1. Introduction

Transient wave, also known as water hammer wave, in a water supply pipe system is a pressure surge caused by a sudden fluid perturbation, e.g., a valve operation. This pressure wave can cause major problems, for example, pipe breakage and structural fatigue (Schmitt et al., 2006). An engineer should always assess these risks, where at least a rough forecast of the water hammer pressure rise is needed and can be computed via the Joukowsky equation (Keramat & Haghighi, 2014). Transient waves can also be employed to detect defects and anomalies in pipes, such as leakage, blockage, and deterioration (Colombo et al., 2009; Duan et al., 2013; Ferrante et al., 2007; Gong et al., 2013; Huang et al., 2020; Lee et al., 2005; Meniconi et al., 2015; Nguyen et al., 2018; Wang & Ghidaoui, 2018a, 2018b, 2019; Wang et al., 2021; Wang, 2021; Wang, Ghidaoui, & Lin, 2019). This is realized by matching the transient measurement with its physical model where defects are assumed to be free parameters to be optimized. For this purpose, a transient wave forecast with high fidelity is desired, which can be realized by a numerical simulation using the method of characteristics (MOC; Chaudhry, 2014; Wylie & Streeter, 1978), as long as all the model input information can be precisely given.

However, real water supply pipeline systems are always highly complex and involve a large number of uncertainties, either aleatoric or epistemic (Wang, Waqar, et al., 2020), including traffic noise, mechanical device error, turbulence (Wang, Lin, & Ghidaoui, 2020), air bubbles in water (Zhou et al., 2018), imprecise wave speed (Wang & Ghidaoui, 2018b; D. Wiggert, 1999), imprecisely known pipe bends (D. C. Wiggert et al., 1985), internal partial blockage or roughness (Zhang et al., 2018), pipe material properties such as the Young's modulus and viscoelastic parameters (Pan et al., 2020, 2021; Wang, Lin, et al., 2019; H. C. Yan et al., 2018), and initial and boundary conditions such as the boundary pressure and pre-transient flow velocity (Alawadhi et al., 2018; Alawadhi & Tartakovsky, 2020; Meniconi et al., 2018; Wang, Che, & Ghidaoui, 2020). The influence of environmental uncertainty on transient wave propagation and defect detection has been noticed in (Covas & Ramos, 2010; Liggett & Chen, 1994; Wang & Ghidaoui, 2018b).

Several approaches have been attempted for more systematical uncertainty quantification (UQ) of transient wave propagation. Considering the uncertainties on pre-transitional (steady-state) pressure or flow velocity, the polynomial chaos expansion (PCE; Sattar & El-Beltagy, 2017) and the method of distributions (Alawadhi et al., 2018) are respectively employed to model the uncertainties of transient pressure. However, the methodologies in (Alawadhi et al., 2018; Sattar & El-Beltagy, 2017) use, explicitly or implicitly, the property that the transient wave varies linearly with the steady-state pressure and flow velocity, thus cannot...
be easily extended for other nonlinear uncertain parameters. The Monte-Carlo (MC) simulation approach generates random samples of uncertain parameters, computes the corresponding model outputs (numerically, using MOC), and then identifies the statistical characteristics of transient wave (Duan, 2015; Duan, Tung, & Ghidaoui, 2010; Huang et al., 2017; Zhang et al., 2011). However, this method suffers from a slow convergence rate (Kazemi & Collins, 2018); it usually requires thousands to millions of runs, especially for complex UQ problems (e.g., in a pipe environment) where many system parameters are uncertain. As a matter of fact, the computational cost of the MC simulation is often not affordable for high-dimensional problems even with modern high-performance computing architectures. Furthermore, the MC method has to redefine probabilistic distributions of uncertain parameters; this information is not always available, particularly for the cases that some uncertain parameters may have strong interdependence. The first-order-second-moment (FOSM) method estimates the variation of model output due to uncertain parameters via the first-order Taylor expansion and quantifies the uncertainty and sensitivity by the variance (the second-order central moment) of model output (Duan, 2015; Ferrante et al., 2002; Rougier & Goldstein, 2001; Wang, Quost, et al., 2016). Since the linear approximation is only valid for a small change of input parameter, the FOSM method is only applicable for parameters with a low level of uncertainty and for local sensitivity (Duan, 2016, 2018), which is not the case of pipeline systems where each parameter may vary in a large interval (e.g., the wave speed in a pipe typically varies by 10%–15% (Wang & Ghidaoui, 2018b; Wylie & Streeter, 1978)). In summary, for the problem of transient wave propagation in water supply pipe systems, a more advanced UQ methodology is desired which should have the following attributes: (i) It can quantify a large number of uncertain parameters, each of which may vary in a large range (consequently, the global sensitivity analysis (GSA) that measures sensitivity across the whole input space can be realized); (ii) it is not limited to a specific predefined probabilistic distribution of uncertain parameters; (iii) it has a low computational cost, in other words, the required number of numerical simulations of transient wave propagation is minimized.

In order to solve the aforementioned problems, surrogate models, also known as meta-models, can be employed, which replace the original computational model with an easy-to-evaluate model (Forrester et al., 2008; Razavi et al., 2012). The surrogate model techniques can realize fast computations of risk/reliability assessment, optimization, uncertainty quantification, and sensitivity analysis, which always require a large number of model evaluations (Chen et al., 2017; Margheri & Sagaut, 2016; Queipo et al., 2005). Using a sampling strategy to select discrete representatives of input uncertain parameters, such as the Sobol sequences which is an example of the quasi-Monte-Carlo (QMC) samples (Sobol’, 1967), and computing the corresponding transient wave responses, an analytical surrogate model (a continuous function of uncertain parameters) can be constructed. PCE and Kriging are the two most popular surrogate modeling approaches. The PCE method approximates the original model by a series of orthogonal multi-variate polynomials (Xiu & Karniadakis, 2003), which is good at capturing the global behavior of the model. The Kriging approach assumes that the model is a summation of a trend part and a realization of a Gaussian random field (Krige, 1951), the latter can effectively describe the local variability of the model. Ordinary Kriging is the most widely used Kriging method, which assumes the trend term as an unknown constant. A recent version of Kriging, known as the polynomial-chaos-Kriging (PC-Kriging; Kersaudy et al., 2015; Schobi et al., 2015), uses the PCE result as the trend term of Kriging and, thus, inherits the attributes of both the PCE and Kriging methods, although more unknown parameters increase the risk of over-fitting. When a reliable surrogate model is built up, uncertainty propagation with any assumption of uncertain parameter distribution and GSA can be rapidly computed; the variance-based sensitivity analysis (often referred to as the Sobol method or Sobol index) is a powerful tool to analyze the global sensitivity of each uncertain parameter (Abbiati et al., 2021; Saltelli et al., 2008, 2010; Sobol, 1993, 2001).

The next section shows the uncertainties of pipe transient wave via experimental data. Then, the full UQ and GSA methodology is introduced in Section 3. In Section 4, illustrative examples are presented to demonstrate the proposed methodology and the influences of uncertainty sources on various transient signal features. Finally, conclusions are drawn in Section 5.
2. Experimental Study of Pipe System Uncertainties

2.1. Experimental Setup and Measured Signals

The uncertainties of system parameters are illustrated via experimental data from a pipe system in the Water Resources Research Laboratory at the Hong Kong University of Science and Technology. The experimental setup is shown in Figure 1. The pipe material is high-density polyethylene (HDPE). The pipe length is \( l = 142 \) m, the internal pipe diameter is \( d = 0.0792 \) m, and the pipe wall thickness is \( e = 0.0054 \) m. The upstream boundary of the pipe is connected to a vertical multistage centrifugal pump. The steady-state water flow discharge is approximately 0.5 L/s. A ball valve is set at the downstream end of the pipe, where transient waves are generated by rapidly closing the valve. Seven pressure sensors (Druck UNIK 5000, Leicester, UK) with a full-scale range of 0.5 MPa (5 bar) and a maximum error of 0.04% of the full-scale are set in the pipe. A National Instruments cRIO-9030 data acquisition system (Austin, Texas) is used to collect the data, with a sampling frequency of 1,000 Hz. Figure 2 (a) shows the measurements from the seven sensors in a transient test. The experiment is repeated 30 times, among which measured signals of five experiments from the most downstream sensor at 139.62 m are plotted in Figure 2 (b). In this figure, each signal has been shifted and aligned with the wave arrival time. In the following, the variabilities or uncertainties of valve closure time, wave speed, and viscoelastic parameters (VEPs) calibrated from the transient wave measurements are investigated and discussed.
2.2. Valve Closure

Figure 2 (b) shows that the valve closure time, denoted by $t_c$, is different in these tests. This was not intended: The experimental setup was unchanged and the same experimentalist tried to close the valve as fast as possible in every test. However, the valve operation is manual, which leads to the uncertainty of valve closure time. The valve closure time in each test can be estimated from the measured signal, more specifically, by the difference between the wave arrival time (aligned in Figure 2 (b)) and the instant that the signal reaches the first local maximum. The interval of $t_c$ from the 30 transient experiments is $[0.059, 0.091]$ s.
2.3. Wave Speed

In viscoelastic pipes, the wave speed is time-dependent, as the viscoelastic effect (retarded strain) grows with time (Wang, Lin, Ghidaoui, Meniconi, Brunone, 2020). The elastic component of wave speed, which corresponds to the immediate response of pipe to transient wave, has the theoretical expression Equation A4 (cf. Appendix A) and can be computed given the precise information of the system, including the Young’s modulus, Poisson’s ratio, and pipe wall thickness. In practice, however, this would not be precise, because the material parameters and pipe geometry change with the surrounding environment and a real-time mechanical test would not be possible for buried pipes. For example, in most cases especially for aged pipes, only a very rough value of the Young’s modulus is available which implies a quite large error of wave speed. Alternatively, a more reliable estimate of the elastic wave speed may be obtained experimentally using the wave arrival time at two different sensors:

\[ \hat{a} = \frac{x_i - x_j}{t_i - t_j}, \]

in which \(x_i\) and \(x_j\) are the locations of the \(i\)-th and \(j\)-th sensors, \(t_i\) and \(t_j\) are the corresponding times the signals reaching a given pressure head level, as shown by the dashed line in Figure 2 (a). Here, the wave speed is computed from two adjacent sensors using the experimental data of the 30 transient tests. The variability of the wave speed estimate is shown in Figure 3, which demonstrates that the wave speed estimation, although more precise than the theoretical forecast, still has strong uncertainties with an error up to approximately 20%. As a matter of fact, a mechanical test on a pipe sample gives the Young’s modulus \(E = 1 \times 10^9\) Pa and the Poisson’s ratio \(\nu = 0.46\); the density and bulk modulus of water are respectively \(\rho = 1,000\) kg/m\(^3\) and \(\kappa = 2.1 \times 10^9\) Pa. According to Equation A4, the theoretical forecast of the elastic wave speed is 288 m/s, which has a very large error according to the experimental result. Besides, a less uncertain wave speed can be obtained with a longer distance between two sensors. However, in practice, it is not always possible to set two sensors very far away from each other, because the entry points of a buried pipe system might be very limited.

2.4. Viscoelastic Parameters

Precise evaluation of pipe material viscoelasticity is key to transient simulation and defect detection in pipes (Wang, Lin, Keramat et al., 2019). The viscoelasticity depends on the temperature, humidity, stress time history, and pipe constraints, therefore cannot be predicted by mechanical tests on pipe samples (Covas et al., 2004). Alternatively, the VEPs can be estimated by matching the transient measurement with its theoretical model, which incorporates the viscoelasticity via the generalized Kelvin-Voigt (K-V) model (Covas et al., 2005; Ferrante & Capponi, 2018; Keramat & Haghighi, 2014; Pan et al., 2020, 2021; Pezzinga et al., 2016; Soares et al., 2008; Wang, Lin, Ghidaoui, Meniconi, Brunone, 2020; Weinerowska-Bords, 2015; Yao et al., 2016). Here, the VEP estimation is performed by the method in (Wang, Lin, Ghidaoui, Meniconi, Brunone, 2020), which is applicable even when unknown defects exist and uses only the first seven steady-state crossing times (instead of the full signal). Three-element and five-element are respectively assumed in the generalized K-V model. In the former case, the initial parameters in the optimization procedure are \(J = 10^{-10} \times [2, 1, 0.5] \) Pa\(^{-1}\) and \(\tau = [0.05, 0.5, 5]\) s; in the latter case, they are \(J = 10^{-10} \times [2, 1, 0.8, 0.5, 0.1] \) Pa\(^{-1}\) and \(\tau = [0.05, 0.5, 1.5, 5, 10]\) s. The estimated VEPs for the 30 tests are shown in Figure 4, which demonstrates the uncertainty of VEPs. The VEPs are much more scattered for the case where only \(N_{KV} = 3\) elements are assumed in the generalized K-V model. Therefore, \(N_{KV} = 5\) is a more appropriate model where the VEP uncertainty has been largely reduced. The interdependence of \(J\) and \(\tau\) is also studied using the data from the 30 tests, which is introduced and applied in UQ and GSA in Section 4.4.
2.5. Problem Statement

The aforementioned results show the distributions of some system parameters. However, there exist other system parameters that are difficult to be calibrated, e.g., the pipe wall thickness which may change with time due to corrosion, biofilm, or deposition (X. F. Yan et al., 2021). Furthermore, in urban water supply systems, the uncertainties are often more complicated and unpredictable than the laboratory experiments introduced in this section, due to the traffic noise, users, pump and valve activities, etc. In terms of uncertainty management, it is natural to have the following questions:

• Is it possible to build up a model for the transient wave \( h(t; \mathbf{u}) \) with varying uncertain parameters \( \mathbf{u} = (u_1, ..., u_M) \) in their practical ranges?

• How to quantify the influences of various uncertain parameters on the transient wave propagation?

In the next section, a full UQ and GSA methodology is introduced to address these problems.

3. Methodology

The proposed methodology builds up a surrogate model of quantity of interest (QOI) with \( \mathbf{u} \), denoted by \( h(\mathbf{u}) = h(u_1, ..., u_M) \). The surrogate model is analytical for any \( u_m (m = 1, ..., M) \) in a given range: \( u_m \in [u_m^{\text{min}}, u_m^{\text{max}}] \). Then, the UQ and GSA which require massive computations of \( h(\mathbf{u}) \) can be rapidly computed.

3.1. Sampling Design

The surrogate modeling is based on \( N \) samples of \( \mathbf{u} \), denoted by \( \mathbf{u}_n = (u_1^n, ..., u_M^n) \) where \( u_m^n \in [u_m^{\text{min}}, u_m^{\text{max}}] \), \( m = 1, ..., M \), \( n = 1, ..., N \), and the corresponding numerical computation results of QOI \( h(\mathbf{u}_n) \). Note that the convergence and reliability of the surrogate modeling depend on the sample size \( N \) and the choice of \( \mathbf{u}_n \) in the parameter space. In order to decrease the computational cost of numerical simulation, the required number \( N \) should be minimized and the distribution of \( \mathbf{u}_n \) should be optimized.

Assuming that \( \mathbf{u} = (u_1, ..., u_M) \) are independent and follow a distribution with support set \( [u_m^{\text{min}}, u_m^{\text{max}}] \), the MC simulation method can be employed to take \( N \) samples of \( \mathbf{u} \). However, this method is inefficient in terms of convergence rate (Khazaie et al., 2019). On the other hand, QMC is able to solve the problem using quasi-random low-discrepancy sequences to more efficiently represent the space of \( \mathbf{u} \) (Morokoff & Caflisch, 1995). Sobol sequences are an example of QMC (Sobol’, 1967), which form successively finer uniform partitions of the unit interval and then reorder the coordinates in each dimension. Figure 5 shows the first to thirtieth samples of Sobol sequences in the two-dimensional unit square and in the three-dimensional unit cube. Let \( (\mathbf{u}_1, ..., \mathbf{u}_N) \) denote the \( n \)-th sample of Sobol sequence in the \( M \)-th dimensional unit hypercube, the associated samples of uncertain parameters are obtained by (\( u_{1m}, ..., u_{Nm} \)), where

\[
u_m = u_m^{\text{min}} + \frac{u_m^{\text{max}} - u_m^{\text{min}}}{N} n = 1, ..., N.
\]
3.2. Surrogate Modeling

With the discrete inputs \( \mathbf{u}_n = (u_{1,n}, \ldots, u_{M,n}), n = 1, \ldots, N \), and the corresponding outputs \( h(\mathbf{u}_n) \), a surrogate model can be built, which is an analytical function \( h(\mathbf{u}) \) taking values continuously on \( \mathbf{u} \). Here, three surrogate modeling techniques, PCE, Kriging, and PC-Kriging, are respectively introduced.

3.2.1. Polynomial Chaos Expansion

It is assumed that the output \( h \) is a second-order random variable, i.e., \( \mathbb{E}(h^2) < \infty \), PCE approximates \( h \) by a series of orthogonal polynomials of the uncertain (input) parameters \( \mathbf{u} \):

\[
 h \approx \mathcal{M}^{PCE}(\mathbf{u}) = \sum_{\alpha \in \mathcal{A}} k_\alpha f_\alpha(\mathbf{u}) = \mathbf{k}^\top \mathbf{f}(\mathbf{u}).
\]  

Here, \( \mathbf{f}(\mathbf{u}) = \text{vec}\{f_\alpha(\mathbf{u}) : \alpha \in \mathcal{A}\} \) are the orthogonal polynomials with indices \( \alpha = (\alpha_1, \ldots, \alpha_M) \in \mathcal{A} \) and coefficients \( \mathbf{k} = \text{vec}\{k_\alpha \in \mathcal{A}\} \). \( \mathcal{A} \subset \mathbb{N}_M^M \) is the index set, which is selected to be (Blatman & Sudret, 2010; Schobi et al., 2015):

\[
 \mathcal{A} = \mathcal{A}_M^{M,p} = \{\alpha \in \mathbb{N}_M^M : \|\alpha\|_q \leq p\},
\]  

where \( \|\cdot\|_q \) is the \( q \)-norm. Note that when \( q = 1 \), a classical total degree index set is selected which consists of polynomials with total degree not greater than \( p \), i.e., \( \mathcal{A} = \mathcal{A}_M^{M,p} = \{\alpha \in \mathbb{N}_M^M : \sum_{m=1}^M \alpha_m \leq p\} \), that includes all the polynomials \( \alpha \in \mathbb{N}_M^M \) up to order \( p \). For \( q < 1 \), a more restrictive truncation scheme is applied, where only polynomials with low-degree interactions are considered. This reduces the number of polynomials and simplifies the computational complexity in PCE-based surrogate modeling (Blatman & Sudret, 2010; Schobi et al., 2015). The family of the orthogonal polynomials is chosen based on the distributions of the input random variables, among which the Hermite polynomials are used for Gaussian random variables.

**Example 1** The Hermite polynomials for two uncertain parameters (i.e., \( M = 2 \)) and \( p = 3 \) are considered. When \( q = 1 \), \( \mathcal{A} \) includes 10 polynomials: \( f_{0,0} = 1, f_{1,0} = u_1, f_{0,1} = u_2, f_{2,0} = u_1^2, f_{0,2} = u_2^2, f_{1,1} = u_1 u_2, f_{2,1} = u_1^3 - 3u_1, f_{2,0} = u_2^3 - 3u_2, f_{2,1} = u_1(2u_2^2 - 1), f_{1,2} = u_2(2u_1^2 - 1) \). They are orthogonal when Gaussian probability distribution of \( u_1 \) and \( u_2 \) are assumed (Xiu & Karniadakis, 2002). When \( q = 0.75, f_{2,1} \) and \( f_{1,2} \) are removed where \( \|\alpha\|_q = 3.7 > p = 3 \). When \( q = 0.5, f_{1,1} \) is further removed where \( \|\alpha\|_q = 4 > p = 3 \); in this case, all the polynomials with interaction of \( u_1 \) and \( u_2 \) are not considered.

The \( N \) samples \( \mathbf{u}_n = (u_{1,n}, \ldots, u_{M,n}) \) of \( \mathbf{u} \) and the corresponding output \( h_n, n = 1, \ldots, N \), are assumed to follow the PCE model Equation 3; let \( \mathbf{h} = (h_1, \ldots, h_N) \) and \( \mathbf{F} = (\mathbf{f}(\mathbf{u}_1), \ldots, \mathbf{f}(\mathbf{u}_N)) \), we have \( \mathbf{h} = \mathbf{Fk} \). The least squares method gives the estimation of the PCE coefficients:

\[
 \hat{\mathbf{k}} = \arg \min_{\mathbf{k}} \|\mathbf{h} - \mathbf{Fk}\|^2 = (\mathbf{F}^\top \mathbf{F})^{-1} \mathbf{F}^\top \mathbf{h}.
\]  

After checking the reliability of the PCE model (detailed later in Section 3.2.4), the model output with any new input \( \hat{\mathbf{u}} \) can be forecast by
3.2.2. Kriging

The Kriging method (Kleijnen, 2009; Krige, 1951), also known as the Gaussian process regression or the spatial correlation modeling, assumes that the QOI \( h(u) \) is a realization of a Gaussian random field:

\[
\hat{h}(\mathbf{u}) = \mathcal{M}^{(PC)}(\mathbf{u}) = \sum_{a \in A} \hat{k}_a f_a(\mathbf{u}) = \mathbf{k}^\top \mathbf{f}(\mathbf{u}).
\]

(6)

The Kriging method (Kleijnen, 2009; Krige, 1951), also known as the Gaussian process regression or the spatial correlation modeling, assumes that the QOI \( h(u) \) is a realization of a Gaussian random field:

\[
h \approx \mathcal{M}^{(K)}(\mathbf{u}) = \mathbf{k}^\top \mathbf{f}(\mathbf{u}) + \sigma^2 Z(\mathbf{u}),
\]

(7)
in which \( \mathbf{k}^\top \mathbf{f}(\mathbf{u}) = \sum_{w=1}^{W} k_w f_w(\mathbf{u}) \) is the trend part, \( \sigma^2 \) is the variance of the Gaussian field, and \( Z(\mathbf{u}) \) is a stationary Gaussian random field with zero-mean and unit-variance. The random part \( Z(\mathbf{u}) \) is decided by its auto-correlation function (ACF), denoted by \( R(\mathbf{u}, \mathbf{u}') = R(|\mathbf{u} - \mathbf{u}'|, \varnothing) \), where \( \varnothing \) denotes its hyper-parameters. Various ACFs can be used (Khazaie et al., 2016; Klimeš, 2002); in this paper, the Matérn ACF (Matérn, 2013)

\[
R(|\mathbf{u} - \mathbf{u}'|; 1, v) = \prod_{m=1}^{M} \left( 1 + \frac{\sqrt{3} |\mathbf{u}_m - \mathbf{u}'_m|}{l_m} \right)^v \chi_v \left( \frac{\sqrt{3} |\mathbf{u}_m - \mathbf{u}'_m|}{l_m} \right)
\]

(8)
is considered, in which \( \Gamma \) is the Euler Gamma function, \( \chi_v \) is the Bessel function of third kind, \( v \) is the shape parameter (can be given a priori), and \( l = \{l_1, \ldots, l_M\} \) are the scale parameters to be estimated. If \( v = p + 1/2, p \in \mathbb{N} \), the Matérn ACF in Equation 8 can be simplified as a product of an exponential and a polynomial of order \( p \), among which \( v = 1.5 (p = 1) \) is taken in this paper (Roustant et al., 2012; Schobi et al., 2015), where the Matérn ACF becomes

\[
R(|\mathbf{u} - \mathbf{u}'|; 1, 1.5) = \prod_{m=1}^{M} \left( 1 + \frac{\sqrt{3} |\mathbf{u}_m - \mathbf{u}'_m|}{l_m} \right)^{1.5} \exp \left( -\frac{\sqrt{3} |\mathbf{u}_m - \mathbf{u}'_m|}{l_m} \right).
\]

(9)

With different assumptions of the trend term, three types of Kriging are defined:

- **Simple Kriging:** The trend term is a known constant (independent of \( \mathbf{u} \)). More exactly, \( W = 1 \) and \( f_1(\mathbf{u}) = 1 \), thus \( \mathbf{k}^\top \mathbf{f}(\mathbf{u}) = k \).

- **Ordinary Kriging:** the trend term is an unknown constant. More exactly, \( W = 1 \) and \( f_1(\mathbf{u}) = 1 \), thus \( \mathbf{k}^\top \mathbf{f}(\mathbf{u}) = k \), where \( k \) is the unknown parameter to be estimated.

- **Universal Kriging:** the trend term is a linear combination of \( W \) predefined functions \( f_w(\mathbf{u}) \):

\[
\mathbf{k}^\top \mathbf{f}(\mathbf{u}) = \sum_{w=1}^{W} k_w f_w(\mathbf{u})
\]

(10)

where \( k_w, w = 1, \ldots, W, \) are parameters to be estimated.

In this paper, the ordinary Kriging and a specific universal Kriging (the PC-Kriging which is introduced in Section 3.2.3) are applied.

Then, the parameters of the Kriging model are estimated. First, the hyper-parameters \( \varnothing \) in the ACF \( R(|\mathbf{u} - \mathbf{u}'|, \varnothing) \) are estimated by the maximum likelihood method (Marrel et al., 2008), which results in

\[
\hat{\varnothing} = \arg \min_{\varnothing} \left\{ (\mathbf{h} - \mathbf{Fk})^\top \mathbf{R}^{-1} (\mathbf{h} - \mathbf{Fk}) \det(\mathbf{R})^{1/2} \right\}.
\]

(11)

Here, \( \mathbf{R}(\varnothing) = (R(|\mathbf{u}_n - \mathbf{u}_m|, \varnothing))_{n=1, m=1}^{N, N} \) is the correlation matrix, \( \mathbf{F} = (f_w(\mathbf{u}_n))^w_{n=1, w=1} \), and

\[
\mathbf{k}(\varnothing) = \left( \mathbf{F}^\top \mathbf{R}^{-1} \mathbf{F} \right)^{-1} \mathbf{F}^\top \mathbf{R}^{-1} \mathbf{h}.
\]

(12)
Note that Equation 11 involves a multi-dimensional optimization problem; in this paper, it is solved by the Broyden-Fletcher-Goldfarb-Shanno (BFGS) Quasi-Newton algorithm (Fletcher, 2013). After Equation 11 is computed, the trend term in the Kriging model Equation 7 is derived by $\hat{h}(u) = \hat{k}^\top f(u) + r(u)\hat{R}^{-1}(h - F\hat{k})$, where $\hat{k}$ is Equation 12 in which the parameter $\theta$ is replaced by the solution of Equation 11.

For any new uncertain parameter $\tilde{u}$, the Kriging predictor is (Santner et al., 2003; Schobi et al., 2015):

$$\hat{h}(\tilde{u}) = \hat{k}^\top f(\tilde{u}) + r(\tilde{u})\hat{R}^{-1}(\tilde{h} - F\hat{k}),$$

in which $\hat{R} = R(\hat{\theta})$ and $r(u) = \text{vec}[R(|u - u_0|, \hat{\theta}), n = 1, ..., N]$ is an $N$-dimensional vector representing the correlation between the new uncertain parameter and those input parameters used in the Kriging modeling.

### 3.2.3. PC-Kriging

PC-Kriging is a type of universal Kriging that uses the PCE model as the trend term in Kriging (Schobi et al., 2015):

$$h \approx \mathcal{M}^{(PC)}(u) = \sum_{a \in A} k_a f_a(u) + \sigma^2 Z(u).$$

It is found in (Schobi et al., 2015) that PC-Kriging inherits the capacity of modeling the global trend of PCE (with the first term of Equation 14) and the advantage of capturing the local variability of Kriging (with the second term of Equation 14). However, the PC-Kriging has to estimate more unknowns and thus increases the risk of over-fitting (Schobi et al., 2015). Basically, the performance of the PC-Kriging method depends on the shape of the output function. Here, the PCE-based trend term and the Gaussian field are determined sequentially (known as the sequential PC-Kriging in (Schobi et al., 2015)). First, the set of polynomials and their coefficients are obtained as in Section 3.2.1; then, the Kriging parameters are estimated as in Section 3.2.2. The prediction of QOI for any new input parameters $\tilde{u}$ is given by Equation 13.

### 3.2.4. Evaluation of Surrogate Modeling Accuracy

The reliability of a built surrogate model can be evaluated by new samples. Let $\tilde{u}_n$, $n = 1, ..., N_{MC}$, be the forecast samples of uncertain parameters and $h(\tilde{u}_n)$ be the associated numerical simulation results. Let $\hat{h}(\tilde{u}_n)$ denote the forecast of QOI from a surrogate model, its relative forecast error can be computed as:

$$\epsilon = \sqrt{\frac{1}{N_{FC}} \sum_{n=1}^{N_{FC}} \left( \frac{h(\tilde{u}_n) - \hat{h}(\tilde{u}_n)}{h(\tilde{u}_n)} \right)^2}.$$

A surrogate model is deemed as reliable if Equation 15 is less than a given threshold.

### 3.3. Sobol Index for Global Sensitivity Analysis

In this section, the global sensitivity is quantified via the variance-based Sobol sensitivity indices (Sobol, 1993) through MC or QMC (Sobol sequences) estimators (Saltelli et al., 2010). Given $N_{MC}$ MC or QMC samples of $u$, let $\tilde{h}_n$ ($n = 1, ..., N_{MC}$) represent the associated QOI computed from a surrogate model (PCE, ordinary Kriging or PC-Kriging), then the mean estimate $\mu(h)$ of the QOI $h$ is

$$\mu(h) \approx \frac{1}{N_{MC}} \sum_{n=1}^{N_{MC}} \tilde{h}_n.$$

Similarly, the total variance $\sigma^2(h)$ can be estimated:

$$\sigma^2(h) \approx \frac{1}{N_{MC}} \sum_{n=1}^{N_{MC}} \left( \tilde{h}_n - \mu(h) \right)^2.$$
Let $\Psi_1$ and $\Psi_2$ be two independent sample sets of $u$, each of which forms an $N_{MC} \times M$ matrix. The first-order Sobol index for the $m$-th uncertain parameter ($m = 1, \ldots, M$) is

$$S_m = \frac{\sigma_m^2}{\sigma^2},$$

obtained from the estimate of the total variance $\sigma^2$ and the partial variance $\sigma_m^2$:

$$\sigma_m^2 = \text{Var}_{u_m} \left( E_{u_m} (h | u_m) \right) \approx \frac{1}{N_{MC}} \sum_{n=1}^{N_{MC}} \bar{h}_n (\Psi_2 (\Psi_1) - \bar{h}_n (\Psi_1)),$$

where $u_m$ denotes the set of all variables except $u_m$, $\Psi_1^m$ is the first sample set $\Psi_1$ where the $m$-th column has been replaced by the $m$-th column of the second sample set $\Psi_2$. The residual variance can be estimated by the difference between the total variance and the sum of all the first-order partial variances.

The second-order Sobol index

$$S_{m_1m_2} = \frac{\sigma_{m_1m_2}^2}{\sigma^2}$$

(20)

describes the sensitivity of the output to the interactions between the $m_1$-th and $m_2$-th uncertain parameters ($m_1, m_2 \in \{1, \ldots, M\}$, in which

<table>
<thead>
<tr>
<th></th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a$ [m/s]</td>
<td>347</td>
<td>382</td>
</tr>
<tr>
<td>$t_c$ [s]</td>
<td>0.059</td>
<td>0.091</td>
</tr>
<tr>
<td>$f_{ow}$</td>
<td>0.015</td>
<td>0.025</td>
</tr>
<tr>
<td>$e$ [m]</td>
<td>0.0051</td>
<td>0.0057</td>
</tr>
<tr>
<td>$H_{UH}$ [m]</td>
<td>39.975</td>
<td>40.025</td>
</tr>
<tr>
<td>$J$ [$10^{-10}$ Pa$^{-1}$]</td>
<td>[0.74, 0.58, 0.56, 0.29, 0.074]</td>
<td>[1.14, 1.56, 0.97, 0.59, 0.12]</td>
</tr>
<tr>
<td>$\tau$ [s]</td>
<td>[0.056, 0.32, 1.37, 3.93, 9.10]</td>
<td>[0.089, 0.57, 2.01, 5.77, 12.06]</td>
</tr>
</tbody>
</table>

| Table 1 | Ranges of Uncertain Parameters Considered in the Example |

<table>
<thead>
<tr>
<th></th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a$ [m/s]</td>
<td>347</td>
<td>382</td>
</tr>
<tr>
<td>$t_c$ [s]</td>
<td>0.059</td>
<td>0.091</td>
</tr>
<tr>
<td>$f_{ow}$</td>
<td>0.015</td>
<td>0.025</td>
</tr>
<tr>
<td>$e$ [m]</td>
<td>0.0051</td>
<td>0.0057</td>
</tr>
<tr>
<td>$H_{UH}$ [m]</td>
<td>39.975</td>
<td>40.025</td>
</tr>
<tr>
<td>$J$ [$10^{-10}$ Pa$^{-1}$]</td>
<td>[0.74, 0.58, 0.56, 0.29, 0.074]</td>
<td>[1.14, 1.56, 0.97, 0.59, 0.12]</td>
</tr>
<tr>
<td>$\tau$ [s]</td>
<td>[0.056, 0.32, 1.37, 3.93, 9.10]</td>
<td>[0.089, 0.57, 2.01, 5.77, 12.06]</td>
</tr>
</tbody>
</table>

Figure 6. Range of pressure head signal at the downstream of the pipe in the presence of uncertainty. The ranges of the uncertain parameters are given in Table 1.
\[ \sigma_{m_1m_2}^2 = \text{Var}_{u_{m_1}, u_{m_2}} \left( E_{u_{m_1}, u_{m_2}}(h | u_{m_1}, u_{m_2}) \right) \approx \frac{1}{N_{MC}} \sum_{n=1}^{N_{MC}} \tilde{h}_n(\bar{\Psi}_1) \tilde{h}_n(\bar{\Psi}_{m_1m_2}) - \bar{h}_n(\bar{\Psi}_1) \]  

\[ (21) \]

\( \sigma_{m_1m_2}^2 \) denotes the partial variance corresponding to the interaction, \( \Psi_1^{m_1m_2} \) is the first sample set \( \Psi_1 \) where the \( m_1 \)-th and \( m_2 \)-th columns are replaced by the corresponding columns of the second sample set \( \Psi_2 \). Higher-order interactions can be similarly considered; in total \( 2^M - 1 \) Sobol indices can be defined. The sum of all the \( 2^M - 1 \) Sobol indices equals to 1.
Then, the total Sobol indices (Homma & Saltelli, 1996; Saltelli et al., 2010; Sobol, 2001), which describe the total contribution of each uncertain parameter to the uncertainty of QOI, can be computed by

\[
S_{Tm} = \mathbb{E}_{\mathbf{u}_{\sim m}} \left( \frac{\text{Var}_{\mathbf{u}_{\sim m}}(h | \mathbf{u}_{\sim m})}{\sigma^2} \right)
= S_m + \sum_{m_1 \neq m} S_{mm_1} + \sum_{m_1 \neq m, m_2 \neq m, m_1 > m_2} S_{mm_1m_2} + \cdots
\] (22)

for \( m = 1, \ldots, M \). Note that for a large number \( M \) of uncertain parameters, this computation is troublesome which needs to compute all the \( 2^M - 1 \) Sobol indices. Alternatively, the total Sobol indices can be approximated by (Saltelli et al., 2010):

\[
S_{Tm} \approx \frac{1}{2\sigma^2 N_{MC}} \sum_{n=1}^{N_{MC}} \left( \tilde{h}_e(\Psi^m_1) - \tilde{h}_e(\Psi_1) \right)^2.
\] (23)

Figure 8. Relative forecast error of transient signal using PCE, ordinary Kriging and PC-Kriging with (a) \( N = 40 \) and (b) \( N = 100 \) samples. PCE, polynomial chaos expansion; PC-Kriging, polynomial-chaos-Kriging.
Results

4.1. Numerical Setup and Uncertainty Assumption

In this section, a numerical example is considered to illustrate the surrogate modeling, UQ, and GSA problems. A reservoir-pipe-valve system with pipe length $l = 142$ m and internal diameter $d = 0.0792$ m is considered. In order to study the influence of uncertainties on the transient-based defect detection, a leak is also assumed in the pipe whose location is $x = 45$ m from the upstream boundary and the effective leak size (the lumped leak parameter) is $5 \times 10^{-5}$ m$^2$. In total 15 system parameters are assumed as uncertain, i.e., $M = 15$, as is shown in Table 1. The ranges of the (elastic) wave speed $a$, the valve closure time $t_c$ and the VEPs $J$ and $\tau$ are given according to the experimental results in Section 2.1; the range of the pressure head at the upstream reservoir $H_{UH}$ is estimated from the fluctuation interval of the pre-transient signal in the experimental results; the ranges of the Darcy-Weisbach friction factor $f_{DW}$ and the pipe wall thickness $e$ are given empirically. The corresponding ranges of the pressure head outputs at the downstream end of the pipe, given the uncertainties of system parameters in Table 1, are shown in Figure 6. Three samples of pressure head signal given specific input parameters are also plotted. It is clear that, as the wave propagates, the magnitude of uncertainty increases. In the following, the causes and the contributions of various uncertainty sources are analyzed using the proposed methodology.

4.2. Surrogate Modeling and its Reliability Assessment

Here, the Sobol sequence $u_n = (u_{1n}, ..., u_{6n}), n = 1, ..., N$, is taken from the 15-dimensional space of uncertain parameters. For each $u_n$, the system output, i.e., the pressure head signal, is numerically simulated via MOC (cf. Appendix A). Here, the simulation lasts 5 s and the sampling frequency is 1,000 Hz. For each time step $t$, a surrogate model of pressure head $h(t; u)$ at the pipe downstream end is built up by $h(t; u_n) (n = 1, ..., N)$, using respectively the PCE, ordinary Kriging and PC-Kriging approaches. In PCE, Hermite polynomials with maximum degree $p = 2$ and index set norm $q = 0.75$ are used. The reliability of each surrogate model versus the sample size $N$ is verified by its accuracy of forecast at new samples.

First, a new sample $\bar{u}$ of $u$ is tested where $a = 372.6$ m/s, $t_c = 0.0596$ s, $f_{DW} = 0.0155$, $e = 5.4$ mm, $H_{UH} = 40.014$ m, $J = [1, 0.80, 0.95, 0.52, 0.10] \times 10^{-10}$ Pa$^{-1}$, $\tau = [0.072, 0.45, 1.66, 4.40, 11.42]$ s. This sample is arbitrarily chosen and not one of the samples used in the surrogate modeling. The forecast of the pressure head signal using the PCE, ordinary Kriging, and PC-Kriging models built with $N = 40$ and $N = 100$ samples is shown in Figure 7. It can be seen that, for $N = 40$, PCE and PC-Kriging have a large forecast error, while the ordinary Kriging method can accurately describe the trend of the signal but has many local fluctuations. When
$N = 100$, all the three methods are improved, among which the ordinary Kriging method still outperforms PCE and PC-Kriging and can replicate both the overall trend and details of the transient signal.

Then, in order to more systematically assess the surrogate models’ reliability, $N_{FC} = 100$ samples of uncertain parameters are generated at random. The relative forecast error in Equation 15 versus $t$ for $N = 40$ and $N = 100$ is plotted in Figure 8. Furthermore, the relative forecast error averaged over $t \in [0, 5]$ s versus the sample size for surrogate modeling $N \in \{20, 40, 60, 80, 100\}$ is demonstrated in Figure 9. It is clear that for a low sample size $N$, the PCE and PC-Kriging methods return almost the same result, while as $N$ increases, the PC-Kriging tends to the ordinary Kriging. However, for $N \leq 100$, the ordinary Kriging approach is uniformly optimal and the forecast accuracy is satisfactory. Therefore, the ordinary Kriging-based surrogate model with $N = 100$ is used in the following for UQ and GSA.

Note that, before a surrogate model is applied, its reliability and the required sample size $N$ have to be tested. They are problem-dependent due to the shape (nonlinearity) of system output and the number

---

**Figure 10.** Convergence of (centralized) mean and standard deviation versus the sample size. The uncertain parameters are assumed to be uniformly distributed and independent.
of uncertain parameters. It has been shown in (Khazaie et al., 2019; Schobi et al., 2015) that PC-Kriging is more efficient than the PCE and ordinary Kriging methods for some types of output functions, but for some other problems, PC-Kriging is not the optimal choice due to the effect of over-fitting (Schobi et al., 2015).

4.3. UQ and GSA in the Case of Independent Uncertain Parameters

It is first assumed that only the ranges of the uncertain parameters are known. According to the principle of maximum entropy, the uncertain parameters are assumed to be independent and follow uniform distributions. The supports of the uniform distributions are given as in Table 1. In this case, a 15-dimensional Sobol sequence is taken to provide the new samples of uncertain parameters for UQ and GSA, which follow the independent uniform distributions. The corresponding output signals $h(t)$ are computed from the ordinary Kriging surrogate model. Note that the sample size $N_{UQ}$ for UQ and GSA should be large enough to guarantee the convergence. Here, the convergence of the mean and standard deviation of $h(t)$ with regard to the sample size $N_{UQ}$ is shown in Figure 10, which illustrates that the convergence is satisfactory when $N_{UQ} \geq 100$ for almost all $t$. Then, the full statistical information of the pressure head at each step can be estimated. Probability density functions (PDFs) of pressure head at $t = 0.01, 0.02, ..., 5$ s are plotted in Figure 11. It can be seen that the PDF tends to be flat and low as $t$ increases and at reflection-arriving times which implies a higher level of uncertainty.

The total Sobol indices for the 15 uncertain parameters are computed from Equation 23 and plotted in Figure 12. Figure 13 illustrates the total Sobol indices averaged with time for the whole signal ($t \in [0, 5]$ s), for the leak reflection section ($t \in [0.5, 0.7]$ s) and for the initial surge part ($t \in [0, 0.09]$ s). The latter two time intervals are important because they include most information for leak detection (Wang, Waqar, et al., 2020) and for structural reliability design (Zhang et al., 2011), respectively. It can be seen that in the initial time period, the valve closure time $t_c$ contributes mostly to the output uncertainties, while for the forecast of the whole signal and leak detection, the wave speed $a$ and the VEPs $J$ are more crucial. Generally speaking, the importance of $J_j$ ($j = 1, ..., 5$) decreases with $j$, because a higher order in the generalized K-V model corresponds to a larger retardation time which plays no role in a short time. However, the total Sobol index of $J_2$
is higher than $J_1$, this is because the uncertainty range of $J_2$ is much larger than $J_1$ (cf. Table 1 or Figure 4). Furthermore, the results clearly indicate that the contribution from the retardation times $\tau$ to transient wave uncertainty is much lower than the creep compliance functions $J$ and the boundary condition $H_0^U$ is negligible at any time during the transient period.

**4.4. UQ and GSA in the Case of Correlated Uncertain Parameters**

In Section 4.3, uncertainty and global sensitivity are analyzed based on the assumption that the uncertain parameters are independent. In practice, however, this assumption does not always hold. Especially, the viscoelastic parameters $J$ and $\tau$, which do not have a rigorous physical meaning but jointly simulate the mechanical behavior of pipe wall viscoelasticity by means of a combination of conceptual elements, are proven to have strong interdependence (Wang, Lin, Ghidaoui, Meniconi, Brunone, 2020). In this section, the viscoelastic parameters $J$ and $\tau$ are assumed to be correlated. The mean and variance of all the uncertain parameters are listed in Table 2: They are given according to the experimental results in Section 2.1.
for \( a, t, J, \tau \) and \( H_0^{\uparrow} \) and given empirically for \( f_{\text{low}} \) and \( e \). According to the principle of maximum entropy, all the 15 uncertain parameters are assumed to be Gaussian-distributed. For each uncertain parameter, the mean is much larger than the standard deviation, which guarantees that meaningless parameters (e.g., negative wave speed, pipe wall thickness, pressure head, and valve closure time) would not be generated in the stochastic simulation. It is further assumed that \( a, t, f_{\text{low}}, e \) and \( H_0^{\uparrow} \) are independent with each other and with \( J \) and \( \tau \). The correlation matrix of \((J, \tau)\) is given from the experimental results in Section 2.4, being

Figure 13. Average total Sobol index for the whole signal \((t \in [0, 5] \text{ s})\), the leak reflection section \((t \in [0.5, 0.7] \text{ s})\) and the initial surge part \((t \in [0, 0.09] \text{ s})\). The uncertain parameters are assumed to be independent and uniformly distributed with supports given in Table 1.
The MC simulation with sample size $N_{UQ} = 2,000$ is employed. Note that here the output pressure head signal $h(t)$ is computed from the ordinary-Kriging-based surrogate model, thus the computation is analytical and fast, which is not the case of the MC method in (Duan, 2015; Duan, Tung, & Ghidaoui, 2010; Huang et al., 2017; Zhang et al., 2011) where every computation of $h(t)$ calls the numerical code. The convergence of mean and standard deviation with sample size $N_{UQ}$ is shown in Figure 14. It is clear that the convergence, in this case, needs $N_{UQ} \geq 1,000$ and is much slower than in Section 4.3, this is mainly because the MC simulation is less efficient than the QMC via Sobol sequence. PDFs of pressure head at $t = 0.01, 0.02, …, 5$ s are displayed in Figure 15. It is clear that, in this case, the PDFs are sharper (the variances are lower) than in Figure 11. Besides, in this case, the PDFs look more like Gaussian-distributed especially for the initial time steps; on the contrary, for the independent case in Figure 11, the PDFs for the initial time steps are flatter and have approximately the shape of uniform distribution.

Figure 16 plots the total Sobol indices for $e$, $t_c$, $f_{DW}$, $a$, $H_0^U$, and VEPs. Here, since the VEPs $J$ and $\tau$ are assumed to be correlated, they are considered as a whole. Figure 17 shows the total Sobol indices averaged over the whole signal ($t \in [0, 5]$ s), the leak reflection section ($t \in [0.5, 0.7]$ s) and the initial surge part ($t \in [0, 0.09]$ s). It can be seen that the interdependence of the uncertain parameters does not significantly change the order of contributions: the VEPs and wave speed $a$ are the most crucial parameters for the whole signal and for the leak signature, while the valve closure time $t_c$ dominates the initial pressure surge part.

### 5. Conclusion

The sources of uncertainties in pressurized water-pipe systems and their propagation with transient waves are investigated. The statistical information of uncertain parameters is obtained via experimental data from a laboratory pipe system. Surrogate transient wave models are built, where the PCE, ordinary-Kriging, and PC-Kriging surrogate modeling methods are tested and compared. The efficient

\[
\Sigma_{J, \tau} = \begin{bmatrix}
1 & -0.94 & 0.31 & -0.46 & -0.42 & 0.53 & 0.36 & 0.57 & 0.61 & -0.28 \\
-0.94 & 1 & -0.36 & 0.52 & 0.45 & -0.72 & -0.08 & -0.78 & -0.63 & 0.21 \\
0.31 & -0.36 & 1 & -0.45 & -0.43 & 0.15 & 0.06 & 0.20 & 0.39 & -0.35 \\
-0.46 & 0.52 & -0.45 & 1 & 0.78 & -0.73 & 0.18 & -0.53 & -0.62 & 0.38 \\
-0.42 & 0.45 & -0.43 & 0.78 & 1 & -0.60 & 0.06 & -0.49 & -0.74 & 0.57 \\
0.53 & -0.72 & 0.15 & -0.73 & -0.60 & 1 & -0.50 & 0.86 & 0.63 & -0.26 \\
0.36 & -0.08 & 0.06 & 0.18 & 0.06 & -0.50 & 1 & -0.47 & -0.05 & -0.15 \\
0.57 & -0.78 & 0.20 & -0.53 & -0.49 & 0.86 & -0.47 & 1 & 0.57 & -0.10 \\
0.61 & -0.63 & 0.39 & -0.62 & -0.74 & 0.63 & -0.05 & 0.57 & 1 & -0.68 \\
-0.28 & 0.21 & -0.35 & 0.38 & 0.57 & -0.26 & -0.15 & -0.10 & -0.68 & 1
\end{bmatrix}
\]

(24)

### Table 2

**Means and Standard Deviations of the Uncertain Parameters**

<table>
<thead>
<tr>
<th>Parameter (Unit)</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a$ [m/s]</td>
<td>364.5</td>
<td>10.1</td>
</tr>
<tr>
<td>$t_c$ [s]</td>
<td>0.075</td>
<td>0.0092</td>
</tr>
<tr>
<td>$f_{DW}$</td>
<td>0.02</td>
<td>0.0029</td>
</tr>
<tr>
<td>$e$ [m]</td>
<td>0.0054</td>
<td>0.0002</td>
</tr>
<tr>
<td>$H_0^U$ [m]</td>
<td>40.0</td>
<td>0.0144</td>
</tr>
<tr>
<td>$J$ [10^{-10} Pa^{-1}]</td>
<td>[0.97, 1.04, 0.82, 0.49, 0.10]</td>
<td>[0.13, 0.29, 0.10, 0.055, 0.0089]</td>
</tr>
<tr>
<td>$\tau$ [s]</td>
<td>[0.068, 0.42, 1.64, 5.24, 10.44]</td>
<td>[0.011, 0.070, 0.17, 0.40, 0.52]</td>
</tr>
</tbody>
</table>
surrogate model makes the fast computation of uncertainty quantification and global sensitivity analysis possible. The full probability distribution of transient signal at each time step is derived; the contributions of various system uncertain parameters are quantified via the Sobol indices. The proposed methodology is illustrated via an example of a reservoir-pipe-valve system where uncertainties on 15 parameters are considered. It is found that the wave speed and the creep compliance functions of pipe wall viscoelasticity are dominant in forming the uncertainty of transient signal, especially at the signature of leak. This implies that quantifying the uncertainties of wave speed and pipe viscoelasticity would be essential in the leak detection procedure, while other sources of uncertainty can be reasonably neglected.

In this paper, the uncertainties of transient waves are illustrated via a laboratory pipe system. In real urban water supply systems, however, the uncertainties are often much more complex, where valve and pump activities, water consumption from users, and pipe geometry might be unknown or imprecisely known. In

![Figure 14: Convergence of (centralized) mean and standard deviation versus the sample size where the uncertain parameters are correlated.](image-url)
these cases, more parameters are needed to describe the system uncertainties. Applications of the proposed uncertainty quantification and global sensitivity analysis methods in these more complicated scenarios would be an interesting and logical next step.

Figure 15. Probability density functions of pressure head at $t = 0.01, 0.02, \ldots, 5$ s. The uncertain parameters are assumed to be Gaussian-distributed and correlated.

Figure 16. Total Sobol indices where the uncertain parameters are correlated and Gaussian-distributed with mean and standard deviation given in Table 2.
Appendix A: Physical model of transient wave and its numerical simulation

Here, the physical model of transient wave in viscoelastic pipes is briefly introduced. It is assumed that the creep function $J(t)$ of a viscoelastic material follows the generalized Kelvin-Voigt (K-V) model (Covas et al., 2004, 2005; Duan, Ghidaoui, et al., 2010; Keramat et al., 2012; Meniconi et al., 2013; Pezzinga et al., 2016):

$$J(t) = J_0 + \sum_{j=1}^{N_{KV}} J_j (1 - \exp(-t / \tau_j)).$$

(A1)

where $J_j$ is the creep compliance, $\tau_j$ is the retardation time, and $N_{KV}$ is the number of elements.
Let \( q \) and \( h \) denote the discharge and head, respectively. The unsteady-oscillatory continuity and momentum equations with time \( t \) and spatial coordinate \( x \) read (Wang, Lin, Keramat et al., 2019):

\[
\frac{\partial q}{\partial x} + \frac{qA}{2\xi} \frac{\partial h}{\partial t} + (1 - \nu^2) \rho \frac{gdA}{\xi} \sum_{j=1}^{NKV} \frac{1}{\tau_j} \frac{\partial}{\partial t} \int_0^t h(t - s) \exp(-t / \tau_j) ds = 0; \tag{A2}
\]

\[
\frac{1}{gA} \frac{\partial q}{\partial t} + \frac{\partial h}{\partial x} + \frac{f_{DW}q}{2gA^2} = 0. \tag{A3}
\]

Here,

\[
a = \left( \rho \left( \frac{1}{\kappa} + (1 - \nu^2) \frac{d}{Ee} \right) \right)^{\frac{1}{2}} \tag{A4}
\]

is the elastic component of wave speed, \( \kappa \) and \( \rho \) are the bulk modulus and density of fluid, \( E \) is the Young’s modulus, \( \nu \) is the Poisson’s ratio of pipe, \( A \) is the cross-sectional area of pipe, \( f_{DW} \) is the Darcy-Weisbach friction factor.

The boundary conditions can be given according to the property of the system. If the upstream end \( x^U \) is connected to a reservoir or a pump, the pressure head oscillation is assumed to be \( h(t, x^U) = 0 \). The downstream end \( x^D \) is a wave-generation valve, the boundary condition \( q(t, x^D) \) can be given according to the pattern of valve operation. If a leak exists in the pipe, the pressure and mass conservation across the leak is considered:

\[
h(x_-^L) = h(x^L_+) = h(x^L); \tag{A5}
\]

\[
q(x_-^L) = q(x^L_+) = q(x^L) + s^L \sqrt{2g \left( h(x^L) - z^L \right)}, \tag{A6}
\]

in which \( x_-^L \) and \( x^L_+ \) represent respectively just upstream and downstream of the leak, \( z^L \) is the elevation of the pipe at the leak, and \( s^L \) is the leak size (the lumped leak parameter).

Together with Equation A2, A3, A5, A6, and the boundary conditions \( h(t, x^U) \) and \( q(t, x^D) \), the transient wave propagation in a viscoelastic pipe can be simulated using the method of characteristics (MOC; Covas et al., 2005). The integral in Equation A2 and its time-derivative are numerically computed (Covas et al., 2005).

**Nomenclature**

- \( q \) [m³/s]: transient discharge
- \( h \) [m]: transient pressure head
- \( Q_0 \) [m³/s]: steady-state discharge
- \( H_0 \) [m]: steady-state pressure head
- \( H_0^U \) [m]: pressure head at the upstream boundary
- \( a \) [m/s]: (elastic) wave speed
- \( A \) [m²]: cross-sectional area of pipe
- \( l \) [m]: pipe length
- \( d \) [m]: internal pipe diameter
- \( e \) [m]: pipe wall thickness
- \( t_c \) [s]: valve closure time
- \( J = (J_{1}, ..., J_{NKV}) \) [Pa⁻¹]: creep compliance functions in the generalized K-V model
- \( \tau = (\tau_1, ..., \tau_{NKV}) \) [s]: retardation times in the generalized K-V model
- \( u = (u_1, ..., u_M) \): a sample of \( u \)
- \( M \): number of uncertain parameters
Acknowledgments

This work has been supported by the National Natural Science Foundation of China (Grant No. 52005016) and the State Key Laboratory of Mechanical System and Vibration (Grant No. MSV202101). The author thanks M. S. Ghidaoui, G. A. Camino, M. Y. Lam, C. Y. Sze, and S. H. Chan for the transient experiments and thanks M. Waqar for plotting the 3D pipe system map in Figure 1.

References


Data Availability Statement

The experimental data used in this manuscript are available at Mathwork (https://ww2.mathworks.cn/matlabcentral/profile/authors/4659871).


