



EPiC Series in Engineering

Volume 3, 2018, Pages 335–342

HIC 2018. 13th International
Conference on Hydroinformatics



Spectral Propagation of Uncertainty in Water Age

Mathias Braun^{1,2*}, Olivier Piller¹, Jochen Deuerlein³, Iraj Mortazavi⁴,
Angelo Iollo^{2,5}

¹Irstea, UR ETBX, Dept. of Water, Cestas, France

²INRIA Bordeaux-Sud-Ouest, Team MEMPHIS, France

³3S Consult GmbH, Albtalstrasse 13, 76137 Karlsruhe, Germany

⁴CNAM, EA-7340-M2N-Modelisation Mathematique et Numerique, 75003 Paris, France

⁵Universite Bordeaux, IMB, 33400 Talence, France

mathias.braun@irstea.fr

Abstract

Water distribution networks are critical infrastructures that should ensure the reliable supply of high quality potable water to its users. Numerical models of these networks are generally governed by many parameters for which the true value is unknown. This may be due to a lack of knowledge like for consumer demand or due to a lack of accessibility as for the pipe roughness. For network managers, the effect of these uncertainties on the network state is important information that supports them in the decision-making process. This effect is generally evaluated by propagating the uncertainties using the mathematical model. In the past, perturbation and stochastic collocation methods have been used for uncertainty propagation. However, these methods are limited either in the accuracy of the results or the complexity of the calculation. This paper uses an alternative spectral approach with the polynomial chaos expansion that has the potential to give comparable results to the Monte Carlo sampling through the definition of a stochastic model. This approach is applied to the model of a water distribution network in order to evaluate the influence of uncertain demands on the water age.

1 Introduction

Evaluating the quality of drinking water supplied by a distribution system is usually done by using quality indicators like water age or chlorine concentration. These indicators may be determined from

* Corresponding author.

mathematical network models. These depend on a great number of parameters like water demand, pipe roughness and reservoir levels. However, these parameters are rarely known exactly and thus introduce uncertainty into the model.

Uncertainties in mathematical models are generally divided into aleatory and epistemic uncertainties. Aleatory uncertainties are the result of inherently random processes and are often described by probability distributions that are modelled on observations. These uncertainties are regarded as irreducible since the underlying process cannot be influenced. In contrast, epistemic uncertainties follow from incomplete insight into the system, which in theory could be eliminated through perfect knowledge. However, epistemic uncertainties are often the necessary effect of simplifications in the modelling process and the numerical evaluation.

Parameter uncertainties are one of the most important aspects of hydraulic modelling as they can result in large prediction errors. Hidden parameters like the pipe roughness that cannot be observed directly are often estimated through calibration (Savic et al. (2009), Piller et al. (2010)). But, even for a well calibrated network some uncertainty remains. Other parameters like demand are clearly aleatory in nature.

The objective of Uncertainty Analysis (UA) is the evaluation and mitigation of such parameter uncertainties on the Quantities of Interest (QoI). Central part of this process is the uncertainty propagation by means of the mathematical model. Classical approaches contain First Order Second Moment (FOSM) methods and Monte Carlo sampling. However, the FOSM is limited by the fact that it linearizes the system equation (Cacuci et al. (2005)) and will produce symmetrical confidence interval (Vrachimis et al. (2016)). Monte Carlo sampling does not suffer from these limitations, but for models with many parameters it is constrained by the curse of dimensionality. This paper proposes an alternative approach to the uncertainty propagation with the Polynomial Chaos Expansion (PCE) (Smith (2013), Xiu (2010)) Depending on its application the PCE has the potential to greatly reduce or even eliminate the sampling by use of the mathematical model.

The objective of this paper is to analyse the influence of the demand uncertainties on the water age in a network model that has been provided by the partner Veolia in the framework of the French-German project ResiWater (2018). Following this introduction, the example network will be introduced together with the basic equations for the hydraulic state and water age in section 2. This is followed by a short introduction of the PCE in section 3. Section 4 presents some results of the investigation and section 5 closes with a discussion of the results.

2 Material

2.1 Hydraulic Model

In hydraulic modelling, the simplified topological structure of a water distribution network is described by a directed graph corresponding to **Error! Reference source not found.**Figure 1. In this graph, links represent pipe sections, valves and pumps, and nodes the resource nodes, tanks, demand nodes and connections. The mathematical description of this graph is given by the node-link incidence matrix $\mathbf{A} \in M^{nj \times np}$, where nj is the number of junction nodes and np is the number of links. The coefficients are defined as follows:

$$A_{i,j} = \begin{cases} -1, & \text{if node } i \text{ is the end node of link } j \\ 0, & \text{if link } j \text{ is not connected to node } i \\ +1, & \text{if node } i \text{ is the start node of link } j. \end{cases} \quad (1)$$

Water distribution networks in general have a looped structure and the system state is described by the potentials at the nodes (heads) and the currents on the links (flow rates). The hydraulic model is defined by two sets of equations. First, the mass balance at the nodes:

$$\mathbf{A}\mathbf{q} + \mathbf{d} = \mathbf{0}_{nj} \quad (2)$$

where the node-link incidence matrix \mathbf{A} includes only nodes with unknown heads, $\mathbf{q} \in {}^{np}$ is the vector of link flow rates and $\mathbf{d} \in {}^{nj}$ is the vector of demands at consumption nodes. Second, the energy equation:

$$\Delta\mathbf{h}(\mathbf{r}, \mathbf{q}) - \mathbf{A}^T \mathbf{h} - \mathbf{A}_f^T \mathbf{h}_f = \mathbf{0}_{np} \quad (3)$$

where \mathbf{A}_f describes the incidence matrix of nodes with fixed potential like reservoirs or tanks and $\mathbf{h} \in {}^{nj}$ is the vector containing the unknown piezometric heads. Parameters are given by the head vector $\mathbf{h}_f \in {}^{nf}$ describing fixed heads at special nodes like reservoirs or tanks and the vector $\mathbf{r} \in {}^{np}$ containing the friction coefficients for each link. The function $\Delta\mathbf{h}(\mathbf{r}, \mathbf{q})$ describes the loss in head along a pipe.

2.2 Water Quality Indicators

The quality indicators are based on the convection-diffusion equation:

$$\frac{\partial c}{\partial t} = \nabla(D\nabla c) - \nabla(vc) + R(t, x) = 0, \quad (4)$$

where c is the variable of interest, D is the diffusion coefficient, v the velocity field and R a source term which may be a non-linear function in time and space. For water distribution systems two major simplifications are made. First, it is assumed an incompressible water flow, for which extend period simulation for the hydraulic is accurate enough (no water hammer neither mass oscillation); it follows the hydraulic model is one-dimensional and the velocity vector v simplifies to $U(t, x) = U_{i,j}$ which is calculated for every hydraulic time step t_i and constant along x for each pipe p_j . Second, it is assumed that the regime is turbulent as a consequence the longitudinal diffusion can be neglected as in (Rossman & Boulos (1996), Fabrie et al. (2010)). With these adaptations, the transport equation for water age A may be derived by replacing R with a zero-order source term

$$\frac{\partial A(t, x)}{\partial t} + U_{i,j} \frac{\partial A(t, x)}{\partial x} - 1 = 0. \quad (5)$$

During each hydraulic time step the water age is resolved by a succession of quality time steps. The equation is solved on the domain $x \in [0, L_j]$ for pipe p_j , where L_j is the pipe length and for the time $t \in [t_i, t_{i+1}]$. The initial condition $A(t_i, x_j)$ is defined by the previous time step. The boundary conditions are updated with the perfect mixing assumption for the average water age, or with the min or max value entering.

In the software Porteau (Piller et al. (2011)) water age is evaluated in three variables. Since Porteau keeps track of the minimum and maximum water age when water mixes at a junction, it is possible to evaluate the extreme values for the residence time. The third variable, average water age, is calculated by the weighted sum of the water age in the water columns mixing at the upstream junction.

2.3 Network Model

The model of the network used in this paper is shown in **Error! Reference source not found.**. The network has **Error! Reference source not found.** one loop and three reservoirs with a fixed head. The hydraulic system given and (3) have a total number of 21 degrees of freedom given by 9 nodes in equation (2) and 12 pipes in equation (3).

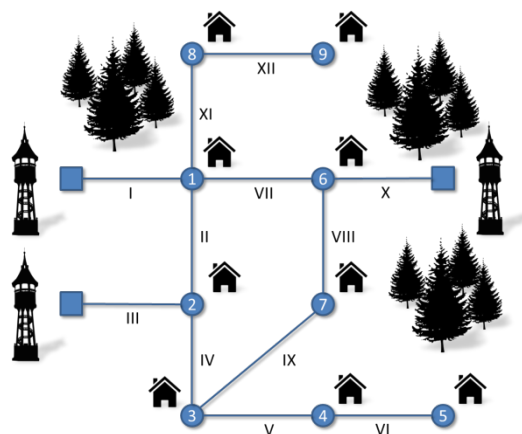


Figure 1: Small looped network.

3 Uncertainty Propagation

The objective of Uncertainty Analysis (UA) is to evaluate the impact of uncertain parameters on the Quantities of Interest (QoI). To do so, the parameters and QoI are modelled as random variables. The central part of UA is the propagation of errors and uncertainties by means of the mathematical model. Classical approaches contain perturbation and stochastic sampling methods. Stochastic spectral methods are a relatively new addition to the field that has been introduced by Xiu (2010).

Perturbation methods calculate the moments for the distribution of the quantity of interest directly from the system equations by means of a truncated Taylor expansion. Typically, the expansions employed are limited to first- or second-order expansions (Cacuci et al. (2005)). This limits their accuracy for highly non-linear models. Stochastic sampling methods like the Monte Carlo Simulation are some of the most prominent algorithms for uncertainty propagation. This is due to the easy implementation and the ability to be used for non-linear systems. However, it can be computationally demanding as its rate of convergence is proportional to $1/\sqrt{M}$, where M is the number of simulations (Cacuci et al. (2005), Smith (2013)). The objective of stochastic spectral methods like stochastic Galerkin and Non-Intrusive Spectral Projection (NISP) is the calculation of a spectral representation of the random variables. Utilizing the smoothness requirement of the polynomial basis leads to a more efficient convergence behavior (Smith (2013), Xiu (2010)).

For the Polynomial Chaos Expansion a random variable X is expressed as a truncated spectral series expansion of the order N based on a basic random variable Z and a set of orthogonal basis polynomials Ψ_k .

$$X_N(Z) = \sum_{k=0}^N x_k \Psi_k(Z) \quad (6)$$

The choice of the basic random variable, also often termed as the germ, and the orthogonal polynomials is joint by the definition of the inner product:

$$E[\Psi_i(Z)\Psi_j(Z)] = \int \Psi_i(Z)\Psi_j(Z)dP_Z(z) = \int \Psi_i(z)\Psi_j(z)f_Z(z)dz = \delta_i \|\Psi_i\|^2. \quad (7)$$

Here, E defines the expected value, P_Z is the cumulative distribution function and f_Z is the probability density function of the germ distribution. The Kronecker delta δ_i is equal to 1 if the indices are equal and 0 for the product of two different polynomials. The coefficients of a random parameter X are calculated by projecting it on the orthogonal basis polynomials:

$$x_k = \frac{\langle X, \Psi_k \rangle}{\|\Psi_k\|^2}. \quad (8)$$

As Equation (7) suggests, the choice of the polynomial basis generally depends on the distribution of the basic random variable. In this paper uncertainties are assumed to follow a Gaussian distribution which is efficiently approximated by the Hermite polynomials. However, the spectral expansion allows to approximate non-Gaussian random variables in the series expansion. Further there exist a number of random distributions that are approximated efficiently by specific sets orthogonal polynomials [6].

Application of the Polynomial Chaos Expansion in general follows an intrusive or a non-intrusive procedure. The intrusive approach requires a reformulation of the system equations and allows for the direct solution in the parameters of the QoI. Although computationally very efficient it may pose certain complications since it is not possible to use existing implementations for the solution of the problem. Further, the handling of nonlinearities may require additional steps that introduce additional errors. Non-intrusive algorithms use a number of samples similar to the Monte Carlo simulation to evaluate the coefficients. This means that in regard to the computation current software and models may be used. Even though the non-intrusive PCE is more efficient than the MCS for high dimensional parameter spaces the problem may become computationally infeasible for applications with high dimensional parameter spaces.

The *non-intrusive* method is a stochastic collocation like Monte Carlo simulation. These methods use a number of M evaluations of the full mathematical model for a limited number of realizations from the random parameters in order to build a surrogate model that is fitted at these collocation points. In contrast to the MCS, the direct evaluations in intrusive methods are used to fit the coefficients of the Polynomial Chaos Expansion to the data. For the vector of QoI for the hydraulic propagation is given by $\mathbf{x} = [\mathbf{q}^T \mathbf{h}^T]^T$, whereas for the water age it is defined as $\mathbf{x} = \mathbf{A}$. The coefficients are estimated by solving the following equation.

$$\begin{bmatrix} \Psi_0(\mathbf{z}^{(1)}) & \Psi_N(\mathbf{z}^{(1)}) \\ \Psi_0(\mathbf{z}^{(M)}) & \Psi_N(\mathbf{z}^{(M)}) \end{bmatrix} \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_N^T \end{bmatrix} = \begin{bmatrix} \mathbf{x}(\mathbf{y}(\mathbf{z}^{(1)}))^T \\ \mathbf{x}(\mathbf{y}(\mathbf{z}^{(M)}))^T \end{bmatrix} \quad (9)$$

The right-hand side is given by the M result vectors $\mathbf{x}(\mathbf{y}(\mathbf{z}^{(i)}))$ of the mathematical model. The left matrix on the left-hand side is formed by the value of the N basis polynomials at the M realizations of the basic random variable. The columns of the second matrix on the left-hand contain the coefficients for the PCE of one of the QoI. Further, $\mathbf{z}^{(i)}$ are realizations of the basic random variable \mathbf{Z} and $\mathbf{y}(\mathbf{z}^{(i)})$ is the vector of realizations of the random model parameters based on the realization of \mathbf{Z} .

4 Results

The setup that is investigated in this paper is defined as follows. The model defines a diurnal demand pattern for a selection of free nodes. This pattern contains a series of demand multipliers. For each hydraulic time-step the according demand factor is multiplied by the base demand that is defined individually for each consumption node. Through this behavioural aggregation a large portion of the network is influenced by a concentrated number of variables. Uncertainty is introduced into the system by an uncertainty multiplier that is applied to the demand pattern. This uncertainty is assumed to be normally distributed with a mean of 1 so that the predefined pattern is equal to the expected value and a standard deviation of 0.1 which in practice ensures that the demand stays positive. From this distribution a sample of 1000 elements is taken to modify the parameters for the Monte Carlo simulation. From the ensemble of these simulations a reduced number of 100 elements are taken to evaluate the Polynomial Chaos Expansion. In the following result plots, the Monte Carlo Simulation is visualized by a histogram. The PCE is illustrated by a red dashed line based on kernel density estimation.

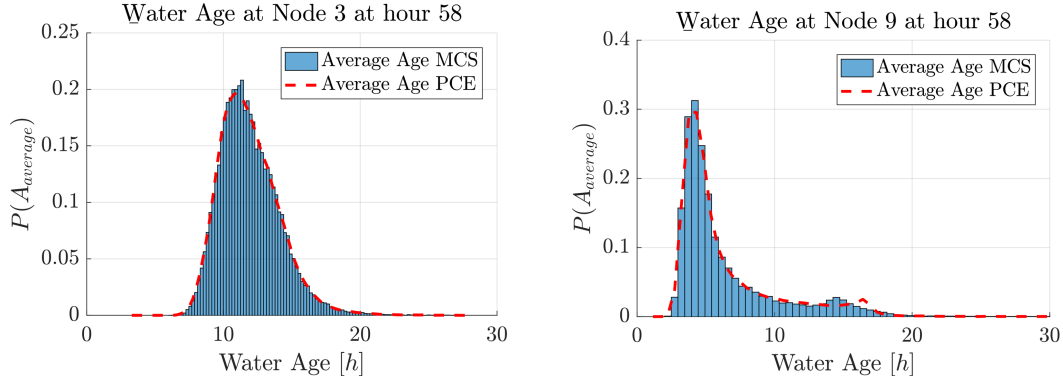


Figure 2: Water age distributions for small network: (a) Node 3 in the meshed part of the system (b) at node 9 at a branch of the network.

Figure 2 shows the probability distribution for two nodes in the small network. Figure 2 (a) gives the distribution at node 3 and Figure 2 (b) the one for node 9. The evaluation is done after a simulated duration of 58 hours at which point the behaviour of the water age has become periodic and is no

longer influenced by the initial condition. The values of the mean values and standard deviations are given in Table 1 along with the expansion order. In the evaluation of the results it has become apparent that for the small network no significant difference could be observed between the minimum, maximum and average water age. The behaviour at the two nodes is very different. The distribution at node 3, while clearly being nonlinear, has still a single peak and a limited standard deviation. This observation is supported by the coefficients in Table 2 which decrease rapidly after the second coefficient. In contrast the distribution at node 9 shows two peaks and the expansion coefficients decrease more slowly. This means that the probable water age in the loop node 3 is higher than the one in the tree node 9 however due to the lower standard deviation the level of uncertainty is lower at node 3. The development of the two peaks in Figure 2 (b) suggests that, depending on the demand multiplier, the flow structure changes and water arrives from a different node, takes a longer route or flows more slowly. Remixing of old water is unlikely due to the structure of the network with three reservoirs that are connected to the loop. A detailed analysis of the origin reservoir has not been performed as a part of this study, but may offer better insight.

| | LTB /100 | Average water age A_{avg} |
|--------|--------------------|-----------------------------|
| Node 3 | Mean | $\mu_3 = 8.1$ |
| | Standard deviation | $\sigma_3 = 4.7$ |
| | Expansion order | $O_3 = 7$ |
| Node 9 | Mean | $\mu_9 = 3.3$ |
| | Standard deviation | $\sigma_9 = 7.8$ |
| | Expansion order | $O_9 = 9$ |

Table 1: Stochastic moments and PCE order for the probability distribution function of average water age at nodes 3 and 9 of the small network.

| Order N | a_0 | a_1 | a_2 | a_3 | a_4 | a_5 | a_6 | a_7 | a_8 | a_9 |
|-----------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| Node 3 | 11.9 | -2.12 | 0.40 | 0.011 | 0.05 | 0.09 | 0.03 | 0.04 | | |
| Node 9 | 6.52 | -3.67 | 2.10 | -1.52 | -0.03 | -1.70 | 0.00 | -1.96 | 0.06 | -0.59 |

Table 2: The expansion coefficients for the polynomial chaos expansions of average water age at nodes 3 and 9 of the small network.

5 Conclusion

The polynomial chaos expansion has been applied to a small water distribution system to evaluate the effect of demand uncertainties on the resident time of water in the network. It has been shown that the Polynomial Chaos Expansion gives comparable results to Monte Carlo simulations while reducing the number of evaluations of the full system by an order of magnitude. This means that for a fixed number of simulations the PCE performs considerably better. Even for the relatively small normally distributed perturbation of the demand used in this paper the probability distribution for the water age is far from normal. This means that perturbation methods are not suitable for this task. Further it has been shown that the size of the network has a considerable influence on the maximum water age.

Acknowledgements

The work presented in the paper is part of the French-German collaborative research project ResiWater that is funded by the French National Research Agency (ANR; project: ANR-14-PICS-0003) and the German Federal Ministry of Education and Research (BMBF; project: BMBF-13N13690).

References

- Savic D. A., Kapelan Z. S., Jonkergouw P. M. (2009). *Quo vadis water distribution model calibration?*. Urban Water Journal 6, (1) 3-22.
- Piller et al. (2010) [2] Piller O., Gilbert D., Van Zyl J. E. (2010). *Dual calibration for coupled flow and transport models of water distribution systems*. Water Distribution Systems Analysis, 722-731.
- Cacuci D. G., Ionescu-Bujor M., Navon I. M. (2005). *Sensitivity and uncertainty analysis, volume II: applications to large-scale systems*. Vol. 2, CRC press.
- Vrachimis, S. G., Eliades D. G., and Polycarpou M. M. (2016). *Real-time hydraulic interval state estimation for water transport networks: A case study*. 14th International computing and control for the water industry (CCWI) conference. Amsterdam, Netherlands.
- Smith R. C. (2013). *Uncertainty quantification: theory, implementation, and applications*. Vol. 12, SIAM.
- Xiu D. (2010). *Numerical methods for stochastic computations: a spectral method approach*. Princeton University Press.
- ResiWater ANR and BMBF project, Innovative Secure Sensor Networks and Model-Based Assessment Tools for Increased Resilience of Water Infrastructures, available at: <http://www.resiwater.eu/>, last access: 9 March 2018.
- Rossman, L. A., and Boulos, P. F. (1996). *Numerical Methods for Modeling Water Quality in Distribution Systems: A Comparison*. J. Water Resour. Plan. Manage., 122(2), 137-146.
- Fabrie, P., Gancel, G., Mortazavi, I., and Piller, O. (2010). *Quality Modeling of Water Distribution Systems using Sensitivity Equations*. Journal of Hydraulic Engineering-ASCE, 136(1), 34-44.
- Piller, O., Gilbert, D., Haddane, K., and Sabatie, S. (2011). *Porteau: An Object-Oriented programming hydraulic toolkit for water distribution system analysis*. Eleventh International Conference on Computing and Control for the Water Industry (CCWI 2011), 5-7 September 2011, Exeter, UK, 1, 27 pp.