

Spectral Propagation of Parameter Uncertainties in Water Distribution Networks

Mathias Braun¹, Olivier Piller¹, Jochen Deuerlein², Iraj Mortazavi³

¹ Irstea, UR ETBX, Dept. of Water, 33610 Cestas, France

² 3S Consult GmbH, Albtalstrasse 13, 76137 Karlsruhe, Germany

³ CNAM, EA-7340-M2N Modélisation Mathématique et Numérique, 75003 Paris, France

¹mathias.braun@irstea.fr

ABSTRACT

The hydraulic state of a water distribution network is governed by a large number of uncertain parameters. These parameters may be given by uncertain consumer demand, valves states, the value of pipe diameters or the roughness of the pipes. In practice, the influence of parameter variations is important in the decision-making process of water utilities, which emphasizes the need for proper quantification of the resulting uncertainties in head and flow. The central step in uncertainty quantification is the propagation of uncertainties through the system. In the past, the influence of parameter uncertainties on the system state has been studied using perturbation methods, stochastic collocation and interval state estimation. This paper presents the results of an alternative spectral approach that has been examined as part of the French-German research project ResiWater. The generalized Polynomial Chaos Expansion is applied to a small looped water distribution network with multiple uncertain input parameters using a non-intrusive projection method. These results are compared to the Monte Carlo simulation as representative of stochastic collocation methods. It is demonstrated that the Polynomial Chaos Expansion is capable to capture a high order of non-linear effects like the Monte Carlo simulation for a considerably lower computational effort.

Keywords: Water Distribution Networks; Polynomial Chaos Expansion; Uncertainty Propagation

1 Introduction

The modelling and state calculation of water distribution networks is an important task for water utilities and has been used for a number of practical applications in the past. A small selection is given by, but not limited to: Network design [5,8], modelling water leakage [11], optimal sensor placement [3,4] and water quality simulation [7]. In all of these applications, parameters of the network model play an important role in determining accurate and reliable results. But, due to limited accessibility and sparse availability of sensors inside the network they cannot always be measured directly and many of the parameters may vary over time. One may assume that, even though considerable effort is put into the calibration of network models [6,9], the estimated parameters are still uncertain. The objective of uncertainty quantification (UQ) is the evaluation of the influence of parameter uncertainties on the Quantities of Interest (QoI) in a mathematical model and it contains three major steps:

- The survey and classification of parameter/measurement uncertainties in the system (sources, range, importance).
- Development of a concept for the estimation of parameter uncertainty and their impact on the reliability of simulation results.
- Estimating of confidence intervals for hydraulic and water quality model results.

Central part of the process and focus of the research done as part of the collaborative French-German project ResiWater [1] is the second step which deals with the propagation of errors and

uncertainties by means of the mathematical model. In the literature a wide spectrum of methods is proposed with their individual benefits and shortcomings. Three of the most common general approaches are given by perturbation methods, sampling methods and a relatively new approach with the stochastic spectral methods.

- **Perturbation Methods:** These methods calculate the mean, variance and other stochastic 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. This limits their accuracy for highly non-linear models.
- **Sampling Methods:** With Monte Carlo Simulations as one of the most prominent representatives for this group, sampling methods are often applied for the propagation of uncertainties in non-linear models. Although, in general, implementation of the method is a straightforward task, its rate of convergence is defined as $1/\sqrt{M}$, where M is the number of simulations [2,13].
- **Stochastic Spectral Methods:** The objective of spectral approaches like stochastic Galerkin and stochastic collocation methods is the calculation of a spectral representation of the random quantity of interest. Utilizing the smoothness requirement of the polynomial basis leads to an efficient convergence behaviour [10,13].

The interval state estimation is an alternative approach that is popular in control applications. It concentrates on the propagation of the confidence intervals which makes it computationally efficient enough for real-time application. However, unlike the previously mentioned methods it does not provide the same amount of detailed information on the resulting random variable. This article explores the application of spectral methods for uncertainty quantification in water distribution networks. The Polynomial Chaos expansion is applied to model uncertain demands at consumption nodes in the hydraulic network equations. To evaluate the results classical methods are compared to the spectral approach. Section 2 gives a short introduction to the deterministic hydraulic model, followed by some details on random variables that will be used for the definition of the stochastic hydraulic model. Then, the Polynomial Chaos Expansion is introduced with some detail on the derivation of the stochastic equations. In section 3 the small looped demonstration network will be introduced with details on the uncertainty in this scenario and the evaluation. The results will be presented followed by a discussion on the methods in section 4.

2 METHODS

2.1 Hydraulic Model

In hydraulic modelling the simplified topological structure of a water distribution network is described by a directed graph as the one shown in Figure 1. In this graph links represent pipe sections and nodes the pipe junctions and connections. The mathematical description of this graph is given by the incidence matrix $\mathbf{A} \in \mathcal{M}^{n_j \times n_p}(\mathbb{R})$, where n_j is the number of nodes and n_p is the number of links. It is defined as $\mathbf{A} = \{A_{i,j}\}_{1 \leq i \leq n_j; 1 \leq j \leq n_p}$. The coefficients are defined as follows:

$$A_{i,j} = \begin{cases} -1 & , \text{if pipe } j \text{ enters node } i \\ 0 & , \text{if pipe } j \text{ is not connected to node } i \\ +1 & , \text{if pipe } j \text{ exits node } i. \end{cases}$$

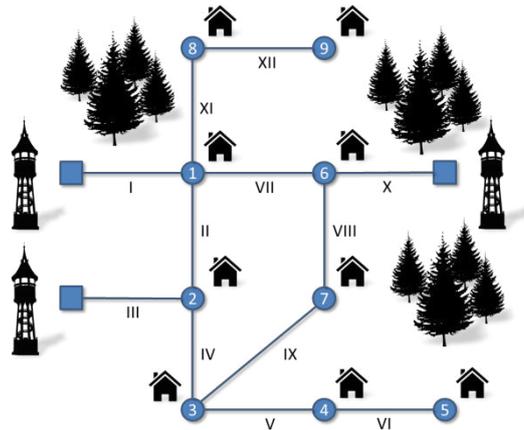


Figure 1. Network graph of a small looped water distribution system.

Water distribution networks in general have a looped structure and the system state is described by the potential at the nodes (head) and the current on the links (flow). The system equations are given by two sets of equations. First the mass balance at the nodes:

$$Aq + d = 0$$

where A is linked to the part of the network that only contains junctions with known demands, $q \in \mathbb{R}^{np}$ is the vector containing the flow rates and $d \in \mathbb{R}^{nj}$ is the vector of demands at consumption nodes. Second the energy equation:

$$\Delta h(r, q) - A^T h - A_f^T h_f = 0.$$

where A_f describes the nodes with fixed potential like reservoirs or tanks and $h \in \mathbb{R}^{nj}$ is the vector containing the piezometric heads. Parameters are given by the potential vector $h_f \in \mathbb{R}^{nj}$ describing fixed head at special nodes like reservoirs or tanks and the vector $r \in \mathbb{R}^{np}$ containing the friction coefficients for each link. The function $\Delta h(r, q)$ describes the loss in head along a pipe and is defined by:

$$\begin{aligned} \Delta h: \mathbb{R}^{np} \times \mathbb{R}^{np} &\rightarrow \mathbb{R}^{np} \\ (r, q) &\mapsto \Delta h(r, q). \end{aligned}$$

It is usually termed the head-loss function. For medium and large Reynolds numbers the head-loss in general is a non-linear function of a friction coefficient r and flow q . In the following application the state vector u consists of the flow rates q and the head h and the system parameters are combined in the vector α .

2.2 Uncertain Parameters and Variables

Above it has been assumed that the values of model parameters are known with their exact value. In real applications it is more likely that they are known by their expected value and a joint probability distribution. As a result it is prudent to assume that the results are also no longer deterministic. To accurately model this in uncertainty quantification, parameters and results are defined by random variables.

A random variable $X = X(\omega)$ assigns a number to each outcome ω of a random experiment. The sample space Ω of such an experiment is defined as the set of all possible outcomes $\Omega = \{\omega\}$. In many technical applications like the one presented in this paper the sample space is directly defined as a random variable. The σ -field or σ -algebra \mathcal{F} is a subset of the sample space that contains all relevant events. In this context an event may be defined as a set of outcomes, including the empty set \emptyset and all combinations of other events in the σ -field. Probability is a concept to measure the

likelihood of occurrence for a certain event $P: \mathcal{F} \rightarrow [0,1]$. It has to satisfy the definitions $P(\emptyset) = 0$, $P(\Omega) = 1$ and if $A_i \in \mathcal{F}$ and $A_i \cap A_j = \emptyset$, $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$. With these components it is possible to define a probability space by the triple (Ω, \mathcal{F}, P) .

A common tool in working with random variables is the cumulative distribution function (cdf). It is defined for $F_X \rightarrow [0,1]$ by

$$F_X(x) = P\{\omega \in \Omega | X(\omega) \leq x\}$$

and describes the probability that a realization of the random variable has a value lower than x .

2.3 Polynomial Chaos Expansion

The polynomial chaos expansion constructs a spectral representation of the uncertain variables of the mathematical model using orthogonal basis polynomial based on a basis random variable Z . Each variable may be represented by the infinite expansion

$$\mathbf{u}(\mathbf{Z}) = \sum_0^{\infty} u_k \Phi_k(\mathbf{Z}), \quad (5)$$

where Φ_k are the orthogonal polynomials. In practical applications a truncated expansion with a total number of $N + 1$ elements is used to approximate random variables

$$\mathbf{u}_N(\mathbf{Z}) = \sum_0^N u_k \Phi_k(\mathbf{Z}). \quad (6)$$

The number of terms in this finite polynomial chaos expansion is dependent on the dimension of the parameter space n and the order of the expansion p . It is calculated as:

$$N + 1 = \frac{(n + p)!}{n! p!}. \quad (7)$$

Polynomial chaos expansion as introduced by Wiener [12] uses orthogonal Hermite polynomials to model arbitrarily distributed random variables based on a Gaussian distribution. Xiu and Karniadakis [14] generalized the method to the use of a wider variety of random processes by introducing broader classes of polynomials defined by the Askey scheme. For each choice of polynomials $\Phi_k(Z)$ a weighted inner product $L^2(\Omega)$ is defined as

$$\langle \Phi_i, \Phi_j \rangle = \int \Phi_i(Z) \Phi_j(Z) dP(Z) = \int \Phi_i(Z) \Phi_j(Z) dP(Z) w(Z) dZ = \delta_{ij} \langle \Phi_i^2 \rangle \quad (8)$$

with respect to the probability density function $w(Z)$.

The first step in any application of the polynomial chaos expansion is the projection of the known random input parameters α on the chosen polynomial basis. The coefficients α_k of their truncated PC expansion

$$\alpha_N(\mathbf{Z}) = \sum_{k=0}^N \alpha_k \Phi_k(\mathbf{Z}) \quad (9)$$

are calculated using the definition in Equation 8

$$\alpha_k = \frac{\langle \alpha(X), \Phi_k \rangle}{\langle \Phi_k^2 \rangle}. \quad (10)$$

The existence and convergence of this projection follows from the classical approximation theory. A general system of coupled non-linear equations like Equations 1 and 2 can be represented in operator form as $\mathcal{P}[\mathbf{u}, \alpha] - \mathcal{Q}[\alpha] = \mathbf{0}$. Here, the dependent random state variables are given by \mathbf{u} and the random input parameters by α . Substituting the random components by their PC expansion results in the stochastic system $\mathcal{P}[\mathbf{u}_N, \alpha_N] - \mathcal{Q}[\alpha_N] = \mathbf{0}$. Projecting these stochastic system equations on the basis polynomials Φ_k for $k = 0 \dots P$ gives a new set of equations that allows for the direct calculation of the stochastic expansion coefficients \mathbf{u}_k for the random representation of the state variable. This Galerkin approach is also known as intrusive spectral projection. This name

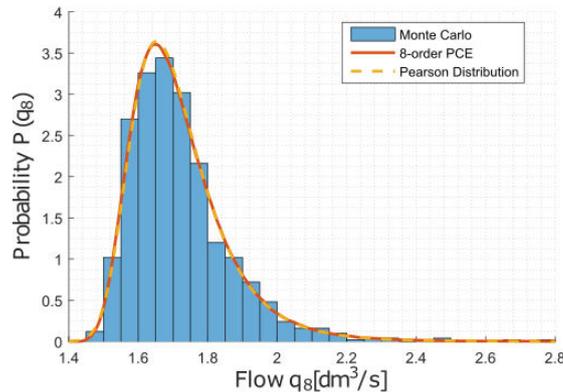


Figure 2. Network graph of a small looped water distribution system.

has been given since the equations of the original problem have to be reformulated. The system of equations depends on a number of factors like the order of the expansion and the dimension of the parameter space and has to be restated if they change. This makes the application in practical applications complicated.

This is one of the main benefits for non-intrusive spectral projection (NISP) approaches. Since the inner product $\langle u(X), \Phi_k \rangle$ is evaluated by sampling the solution space for different realizations \mathbf{Z} these methods are more adaptable to changes. In contrast to the intrusive approach the coefficients are determined by repeated solutions of the original problem.

3 Uncertainty Propagation in a Small Network

This is the For the following demonstration the small looped network shown in Figure 1 will be used. A one dimensional parameter space for the uncertain input is chosen to be represented by the demand d_5 at node 5. It is modelled by a Gaussian distribution with a mean of the deterministic demand and a variance of about 20 percent of the mean to make sure that the non-linear effects are strong enough to influence the result. This parameter may be developed exactly by a first order PCE based on the Hermit polynomials. The main method for the propagation of this random parameter is the PCE. To validate the results a Monte Carlo simulation of the same example will be performed with a sample size of $N_{MC} = 1e5$ simulations. For the calculation of the PCE coefficients a total number of $N_{PCE} = 1e3$ simulations have been performed. The evaluation of the PCE can be done following two different methods. The expansion of the random variable allows for the direct calculation of the first four stochastic moments of its probability distribution through numerical quadrature. Using these moments it is possible to reconstruct the probability density function by the means of the Pearson Distribution. A more general approach is the sampling of the PCE through the basic random variable. This is in general computationally more efficient than the MC simulation since each sample evaluates the polynomial instead of solving the complete system. For the evaluations shown in this section the PCEs are evaluated $N_{PDF} = 1e6$ times. In sum the calculation of the PCE coefficients and the evaluation of the polynomials is considerably faster than the Monte Carlo approach.

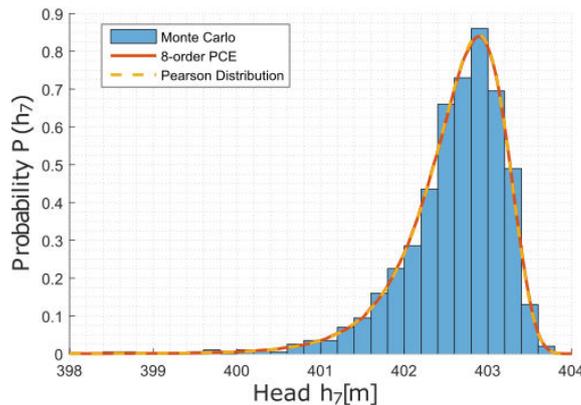


Figure 3. Probability density function for the pressure at node 7 based on the uncertain demand.

Both the Monte Carlo simulation and the PCE produce results for all QoIs in the system. For the discussion of the results a closer look is taken at the flow q_8 at pipe 8 and the head h_7 at node 7 in Figures 2 and 3. Although these two values have been chosen, they are representative for the behaviour of the other variables. In both figures the result of the Monte Carlo simulation is given by the histogram. The red curve shows the kernel density approximation based on the results from sampling the PCE of the random flow. The yellow curve displays the Pearson approximation based on the first four moments.

Figure 2 shows the approximate probability distribution for the flow rate through link 7 $P(q_8)$ for the flow value q_8 in dm^3/s , whereas Figure 3 displays the approximate probability distribution for the head at node 7 $P(h_7)$ for the head h_7 in m . It can be seen that the Monte Carlo simulation and the PCE are in good agreement, but also that the spectral approach gives a smoother solution even so it uses fewer sampling points. Both methods are capable to capture the non-linear influences from the system on the results which distort the distribution into a skewed form. From the view of an system operator it is interesting to see that the head distribution is skewed in the direction of lower heads. This means that due to the uncertainty in the demands it is far more likely to experience a head deficiency than too high heads.

4 Discussion

Regarding the results presented in the previous section two important questions have to be answered. First, a closer look will be taken on the computational complexity of the Monte Carlo

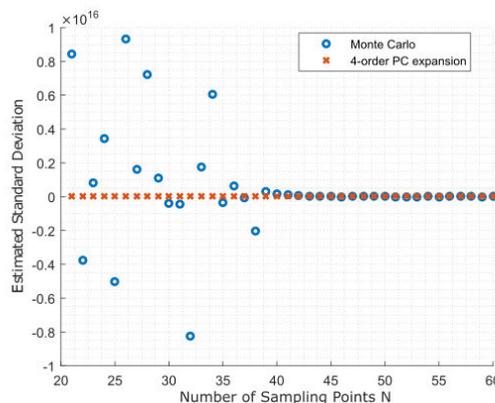


Figure 4. Estimated standard deviation of the pressure at node 7, based on N Samples for classical Monte Carlo and a 4th order Polynomial Chaos expansion.

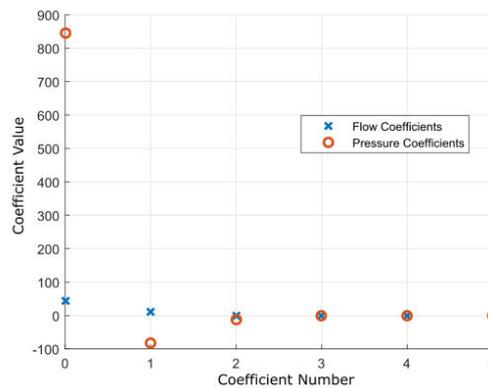


Figure 5. Coefficients for a 5th-order PC expansion for flow and pressure.

simulation compared to the Polynomial Chaos expansion. Second, the evaluation of the necessary order for the Polynomial Chaos expansion will be discussed. Looking at the fact that both the Monte Carlo simulation and the non-intrusive spectral projection may be classified as sampling methods one may ask why the application of the Polynomial Chaos expansion is beneficial. From literature the answer to this question lays in the fact that PC methods use the smoothness of the orthogonal basis polynomials and in effect have a superior convergence behaviour [10]. To illustrate this, Figure 4 shows the value for the standard deviation of the pressure at node 7 over the number of sampling points evaluated. For a low dimensional problem as the one discussed in this article a very small number of points is sufficient to get a good estimation of the PCE coefficients.

An important task for any application of a expansion approach is the evaluation of accuracy for the chosen development order. Since it is not possible to do so a priori this section shows the measures that have been taken based on the estimated coefficients. In a first iteration the expansion order is chosen due to experience. Based on the evaluation it has to be adapted. The appropriate expansion order depends on factors like the non-linear properties of the modelled system and the desired accuracy for the application. Similar to other examples from polynomial approximation theory it is assumed that the expansion converges to the true solution and that the theoretical infinite series may be represented by a truncated series of order N . From this it follows that coefficient values of higher order polynomials should be small and go to zero. Figure 4 illustrates the convergence behaviour of the coefficients for a fourth order PC expansion. The coefficients are shown for the flow through pipe 7 and the pressure at node 7. As expected, their values decline rapidly and are close to zero for higher orders.

5 Conclusion

This paper showed the application of the polynomial chaos expansion on a small water distribution network. The results have been compared to state of the art methods given by the Monte Carlo simulation. It can be concluded that the PCE is capable to capture the non-linear effect of the equations making it superior to sensitivity based perturbation methods. At the same time it is computationally more efficient than the straight forward MC simulation.

For a more efficient evaluation of the PCE in future applications a closer look will be taken at more efficient sampling strategies for the numerical projection and adaptive sparse representations that drastically reduce the number of coefficients in the expansion. As an alternative path more efficient solution methods will be investigated. This includes fast algorithms for the resolution of the hydraulic equations as well as the application of methods used in model order reduction.

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

- [1] ResiWater. <http://www.resiwater.eu/>. Accessed: 2016-10-07.
- [2] George Fishman. Monte Carlo: concepts, algorithms, and applications. Springer Science & Business Media, 2013.
- [3] William E Hart and Regan Murray. Review of sensor placement strategies for contamination warning systems in drinking water distribution systems. *Journal of Water Resources Planning and Management*, 136(6):611–619, 2010.
- [4] Avi Ostfeld and Elad Salomons. Optimal layout of early warning detection stations for water distribution systems security. *Journal of Water Resources Planning and Management*, 130(5):377–385, 2004.
- [5] Avi Ostfeld and Uri Shamir. Design of optimal reliable multiquality water-supply systems. *Journal of Water Resources Planning and Management*, 122(5):322–333, 1996.
- [6] Olivier Piller, Denis Gilbert, and Jakobus E Van Zyl. Dual calibration for coupled flow and transport models of water distribution systems. *Water Distribution Systems Analysis*, pages 722–731, 2010.
- [7] Lewis A Rossman, Paul F Boulos, and Tom Altman. Discrete volume-element method for network water-quality models. *Journal of Water Resources Planning and Management*, 119(5):505–517, 1993.
- [8] Dragan A Savic and Godfrey A Walters. Genetic algorithms for least-cost design of water distribution networks. *Journal of water resources planning and management*, 123(2):67–77, 1997.
- [9] Dragan A Savic, Zoran S Kapelan, and Philip MR Jonkergouw. Quo vadis water distribution model calibration? *Urban Water Journal*, 6(1):3–22, 2009.
- [10] Ralph C Smith. Uncertainty quantification: theory, implementation, and applications, volume 12. SIAM, 2013.
- [11] JE Van Zyl and AM Cassa. Modeling elastically deforming leaks in water distribution pipes. *Journal of Hydraulic Engineering*, 140(2):182–189, 2013.
- [12] Norbert Wiener. The homogeneous chaos. *American Journal of Mathematics*, 60(4):897–936, 1938.
- [13] Dongbin Xiu. Numerical methods for stochastic computations: a spectral method approach. Princeton University Press, 2010.
- [14] Dongbin Xiu and George Em Karniadakis. The wiener–askey polynomial chaos for stochastic differential equations. *SIAM journal on scientific computing*, 24(2):619–644, 2002.