allows reducing the uncertainty of all the state variables, can be viewed as the first essential step when aiming at calibrating WDN model parameters and at identifying the probability and location of leak losses.

References

[1] E. Todini and S. Pilati, A gradient algorithm for the analysis of pipe networks. In B. Coulbeck and C.-H. Orr (eds), Computer Applications in Water Supply: Vol. 1— Systems Analysis and Simulation, pp. 1–20. Research Studies Press Ltd., Letchworth, Hertfordshire, UK, 1988.

[2] C. Hutton, Z.S. Kapelan, L. Vamvakeridou-Lyroudia, D.A. Savić, Dealing with Uncertainty in Water Distribution System Models: A Framework for Real-Time Modeling and Data Assimilation. *J. Water Resour. Plann. Manage*., 140(2) (2014) 169–183.

[3] G. Evensen, The Ensemble Kalman Filter: theoretical formulation and practical implementation. *Ocean Dynamics*, 53 (2003) 343–367.

[4] C. Bragalli, M. Fortini, E. Todini, Enhancing Knowledge in Water Distribution Networks via Data Assimilation. Published on-line Water Resources Management (2016) 1-18. DOI 10.1007/s11269-016-1372-0

[5] J.M. Bernardo, A.F.M. Smith, *Bayesian Theory*. Wiley, Chichester (UK). ISBN 0-471-92416-4. MR 1274699, 1994.

[6] L.A. Rossman, *EPANET 2 users manual*. Rep. No. EPA/600/ R-00/057, Water Supply and Water Resources Division, National Risk Management Research Laboratory, Cincinnati, 2000.

[7] K.C. Chou, *A stochastic approach to multiscale signal processing*. PhD dissertation, Massachusetts Institute of Technology. http://ssg.mit.edu/ssg_theses/ssg_theses_1974_1999/ Chou_PhD_5_91.pdf., 1991.

[8] M. Basseville, A. Benveniste, K.C. Chou, S.A. Golden, R. Nikoukhah, A. Willsky. Modeling and Estimation of Multiresolution Stochastic Processes, *IEEE Transactions on Information Theory*, 38(2) (1992) 766–784.

[9] H. Rauch, F. Tung, C. Striebel Maximum likelihood estimates of linear dynamic systems. *AIAA J.* 3(8) (1965) 1445-1450.

[10] S. Artina, A. Bolognesi, C. Bragalli, C. D'Ambrosio, A. Marchi, Comparison among best solutions of the optimal design of water distribution networks obtained with different algorithms. *CCWI 2011 Urban Water Management: challenges and opportunities* 3 (2011) 985-990, Exeter, ISBN: 0953914089.
[11] H.C. Mahler, C.G. Dean, Chapter 8: Credibility In *Foundations of Casualty Actuarial Science (4th ed.)*. Casualty Actuarial Society. pp. 525–526. ISBN 978-0-96247-622-8, 2001.
[12] N. A. Weiss, *A Course in Probability*, Addison–Wesley, Boston, pp. 385–386, 2006.