Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Pipe Network Parameter Estimation Using a Wave Propagation Path Strategy

Aaron C. Zecchin
The University of Adelaide
Nhu Do
The University of Adelaide
Wei Zeng
The University of Adelaide
Alireza Keramat
alireza.keramat@adelaide.edu.au
The University of Adelaide https://orcid.org/0000-0002-6280-4931
Martin F. Lambert
The University of Adelaide

Research Article

Keywords: Pipeline networks, Parameter estimation, Wave-propagation-path approach, Maximum likelihood estimation, Hydraulic model calibration

Posted Date: July 17th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10153162/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Pipe Network Parameter Estimation Using a Wave Propagation Path Strategy

Aaron C. Zecchin[1,*], Nhu Do[2], Wei Zeng[3], Alireza Keramat[4], Martin F. Lambert[5]

[1]Senior Lecturer, School of Civil Engineering and Construction, Adelaide University, SA 5005, Australia ORCID: https://orcid.org/0000-0001-8908-7023. Email: aaron.zecchin@adelaide.edu.au

[2]Adjunct Lecturer, School of Civil Engineering and Construction, Adelaide University, SA 5005, Australia. Email: nhu.do@adelaide.edu.au

[3]Senior Lecturer, School of Civil Engineering and Construction, Adelaide University, SA 5005, Australia. Email: w.zeng@adelaide.edu.au

[4]Research Fellow, School of Civil Engineering and Construction, Adelaide University, SA 5005, Australia. Email: alireza.keramat@adelaide.edu.au

[5]Emeritus Professor, School of Civil Engineering and Construction, Adelaide University, SA 5005, Australia. Email: martin.lambert@adelaide.edu.au

*Corresponding Author

Abstract

This paper presents a novel parameter estimation strategy in pipeline networks based on a wave-propagation-path approach. Unlike conventional inverse analysis, the method focuses exclusively on the coarse features of pressure signals, including the wavefront delay and attenuation, thus avoiding the need for complex nonlinear hydraulic equations, which are often subject to significant uncertainties. To address parametric uncertainty and measurement error,

the framework employs a probabilistic formulation: Maximum Likelihood Estimation (MLE) is first used to identify the most likely wave propagation paths, followed by maximum a posteriori estimator to efficiently estimate wave speeds and attenuation factors for individual pipes. Two case studies were conducted to demonstrate the capability of the methodology. The first provides detailed insights into the estimation process and demonstrate robustness to uncertainty, while the second applies the method to a realistic-scale urban water distribution system. The results show that the approach can significantly improve estimates of wave speeds and friction factors, particularly for pipes located along dominant propagation paths. The proposed framework provides a scalable and efficient alternative for hydraulic model calibration in complex networks, with the estimated wave speeds offering valuable information for network-level condition assessment.

Keywords: Pipeline networks; Parameter estimation; Wave-propagation-path approach; Maximum likelihood estimation; Hydraulic model calibration

INTRODUCTION

Managing and operating a water distribution system (WDS) requires reliable knowledge of network characteristics and hydraulic conditions. However, only a small portion of a real WDS can be directly observed or measured because of its spatial complexity, buried infrastructure, and limited instrumentation. Therefore, the overall hydraulic state of a network is commonly inferred using hydraulic models, which must be calibrated before being used for operational decision-making, planning, diagnosis, and asset management (Zhang et al., 2020).

Calibration is the process of defining hydraulic model parameters so that computed model responses match measured system data under specified hydraulic conditions (Savic et al., 2009). Conventionally, calibration is performed by solving a set of hydraulic equations, including continuity, energy, and momentum equations, under steady-state, extended-period simulation (EPS), or transient conditions (Lee et al., 2024; Zhou, 2018). A simpler version of

calibration can be implemented using steady-state hydraulic data, although this is mainly applicable to estimating pipe resistance coefficients, such as friction factors, and demand-related parameters (Asadzadeh et al., 2010; Khedr et al., 2015).

Steady-state and EPS models simplify the hydraulic behaviour by assuming that system variables, such as flow rates and hydraulic grades, are time-invariant or quasi-steady. This simplification enables network-scale calibration using either explicit solution techniques, such as Newton–Raphson-based hydraulic solvers, or implicit approaches in which calibration is formulated as an optimization problem. Various parameters may be considered in this process, including water demand (Do et al., 2017a; Preis et al., 2009; Shang et al., 2008), pipe roughness (Alvisi & Franchini, 2010; Lansey et al., 2001; Ormsbee et al., 1986), and valve status (Do et al., 2018; Piller & van Zyl, 2014). Despite their practical value, steady-state and EPS calibration methods face two major limitations: first, the number of identifiable calibration parameters is restricted by the number and quality of available measurements; and second, measurement errors, uncertain boundary conditions, and model structural errors are often ignored or only partially represented. These limitations can lead to non-unique or approximate representations of real systems, reducing the reliability of calibrated models for operational use.

Transient-based methods provide an alternative source of information because pressure waves carry dynamic signatures related to pipe properties, hydraulic parameters, defects, and boundary conditions. Several studies have demonstrated the diagnostic value of transient pressure waves in pipe systems. Meniconi et al. (2011) showed that small-amplitude sharp pressure waves can be used as a minimally invasive diagnostic tool for pipe systems. Massari et al. (2013) developed a stochastic successive linear estimator for pipe-system diagnosis, demonstrating the relevance of probabilistic inverse estimation when hydraulic response data are uncertain. Duan (2015) further showed that modelling and parameter uncertainties can significantly affect transient-flow modelling and transient-based leak detection, highlighting the need for uncertainty-aware diagnostic methods. The attenuation of transient pressure waves

has also been investigated as a useful diagnostic feature; for example, Capponi et al. (2020) analysed laboratory experiments on transient-pressure damping in a leaky polymeric pipe and showed that damping behaviour contains valuable information about leakage and pipe-wall response. More recently, Zhang et al. (2020) formulated pipeline condition assessment as a Bayesian inverse transient analysis problem, enabling parameter estimation and uncertainty quantification within a single probabilistic framework. In a related uncertainty-aware context, Blocher et al. (2021) investigated the influence of prior assumptions on leak localisation in water distribution networks, showing that prior information can strongly affect the reliability of inverse hydraulic assessments.

In terms of transient modelling, calibration and parameter identification have mainly been applied in the context of pipe condition assessment, leak detection, and defect localisation. Many of these approaches have focused on single pipeline systems with either measured or known boundary conditions (e.g., Meniconi, 2022; Wang et al., 2002). Only a limited number of methods have dealt with fluid-line networks of general topology, including the time-domain inverse transient method (Liggett & Chen, 1994; Sarkamaryan et al., 2018; Zecchin et al., 2013), the frequency-domain impedance matrix method (Kim, 2016), and maximum likelihood estimation approaches based on Laplace-domain network admittance matrix models (Zecchin et al., 2009, 2013). Although these methods have advanced transient-based diagnosis, they often require detailed observation-model matching and the solution of nonlinear hydraulic transient equations. Consequently, they may encounter significant challenges associated with uncertain boundary conditions, transient modelling errors, non-convex optimization, and a large number of unknown parameters (Bohorquez et al., 2020; Keramat et al., 2022; Zeng et al., 2018).

These challenges become more important in complex and realistic water distribution systems. In divide-and-conquer approaches, for instance, a city-scale network may be divided into smaller target networks for local calibration or diagnosis. However, these target networks are typically connected to surrounding systems containing unknown or uncertain hydraulic

components. This can introduce errors into classical observation-model matching procedures, especially when full transient pressure histories are used. Although recent studies have attempted to reduce non-uniqueness and improve uncertainty representation, many existing methods still rely on detailed transient signals that are inherently affected by both epistemic and aleatory uncertainties (Hutton et al., 2014). Therefore, there is a need for parameter estimation methods that use robust and informative signal features while avoiding excessive dependence on detailed nonlinear transient modelling.

To address this need, this paper proposes a wave-propagation-path-based parameter estimation strategy for pipeline networks. Instead of matching complete transient pressure histories, the proposed method focuses on coarse but physically meaningful features of the pressure response, namely wavefront delay and attenuation. These features are extracted from transient events and interpreted within a graph-theoretic framework based on wave propagation paths and shortest-path calculations (Dijkstra, 1959). In this way, the method avoids the need to solve the full set of nonlinear hydraulic transient equations while retaining the key wave-propagation information required for estimating network parameters.

The proposed framework incorporates both parametric uncertainty and measurement error. Maximum likelihood estimation (MLE) is first used to identify the most likely wave propagation paths associated with the observed transient features. A maximum a posteriori (MAP) estimator is then developed to estimate pipe-level wave speeds and attenuation factors by combining measurement information with prior knowledge. The estimated wave speeds provide useful information for network-level condition assessment, while the attenuation factors offer additional insight into the dissipative behaviour of the system. The main contribution of this study is therefore a scalable probabilistic framework for estimating hydraulic wave-propagation parameters in complex pipe networks using only wavefront delay and attenuation information. Compared with conventional inverse transient analysis, the method reduces computational demand, avoids detailed full-signal matching, and explicitly incorporates prior information and measurement uncertainty.

The remainder of the paper is organized as follows. Section 2 revisits the parameter estimation problem in pipeline networks and highlights the main challenges associated with transient-based identification under uncertainty. Section 3 presents the proposed wave-propagation-path framework, including the MLE- and MAP-based formulations. Section 4 provides two illustrative case studies to demonstrate the estimation process, assess robustness to uncertainty, and examine the applicability of the method to realistic-scale water distribution systems. Finally, Section 5 summarizes the main conclusions and discusses future research directions.

REVISITING THE PROBLEM OF PARAMETER ESTIMATION

Consider a pipeline network represented as the couple $\{G(\boldsymbol{N},\boldsymbol{\Lambda}),\boldsymbol{\theta}\}$ where the graph $G(\boldsymbol{N},\boldsymbol{\Lambda})$ consists of node set N (i.e. the pipeline junctions $N_1,\ldots,N_n$) and link set $\boldsymbol{\Lambda}$ (the set of network pipes $L_1,\ldots,L_m$) and describes the network topology, and the parameter set $\boldsymbol{\theta}$ contains all important network parameters (e.g. pipeline wave speed and roughness). For this network, there exists a series of hydraulic transient experiments, numbered $i = 1, \ldots, E$ with a pressure generator located at node m_{i0} and sensors located at nodes m_{ij}, $j = 1, \ldots, E_i$. The hydraulic transient experiments involve the injection of a step pressure wave into the network, and the measurement of the pressure response at the sensor nodes, and involve the data sets.

For parameter estimation problems, a general form of the estimation approaches have involved estimators of the form

$$\hat{\theta} = \arg\min_{\theta} J\left(\Phi\left(\mathbf{Y}\right), M\left(\theta,\theta',G\right)\right) \tag{1}$$

where $\Phi\left(\mathbf{Y}\right)$ represents some transform of the data $\mathbf{Y}$, $M\left(\theta,\theta',G\right)$ represents model predictions as dependent on the parameters, ancillary parameters θ' (e.g. boundary conditions, or noise statistics) and the network topology G, and J represents an error function between the two (note, from hereon, the symbol $\hat{z}$ will be used to signify an estimate of the parameter z). For example: for ITA applications, the data contains no transform, $M\left(\theta,\theta',G\right)$

are the model state predictions based on the time-domain simulator, and J is the typical sum of squared errors; for the frequency-domain maximum likelihood estimator, the Fourier transform is used on the data, the network admittance matrix model is used, and the frequency-domain likelihood function is used.

With reference to Eq. (1), the simulation-based calibration approaches (i.e. $M(\theta,\theta',G)$ is either a PDE solver or a frequency-domain model, which involves matching the detailed features of the model predictions to the measured time series data. In other word, previous hydraulic transient based parameter estimation techniques seek for the full solution of the water hammer's continuity and momentum equations elaborated in Appendix A. Note that the wave speed is directly related to the wall thickness and Young's modulus of elasticity (see Appendix A) which are key indicators of pipe wall structural conditions. Thus, the wave speed estimation in a network scale can provide valuable information for pipe condition assessment in a water network. In our statistical framework, this equation yields the mean values of the wave speeds, which, along with the associated uncertainty, will define the prior Gaussian distribution.

While the ITA approach makes the most use of all available information, it suffers significantly from modelling errors, as transient solvers can hardly represent complete dynamic behaviours including dissipation, dispersion, etc., especially in networks (Vrachimis et al. 2022). In addition, the boundary and initial conditions for solving Eq. (1) must be known to obtain a unique set of the model parameters (Imani et al. 2025; Zeng et al. 2025). However, on a network scale, these conditions are often undetermined, and assumptions made for these conditions directly affect the calibration results. However, the proposed method utilizes arrival times that do not require a sophisticated transient model for prediction, thus enabling a rapid network calibration process.

PROPOSED APPROACH

The objective with the proposed strategy is to use only coarse features of the measured signals to estimate key system parameters. The coarse features of interest are the wave propagation delay (as this is dependent on the pipeline wave speeds), and attenuation (as this is dependent on the pipeline resistance and visco-elastic properties).

Specifically, the proposed approach deals exclusively with properties of the initial wave front, namely its arrival time, and attenuation. The duration for a hydraulic disturbance to propagate along a pipe of length l is given by $T = \frac{l}{a}$, which is termed the pipeline delay time (pipeline period, time of flight) and is a key pipeline parameter in this work.

Concerning the attenuation, considering a semi-infinite pipeline of quiescent conditions with a boundary pressure disturbance $h(0,t) = u(t)$ initiated at $t = t'$, at some point $x = l$ downstream, the pressure wave form will arrive at time $t = T + t'$, for which its leading edge will have the form

$$h(l, T + t') = e^{-\alpha} u(t'), \qquad \alpha = \frac{1}{2}\left(\frac{R}{L} + \frac{C_R(0)}{C}\right) T \tag{2}$$

where $L = 1/gA$ is the hydraulic inertance (g is gravitational acceleration, A is cross-sectional area), $R = fv_0/gAD$ is the linear resistance (f is the Darcy-Weisbach friction factor, v_0 is the steady-state mean fluid velocity , and D is the diameter), $C = gA/a^2$ is the hydraulic capacitance, and $C_R = C_R(t)$ is the retarded capacitance function, which have been further discussed in Appendix A. The derivation of Eq. (2) is provided Appendix B.

In terms of field experiments, given a sharp wave injection at node m_{i0}, the properties of the initial wavefront will be presented by delay and attenuation measures. The experimental delay time estimates for experiment i and sensor node m_{ij}, are given by $\tilde{\tau}_{ij} = \tilde{t}_{ij} - \tilde{t}_{i0}$, where $\tilde{\tau}_{ij}$ is the observed delay for sensor j and experiment i (subscripts i and j have constant definition

for all terms), $\tilde{t}_{ij}$ is the observed arrival time for sensor *j* (the case of *j* = 0 represents the generator node; therefore, we will refer to the travel time between the generator and the sensor as the delay time). Additionally, the attenuation factor is given by $\tilde{\gamma}_{ij} = \log(\tilde{h}_{ij}) - \log(\tilde{h}_{i0})$ where $\tilde{h}_{ij}$ is the observed wave amplitude for sensor *j* (for reasons to be made clear, the natural logarithm of the ratio $\tilde{h}_{ij}/\tilde{h}_{i0}$ is used).

Propagation paths in networks

Pipe networks are simply systems of interconnected pipelines, whose individual dynamics are governed by Eq. (2), while their boundary conditions are governed by conservation of mass and equivalence of pressure (i.e. the junctions are assumed to be lossless). Point disturbances in the hydraulic state of pressure or flow within a single pipe will propagate throughout the network by travelling along the individual line and transmitting to incident pipelines through nodal connections. In this sense, a single disturbance within the network will cause a cascade of transmitted waves propagating throughout the entire network.

Given the connectivity of the pipes, and the transmission of pressure waves through the network junctions, a pressure disturbance initiated at node $i \in N$ will first reach a node $j \in N$ by following the shortest-delay path through the network, where a path is simply defined as a sequence of links. To describe this path first consider the following: for a given path *P* through graph $G(N,\Lambda)$, the time taken for a wave to travel along this path $\tau(P)$ is given by $\tau(P) = \sum_{\lambda \in P} T_\lambda$, where T_λ is the delay period for pipe λ (that is, the time for a wave to propagate the length of a pipe). Using this notation, a disturbance at node *i* occurring at time *t* = 0 will be first observed at node *j* at time $t = \tau(P^*)$ where P^* is the shortest delay path, and is given by $P^* = \arg\min_{P \in \boldsymbol{P}(i,j)} \{\tau(P)\}$, where $\boldsymbol{P}(i,j)$ is the set of all paths linking nodes *i* and *j*.

As a pressure wave propagates along a path throughout a network, it reduces in amplitude as it transmits through a junction part of the wave energy is reflected back up the pipeline, the

remaining energy is divided and transmitted into all pipes incident to the junction. Considering only the instantaneous elastic response of a junction, a pressure wave propagating through node *n* from pipe λ_i to pipe λ_j possesses the following transmission coefficient: $S_{ij} = \frac{2}{B_{\lambda_j} \sum_{\lambda \in \Lambda_n} \frac{1}{B_\lambda}}$

where $B = a/(gA)$ is the hydraulic series impedance (subscripts to the pipe numbers are omitted where the association is clear), and Λ_n is the set of pipes connected to node *n* (clearly $\lambda_i, \lambda_j \in \Lambda_n$).

Given the wave attenuation and junction losses, the entire attenuation along path $P = \left(\lambda_1, \ldots, \lambda_{|P|}\right)$ is given by

$$g(P) = \prod_{k=1}^{|P|-1} S_{\lambda_k \lambda_{k+1}} \prod_{\lambda \in P} e^{-\frac{1}{2}\left(\frac{R_\lambda}{L_\lambda} + \frac{C_{R\lambda}(0)}{C_\lambda}\right) T_\lambda} \tag{3}$$

where the first product relates to the junction losses, and the second product relates to the propagation losses.

Parametric Uncertainty

Within all real-world pipe networks, there is some uncertainty about the exact parameter values within the actual network. For example, the fluid properties vary with temperature and air content, the pipe material parameters also vary due to temperature effects, deterioration of the pipe wall, or variations in the material properties due to manufacturing imperfections, and variations in the restraint conditions due to improperly defined pipe wall restraints (i.e. soil interaction), degradation of the pipe restraints, or differences between plan specification and actual installation. This uncertainty can be considered as a knowledge uncertainty, as it pertains to the limitation in our knowledge of the system, as opposed to any inherent randomness in the system itself.

The key to calibration approaches is to determine a mapping from the parameters, through some model, to estimates of features of the measured data. For example, in simulation-based approaches, the data is either in time-domain or frequency-domain form, and the mapping from the parameters to estimates of the data occurs through the simulation model itself.

The formation of θ depends on the modelling strategy used, and what parameters are uniquely identifiable. Given that the proposed strategy is based purely on the properties of wave front as given by the water hammer model, the key processes that are being modelled are propagation delay, as represented by $T = l/a$, and attenuation which is quantified by the attenuation factor α introduced in Eq. (2). So, the key parameters that require estimation are T and α, which are aggregations of the fundamental physical parameters.

Typically, physical parameters can be prescribed lower and upper bounds, based on a priori reasoning (e.g. the friction factor can only exist within a limited range), and these provide the basis for determining the upper and lower bounds on the values that T and α can take. For the purpose of the ensuing analysis, it is convenient to represent the parametric uncertainty as being normally distributed with means $\mathrm{E}\left[\breve{T}\right] = \bar{T}$ and $\mathrm{E}\left[\breve{\alpha}\right] = \bar{\alpha}$ determined directly from a best estimate of the physical parameters, and variances $\mathrm{Var}\left[\breve{T}\right] = \sigma^2$ and $\mathrm{Var}\left[\breve{\alpha}\right] = \zeta^2$ determined from the upper and lower bounds of the physical parameters[1] (note, the symbol $\breve{z}$ is used to represent whenever the constant z is considered as a random variable).

For the purpose of the proposed method, it is convenient to characterise the wave speed uncertainty in terms of a propagation delay uncertainty. That is, a pipe's time-delay $\breve{T}$ is distributed according to the prior f_T, which encapsulates the knowledge uncertainty associated

[1] For example, the upper and lower bounds for the physical parameters can be mapped through to upper and lower bounds for T and α, and these bounds can be used to select σ and ζ such that the majority of the probability weighting for T and α, is within these bounds (e.g. setting $1.96\sigma = \left(T_{upper} - T_{lower}\right)/2$ allows for 95% of the probability weight for T to fall between the upper and lower bounds.

with the system parameters. For the purposes of the analysis, it is convenient to take $f_T = N\left(\overline{T}, \sigma^2\right)$, which enables the time-delay along path P to be normally distributed with

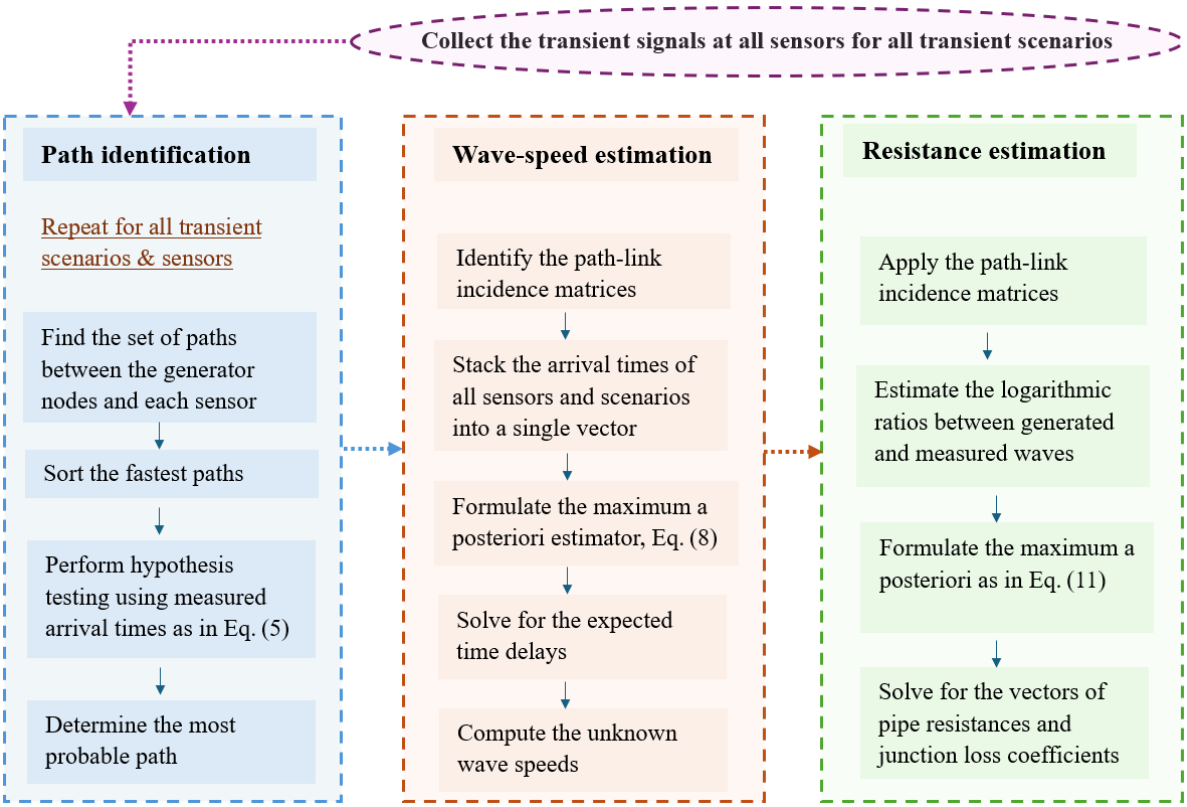
$$\mathrm{E}\left[\breve{\tau}\left(P\right)\right] = \sum_{\lambda \in P} \overline{T}_\lambda, \qquad \mathrm{Var}\left[\breve{\tau}\left(P\right)\right] = \sum_{\lambda \in P} \sigma_\lambda^2 \tag{4}$$

That is, the distribution of the path wave propagation delay can be expressed as a function of the individual pipe wave propagation delay terms.

For the attenuation, it is noted that there are the terms associated with the junction losses, and those associated with the propagation losses. In the proposed approach, the delays are used to first estimate the wave speeds, and so the estimation of the attenuation losses can be considered as conditional on the estimation of the propagation delays. As such, taking the natural logarithm of Eq. (3) yields $\gamma\left(P\right) = \log\left(g\left(P\right)\right) = \sum_{k=1}^{|P|-1} \log\left(S_{\lambda_k \lambda_{k+1}}\right) - \sum_{\lambda \in P} \alpha_\lambda$, where the distribution of $\breve{\gamma}\left(P\right)$, conditioned on the delays $\breve{T}_\lambda = T_\lambda, \lambda \in P$, is normally distributed with mean $\mathrm{E}\left[\breve{\gamma}\left(P\right) \middle| \left(\breve{T}_\lambda = T_\lambda\right)_{\lambda \in P}\right] = \sum_{k=1}^{|P|-1} \log\left(S_{\lambda_k \lambda_{k+1}}\right) - \sum_{\lambda \in P} \overline{\alpha}_\lambda$ and variance $\mathrm{Var}\left[\breve{\gamma}\left(P\right) \middle| \left(\breve{T}_\lambda = T_\lambda\right)_{\lambda \in P}\right] = \sum_{\lambda \in P} \varsigma_\lambda^2$.

Proposed Strategy

As shown in Figure 1, the proposed parameter identification framework involves three major steps which start with estimating the most likely wave propagation path for each generator-sensor pair in each experiment. This is achieved using a multi-array hypothesis testing approach corresponding to each feasible path. The second step is to estimate the pipe delays, which can be achieved using a maximum a posteriori estimator (MAP) approach (Garthwaite and Jolliffe 2002). Similar MAP process can be conducted in the third step to estimate the pipe attenuation factors.



- *Figure 1: Summary of the proposed parameter estimation process*

Path Identification

The purpose of this stage is to identify the propagation path for each generator-sensor pair for each experiment. That is for each experiment $i = 1, \ldots, E$ and each sensor $j = 1, \ldots, E_i$, there exists an observed propagation delay $\tilde{\tau}_{ij}$ from which the propagation path P_{ij} needs to be identified (recall that for experiment i and sensor j, the propagation path is from the generator node m_{i0} to the sensor node m_{ij}). For the case of complete parametric knowledge, P_{ij} is simply the shortest-delay path between the nodes as governed by the delay parameters. However, when the parameters are unknown (and priors are used to characterise the available information), the actual propagation path may differ from the path implied by the mean parameter values.

Specifically, in real practice, when the pipe network has blocked pipes (e.g., closed valves) preventing waves from propagating through the shortest path or when the estimated wave

speeds calculated using the theoretical formula (see Appendix A) are significantly inaccurate, other paths may serve as the actual propagation route. For example, let $P_{ij}^{[1]}$ and $P_{ij}^{[2]}$ be the firsts and second shortest paths as implied by setting the link delay to their mean values, then given the uncertainty associated with the pipe propagation delays, there is a non-zero probability that $\breve{\tau}\left(P_{ij}^{[1]}\right) > \breve{\tau}\left(P_{ij}^{[2]}\right)$, meaning that $P_{ij}^{[2]}$ is the shortest path.

For the experimental setup, consider the set of paths between the generator node m_0 and the sensor node m_j, $\left\{P^{[0]},\ldots,P^{[M-1]}\right\}$, which represent the fastest M paths according to the mean delays, that is $\mathrm{E}\left[\breve{\tau}\left(P^{[k]}\right)\right] < \mathrm{E}\left[\breve{\tau}\left(P^{[k+1]}\right)\right]$ for k = 0, …, M-2. Given the observation $\tilde{\tau}_{ij}$, it is required to identify the most likely path for each experiment i and sensor j. While, in general, the paths of two different sensors may influence each other's identifiability, for simplicity we assume they are statistically independent. This assumption allows us to determine the shortest path for each sensor separately. Consequently, the subscripts *i* and *j* (denoting the experiment number and sensor node, respectively) are omitted hereafter.

To identify the shortest path for each measured time $\tilde{\tau}$, an M-ary hypothesis testing approach is adopted, where H_m is the hypothesis that path $P^{[m]}$ is the actual path, m = 1, .., M-1. That if, for each hypothesis, the observation $\tilde{\tau}$ is assumed to take the form $\breve{\tilde{\tau}} = \breve{\tau}\left(P^{[m]}\right) + \breve{\varepsilon}$ where $\breve{\varepsilon} \sim N\left(0,\sigma_\tau^2\right)$ indicates the measurement error. Under the m-th hypotheses, the expectation is estimated using $\mathrm{E}\left[\breve{\tilde{\tau}}\middle| P^{[k]}\right] = \sum_{\lambda\in P^{[k]}} \bar{T}_\lambda$ and variance of the observation is determined as $\mathrm{Var}\left[\breve{\tilde{\tau}}\middle| P^{[k]}\right] = \sum_{\lambda\in P^{[k]}} \sigma_\lambda^2 + \sigma_\tau^2$. Defining the probability density function $f_\tau\left(\tau\middle|P\right) = N\left(\sum_{\lambda\in P}\bar{T}_\lambda, \sum_{\lambda\in P}\sigma_\lambda^2 + \sigma_\tau^2\right)$, the hypothesis set can be constructed as

$H_m : \tilde{\tilde{\tau}} \sim f_\tau\left(\tau \middle| P^{[m]}\right)$, $m = 0,..,M-1$ where the null hypothesis H_0 states that the deterministic shortest path is the actual shortest path.

If it is possible to define prior probabilities π_m for each hypothesis *m*=1, …, *M-1* (where $\sum_{m=0}^{M-1} \pi_m = 1$), then the following maximum a posteriori estimator can be used for estimating the true hypothesis $H_{\hat{m}}$

$$\hat{m}_{\text{MAP}} = \underset{m\in[0,\ldots,M-1]}{\text{argmax}} f_\tau\left(\tilde{\tau} \middle| P^{[m]}\right)\pi_m, \qquad f_\tau(\tau|P) = N\left(\sum_{\lambda\in P} \bar{T}_\lambda, \sum_{\lambda\in P} \sigma_\lambda^2 + \sigma_\tau^2\right) \tag{5}$$

where the prior π_m can be defined based on an engineering insight or via the results of previous experiments or iterations. Appendix C provides further discussion about potential priors.

If there is no prior belief about path *m* (i.e., treating π_m as flat or uniform), then Eq. (5) reduces to the maximum likelihood estimation. Note that the null hypothesis that corresponds to the shortest path (m=0) is independent of the observation $\tilde{\tau}$ and hence its variance, yet this algorithm finds the path that is most feasible considering the observed delay time of the path and associated variances.

The outcome of this stage computed from Eq. (5) is the identification of the path set $\widehat{\boldsymbol{P}} = \{\widehat{\boldsymbol{P}}_1, \ldots, \widehat{\boldsymbol{P}}_E\}$ for all experiments and sensors, where for experiment *i*, $\widehat{\boldsymbol{P}}_i = \{\hat{P}_{i1}, \ldots, \hat{P}_{iE_i}\}$, where E_i is the number of sensors used in the i-th experiment.

Wave speed estimation

For the wave speed estimation, the problem now is a parameter estimation as opposed to a model identification problem. To undertake this, the parameter estimation problem is formulated using the MAP estimator which requires to define a prior distribution for the delay parameters, and a joint distribution for the observations. Defining the parameter vector $\boldsymbol{T} = \begin{bmatrix} T_{\lambda_1} & \cdots & T_{\lambda_{|\Lambda|}} \end{bmatrix}^T$, the prior distribution for $\boldsymbol{T}$ can be given as

$$f_T(\boldsymbol{T}) = (2\pi)^{-\frac{|\Lambda|}{2}} \det(\boldsymbol{\Sigma}_T)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(\boldsymbol{T} - \overline{\boldsymbol{T}})^T \boldsymbol{\Sigma}_T^{-1} (\boldsymbol{T} - \overline{\boldsymbol{T}}) \right) \quad (6)$$

where $\boldsymbol{\Sigma}_T = \text{diag}\left(\sigma^2_{\lambda_1}, ..., \sigma^2_{\lambda_{|\Lambda|}} \right)$. As for the observation vector for all experiments and sensors we have $\tilde{\boldsymbol{\tau}} = \left[\tilde{\boldsymbol{\tau}}_1^T \ \cdots \ \tilde{\boldsymbol{\tau}}_E^T \right]^T$, where the individual experimental observation vector for the i-th experiment is $\tilde{\tau}_i = \left[\tilde{\tau}_{i1} \cdots \tilde{\tau}_{iE_i} \right]^T$. For convenience, the path-link incidence vector generator $\boldsymbol{P}(P)$ is defined as the binary column vector $\{\boldsymbol{P}(P)\}_{1,j}$ of size $|\Lambda| \times 1$ where it is 1 if $\lambda_j \in P$, (λ_j is the link number) and it hold 0 otherwise. Defining $\boldsymbol{T} = \left[T_{\lambda_1}, \cdots, T_{\lambda_{|\Lambda|}} \right]^{\mathrm{T}}$, it holds that the path delay can be expressed as the vector multiplication $\tau(P) = \boldsymbol{P}(P)^{\mathrm{T}} \boldsymbol{T}$. For experiment *i*, the path vectors from $\boldsymbol{P}_i$ can be arranged to make the path-link incidence matrix $\mathbf{E}(\boldsymbol{P}_i) = \left(\boldsymbol{P}(P_{i1}), \ldots, \boldsymbol{P}(P_{iE_i}) \right)^{\mathrm{T}}$, which is a binary $(E_i \times |\Lambda|)$ matrix. Collating the sub-matrices for all experiments yields the total path-link incidence matrix $\boldsymbol{E}_{\boldsymbol{P}} = \left(\boldsymbol{E}(\boldsymbol{P}_1), \ldots, \boldsymbol{E}(\boldsymbol{P}_E) \right)^T$.

Given the foregoing notation, it can be shown that the observation vector has the assumed form $\breve{\tilde{\boldsymbol{\tau}}} = \boldsymbol{E}_{\boldsymbol{P}} \breve{\boldsymbol{T}} + \breve{\boldsymbol{\varepsilon}}$ where $\breve{\boldsymbol{\varepsilon}}$ is a vectorisation of the observation errors. Consequently, the joint distribution for $\breve{\tilde{\boldsymbol{\tau}}}$ conditioned on $\breve{\boldsymbol{T}} = \boldsymbol{T}$ and $\boldsymbol{P} = \hat{\boldsymbol{P}}$ is given as

$$f_\tau\left(\tilde{\tau} \middle| \boldsymbol{T}, \hat{\boldsymbol{P}} \right) = (2\pi)^{-\frac{1}{2}\sum_{i=1}^{E} E_i} \det(\boldsymbol{\Sigma}_\tau)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\left(\tilde{\boldsymbol{\tau}} - \boldsymbol{E}_{\hat{\boldsymbol{P}}} \boldsymbol{T} \right)^T \boldsymbol{\Sigma}_\tau^{-1} \left(\tilde{\boldsymbol{\tau}} - \boldsymbol{E}_{\hat{\boldsymbol{P}}} \boldsymbol{T} \right) \right) \quad (7)$$

where $\boldsymbol{\Sigma}_\tau = \text{Cov}\left[\breve{\boldsymbol{\varepsilon}} \right] = \text{diag}\left(\sigma^2_{\tau_{11}}, ..., \sigma^2_{\tau_{EE_E}} \right)$ is the observation covariance matrix.

Given these definitions, the MAP estimator can now be defined as

$$\hat{\boldsymbol{T}}_{\text{MAP}} = \arg\max_{\boldsymbol{T}} f_\tau\left(\tilde{\tau} \middle| \boldsymbol{T}, \hat{\boldsymbol{P}} \right) f_T(\boldsymbol{T}) \quad (8)$$

Taking the logarithm of the term maximised in Eq. (8), and noting that the resulting log-likelihood function is strictly increasing in T, so that maximising the original term is equivalent to minimising its negative. The minimiser is then obtained by setting the partial derivative $\partial/\partial T$ to zero. Finally, after standard matrix-calculus methods to the following estimator is obtained,

$$\widehat{\boldsymbol{T}}_{\mathbf{MAP}} = \left(\boldsymbol{E}_{\widehat{\boldsymbol{P}}}^{\boldsymbol{T}}\boldsymbol{\Sigma}_{\boldsymbol{\tau}}^{-\mathbf{1}}\boldsymbol{E}_{\widehat{\boldsymbol{P}}} + \boldsymbol{\Sigma}_{\boldsymbol{T}}^{-\mathbf{1}}\right)^{-\mathbf{1}}\left(\boldsymbol{E}_{\widehat{\boldsymbol{P}}}^{\boldsymbol{T}}\boldsymbol{\Sigma}_{\boldsymbol{\tau}}^{-\mathbf{1}}\tilde{\boldsymbol{\tau}} + \boldsymbol{\Sigma}_{\boldsymbol{T}}^{-\mathbf{1}}\bar{\boldsymbol{T}}\right) \tag{9}$$

which represents a weighted average of the observed delays $\tilde{\boldsymbol{\tau}}$ and the expected delays $\bar{\boldsymbol{T}}$. The observation weights are governed by the sensor data variances that are the diagonal elements of $\Sigma_{\boldsymbol{\tau}}$ considering that off-diagonals (covariances) are usually zero. The uncertainty of the calculated travel times $\Sigma_{\boldsymbol{T}}$ (which are proportional to the wave speed standard deviations) will serve as the weights of the delay times.

For Eq. (9) to have a unique solution and not be underdetermined, the determinant of matrix $\left(\boldsymbol{E}_{\widehat{\boldsymbol{P}}}^{\boldsymbol{T}}\Sigma_{\tau}^{-1}\boldsymbol{E}_{\widehat{\boldsymbol{P}}} + \Sigma_{T}^{-1}\right)$ must be non-zero. For a deterministic modelling, i.e., $\Sigma_T = \Sigma_\tau = 0$, this condition indicates that the number of measured travel times must be equal to or greater than the number of unknown delay times (wave speeds).

Resistance estimation

Defining the parameter vector $\boldsymbol{\alpha} = \left[\alpha_{\lambda_1} \cdots \alpha_{\lambda_{|\Lambda|}}\right]^T$, the prior distribution for $\boldsymbol{\alpha}$ can be given by $\breve{\boldsymbol{\alpha}} \sim f_\alpha = N\left(\bar{\boldsymbol{\alpha}}, \boldsymbol{\Sigma}_\alpha\right)$ where $\boldsymbol{\Sigma}_\alpha = \mathrm{diag}\left(\varsigma_{\lambda_1}^2, \ldots, \varsigma_{\lambda_{|\Lambda|}}^2\right)$. For the estimation of the resistances the observations that are used are the log ratios of the generated and measured waves, defined as $\tilde{\gamma} = [\tilde{\gamma}_1^T \cdots \tilde{\gamma}_E^T]^T$ where the individual experimental observation vectors are $\tilde{\boldsymbol{\gamma}}_i = \left[\tilde{\gamma}_{i1} \cdots \tilde{\gamma}_{iE_i}\right]^T$. Defining the natural logarithm of the junction-based attenuation along path P as $S(P) = \sum_{k=1}^{|P|-1} \log\left(S_{\lambda_k \lambda_{k+1}}\right)$, the vector of junction losses for experiment i can be denoted as

$\boldsymbol{s}(\boldsymbol{P}_i) = \left(S(P_{i1}), \ldots, S\left(P_{iE_i}\right)\right)^T$, and collating the sub-vectors for all experiments yields $\boldsymbol{s_P} = \left(\boldsymbol{s}(\boldsymbol{P}_1), \ldots, \boldsymbol{s}(\boldsymbol{P}_E)\right)^T$. With

Using the foregoing notation for the attenuations and considering their theoretical expected values, it can be shown that the observation vector has the assumed form $\breve{\tilde{\gamma}} = \boldsymbol{s}_{\boldsymbol{p}} - \boldsymbol{E}_{\boldsymbol{P}}\breve{\boldsymbol{\alpha}} + \breve{\tilde{\xi}}$ where $\breve{\tilde{\xi}} \sim N\left(0, \Sigma_\gamma\right)$ is a vectorisation of the observation errors, where $\Sigma_\gamma = \mathrm{Cov}\left[\breve{\tilde{\xi}}\right] = \mathrm{diag}\left(\sigma^2_{\gamma_{11}}, \ldots, \sigma^2_{\gamma_{EE_E}}\right)$. Consequently, the joint distribution for $\breve{\tilde{\gamma}}$ conditioned on $\breve{\boldsymbol{T}} = \hat{\boldsymbol{T}}$ and $\boldsymbol{P} = \hat{\boldsymbol{P}}$ is given as

$$f_\gamma\left(\tilde{\gamma}\middle|\boldsymbol{\alpha}, \hat{\boldsymbol{T}}, \hat{\boldsymbol{P}}\right) = \left(2\pi\right)^{-\frac{1}{2}\sum_{i=1}^{E} E_i} \det\left(\Sigma_\gamma\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\left(\tilde{\gamma} - \boldsymbol{s}_{\hat{\boldsymbol{P}}} + \boldsymbol{E}_{\hat{\boldsymbol{P}}}\boldsymbol{\alpha}\right)^T \Sigma_\gamma^{-1}\left(\tilde{\gamma} - \boldsymbol{s}_{\hat{\boldsymbol{P}}} + \boldsymbol{E}_{\hat{\boldsymbol{P}}}\boldsymbol{\alpha}\right)\right) \tag{10}$$

Given these definitions, the MAP estimator can now be defined as

$$\hat{\boldsymbol{\alpha}}_{\mathrm{MAP}} = \arg\max_{\boldsymbol{\alpha}} f_\gamma\left(\tilde{\gamma}\middle|\boldsymbol{\alpha}, \hat{\boldsymbol{T}}, \hat{\boldsymbol{P}}\right) f_\alpha\left(\boldsymbol{\alpha}\right) \tag{11}$$

where using standard techniques, as for Eq. (9) leads to

$$\hat{\boldsymbol{\alpha}}_{\mathrm{MAP}} = \left(\boldsymbol{E}_{\hat{\boldsymbol{P}}}^T \Sigma_\gamma^{-1} \boldsymbol{E}_{\hat{\boldsymbol{P}}} + \Sigma_\alpha^{-1}\right)^{-1}\left(\boldsymbol{E}_{\hat{\boldsymbol{P}}}^T \Sigma_\gamma^{-1}\left(\boldsymbol{s}_{\hat{\boldsymbol{P}}} - \tilde{\gamma}\right) + \Sigma_\alpha^{-1}\bar{\boldsymbol{\alpha}}\right) \tag{12}$$

Like the previous step, this solution can be expressed as consisting of two components: one derived from the observations $(\boldsymbol{s}_{\hat{\boldsymbol{P}}} - \tilde{\boldsymbol{\gamma}})$ and the other based on the assumed mean $\bar{\boldsymbol{\alpha}}$, which are weighted by their variances.

In practice, the estimated wave speeds (delay times) and resistant parameters can be improved by repeating experiments and their associated precision. Hence, the estimated time delay and attenuation factor formulation modified based on the number of experiments N_e can be expressed as follows:

$$\bar{\bar{\boldsymbol{T}}} = \left(\boldsymbol{E}_{\hat{\boldsymbol{P}}}^{T}\Sigma_{\boldsymbol{\tau}}^{-1}\boldsymbol{E}_{\hat{\boldsymbol{P}}} + \Sigma_{\boldsymbol{T}}^{-1}\right)^{-1}\left(\boldsymbol{E}_{\hat{\boldsymbol{P}}}^{T}\Sigma_{\boldsymbol{\tau}}^{-1}\frac{\Sigma_{k}^{N_e}\tilde{\boldsymbol{\tau}}_{k}}{N_e} + \Sigma_{\boldsymbol{T}}^{-1}\bar{\boldsymbol{T}}\right) \tag{13}$$

$$\bar{\hat{\boldsymbol{\alpha}}} = \left(\boldsymbol{E}_{\hat{\boldsymbol{P}}}^{T}\Sigma_{\boldsymbol{\gamma}}^{-1}\boldsymbol{E}_{\hat{\boldsymbol{P}}} + \boldsymbol{\Sigma}_{\boldsymbol{\alpha}}^{-1}\right)^{-1}\left(\boldsymbol{E}_{\hat{\boldsymbol{P}}}^{T}\Sigma_{\boldsymbol{\gamma}}^{-1}\left(\boldsymbol{s}_{\hat{\boldsymbol{P}}} - \frac{\Sigma_{k}^{N_e}\tilde{\boldsymbol{\gamma}}_{k}}{N_e}\right) + \Sigma_{\boldsymbol{\alpha}}^{-1}\bar{\boldsymbol{\alpha}}\right) \tag{14}$$

which intuitively suggests that the mean values (expectations) of the observations are used to minimize their standard deviations.

NUMERICAL EXPERIMENTS

The path-based parameter estimation methodology is applied to two numerical case studies. The first case study, adopted from Zecchin et al. (2022), is used to provide insight into this proposed method. This includes a detailed construction of various terms within the proposed method. In this case study, uncertainty from measurement errors, assumption of prior distributions as well as other assumptions are studied and discussed. The second case study is used to demonstrate the capability of the proposed method to a realistic-sized network.

Case study 1

The network schematic of the first case study is depicted in Figure 2a, which consists of 15 pipes, 11 nodes and a reservoir. Two transient generation sources for the calibration method are assumed to take place at node 10 and 11, while three sensor locations are at nodes 3, 5 and 6 respectively. Pipe characteristics, including pipe length, diameter, initial velocity, the true wave speed and the true friction factor in each pipe (to be estimated) are shown in Table 1, which are used to compute the hypothetical measurements, i.e., observed arrival times.

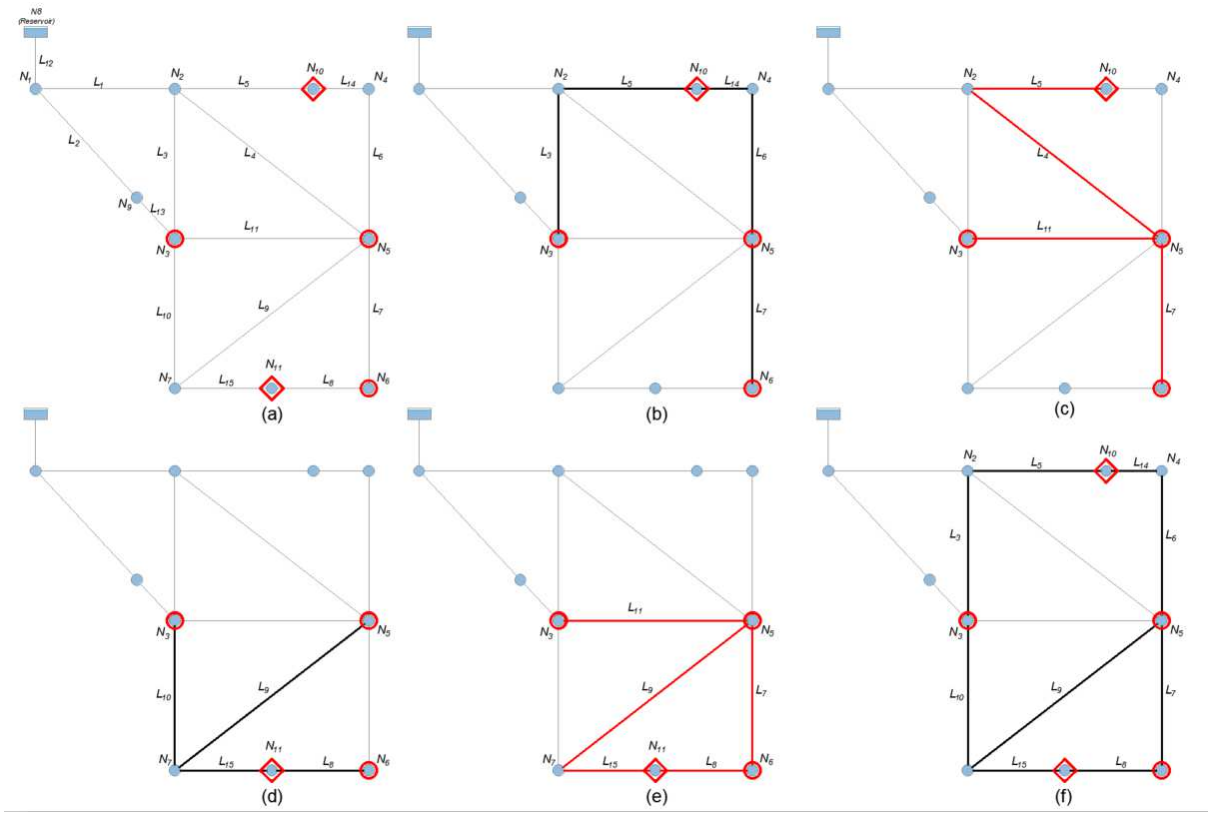


Figure 2: (a) Network schematic of the 1st case study indicating transient sources and sensor sites; (b) and (c) 1st and 2nd fastest paths from transient source N_{10} to sensor nodes N_3, N_5 and N_6; (d) and (e) 1st and 2nd fastest paths from transient source N_{11} to sensor nodes N_3, N_5 and N_6; (f) the most likely fastest paths from two transient sources to the sensor nodes.

Input for the calibration method

In practice, the input data for the calibration process are obtained from transient field tests, where transient sources are injected into specific locations of the network while their traces are measured at locations of interest. For this numerical case study, the input data are synthetically generated by assuming that the true states of the system are known, as shown in Table 1. The time delays and attenuation factors from the sources and the measurements are computed based on this assumption. To introduce measurement errors, random white noises (i.e. $N(0, \sigma_\tau = 0.005)$ and $N(0, \sigma_\gamma = 0.001)$ for time delay and path attenuation factor, respectively) are added to the synthetic measurement data.

Table 1: Pipe characteristics and true states of case study 1 network

Pipe ID	Diameter $D(m)$	Length $L\ (m)$	Init. Velocity $V_0\ (m/s)$	Wave speed $a_{true}\ (m/s)$	Friction factor f_{true}
L1	0.1	1350	0.59	1209	0.025

L2	0.1	1500	0.49	1216	0.024
L3	0.1	1500	0.01	1206	0.04
L4	0.1	1650	0.15	1164	0.031
L5	0.1	600	0.12	1210	0.034
L6	0.1	1730	0.12	1235	0.032
L7	0.1	1355	0.19	1180	0.03
L8	0.1	400	0.07	1155	0.037
L9	0.1	1210	0.19	1228	0.028
L10	0.1	1355	0.26	1198	0.028
L11	0.1	800	0.23	1243	0.03
L12	0.1	500	1.08	1206	0.024
L13	0.1	500	0.49	1162	0.024
L14	0.1	200	0.12	1203	0.034
L15	0.1	400	0.07	1207	0.037

For this case study, the prior information of the time delay (i.e. mean (μ_{λ_i}) and standard deviation (σ_{λ_i})) for each link in the network is computed based on an assumption that the wave speeds are varied ± 50 around their means of 1200 (m/s) (i.e. a maximum relative error of 4.2% is considered for the wave speed). Similarly, the attenuation factor prior is computed based on this wave speed variation as well as the variation of the friction factors, which is ± 0.02 around the mean of 0.04 (i.e. a maximum relative error of 50% is considered for the friction factor). Note that the actual friction factors are generated within the range 0.03 ± 0.01, which is designed to show distinctive benefits of the proposed method even if the prior knowledge is far from the truth. Table 2 summarises the prior distribution characteristics of the time delays and attenuation factors given the aforementioned assumptions.

Table 2: Prior probability of time delay and attenuation factor given variations of wave speeds and friction factors

Pipe ID	Time delay		Attenuation factor	
	Mean	Standard deviation	Mean	Standard deviation
L1	1.1270	0.02396	0.1358	0.03675
L2	1.2522	0.02662	0.1253	0.03391
L3	1.2522	0.02662	0.0026	0.00069
L4	1.3774	0.02928	0.0422	0.01142
L5	0.5009	0.01065	0.0123	0.00332
L6	1.4442	0.03070	0.0354	0.00958
L7	1.1311	0.02405	0.0439	0.01188
L8	0.3339	0.00710	0.0048	0.00129
L9	1.0101	0.02147	0.0392	0.01061
L10	1.1311	0.02405	0.0600	0.01626

L11	0.6678	0.01420	0.0314	0.00849
L12	0.4174	0.00887	0.0920	0.02492
L13	0.4174	0.00887	0.0418	0.01130
L14	0.1670	0.00355	0.0041	0.00111
L15	0.3339	0.00710	0.0048	0.00129

Identification of K fastest paths

The Yen's K-shortest paths finding algorithm (Yen 1971) is adopted to find up to K number of shortest paths (herein, refer to as the fastest paths) within the proposed methodology. The Yen's algorithm firstly finds the fastest path between the transient source and the sensor using a shortest path finding algorithm, e.g. the Djikstra algorithm (1959), and secondly identify the subsequent fastest paths (2^{nd}, 3^{rd},…,k^{th}) by iteratively searching through sub-graph networks that are not contain (1…k^{th-1}) previously defined shortest paths. Details of the method can be found in Yen, 1971.

Figure 2b, 2c, 2d and 2e show the results of the 2-fastest paths from the transient generation source N_{10} and N_{11} to sensor nodes N_3, N_5 and N_6. For example, the 1^{st} and 2^{nd} fastest paths from N_{10} to N_3 are L_5-L_3 and L_5-L_4-L_{11}, respectively, ranked in ascending order of the time delay. Table 3 lists all possible fastest paths from two transient sources to these three sensors in the form of a k-paths-link incident matrix (i.e. a matrix that combines all potential path-link incident matrices together)

Table 3: Possible fastest paths represented in terms of a potential path-link incident matrix and the corresponding likelihood estimate scores.

Path	Rank	L_1	L_2	L_3	L_4	L_5	L_6	L_7	L_8	L_9	L_{10}	L_{11}	L_{12}	L_{13}	L_{14}	L_{15}	LE Score
N_{10} to N_3	**1st**	**0**	**0**	**1**	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**12.89**
	2nd	0	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0.00
N_{10} to N_5	**1st**	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**5.60**
	2nd	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0.00
N_{10} to N_6	**1st**	**0**	**0**	**0**	**0**	**0**	**1**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**8.77**
	2nd	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0.00
N_{11} to N_3	**1st**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**0**	**1**	**15.89**
	2nd	0	0	0	0	0	0	0	0	1	0	1	0	0	0	1	0.00
N_{11} to N_5	**1st**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**1**	**9.91**
	2nd	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0.00
N_{11} to N_6	**1st**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**12.93**

	2nd	0	0	0	0	0	0	1	0	1	0	0	0	0	0	1	0.00

The likelihood score associate with each potential path for estimating the true hypothesis, see Eq. (5), is also shown in table 2. It is seen that the 1st ranking fastest paths have highest score, hence, the null hypothesis cannot be rejected, and the fastest path is the actual path.

Wave speed and friction factor estimation

Once the actual paths are identified, the path-link incident matrix (i.e., the term $E_{\hat{p}}$ in Eq. (10) and Eq. (12)) is defined, which is the collated binary values highlighted in bold in Table 2. The joint distributions for $\tilde{\tilde{\tau}}$ and $\tilde{\tilde{\gamma}}$ given by Eq. (7) and Eq. (10) are also determined, which enable Eq. (9) and Eq. (12) to be solved. As a result, the parameter of interest, in this case, the wave speeds and friction factors along the shortest paths can be computed as shown in Figures 3 and 4.

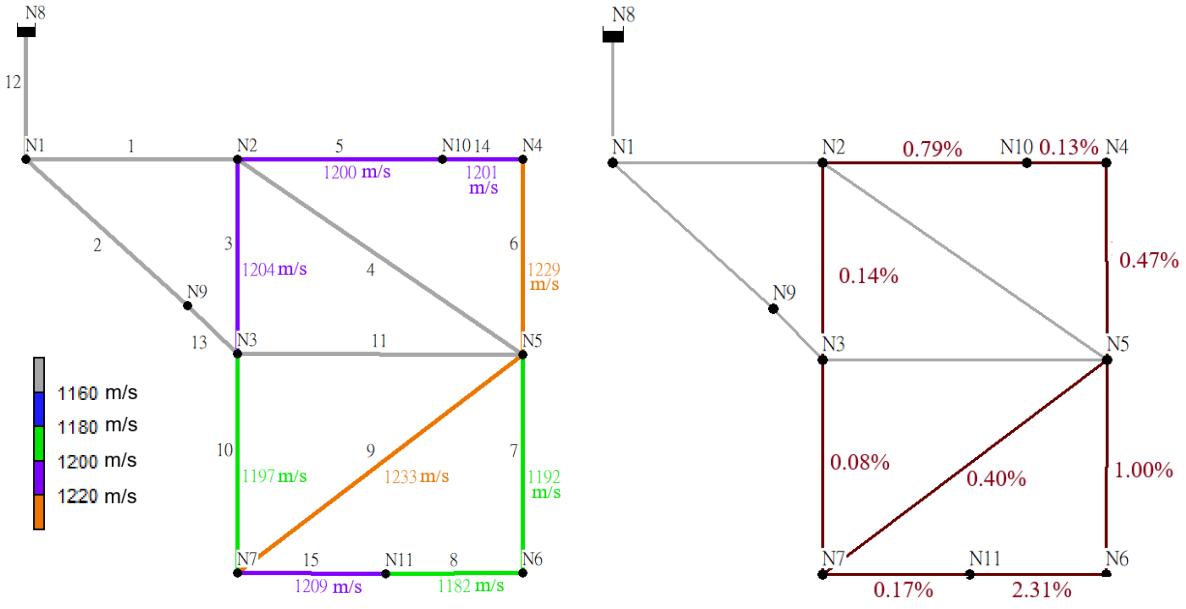


Figure 3: Estimated wave speeds (left) and their error in comparison with the true values (right); see Table 4 in Appendix D for further analysis details. Note that the grey-coloured pipes are undetectable given the incorporated transient source and sensor arrangement.

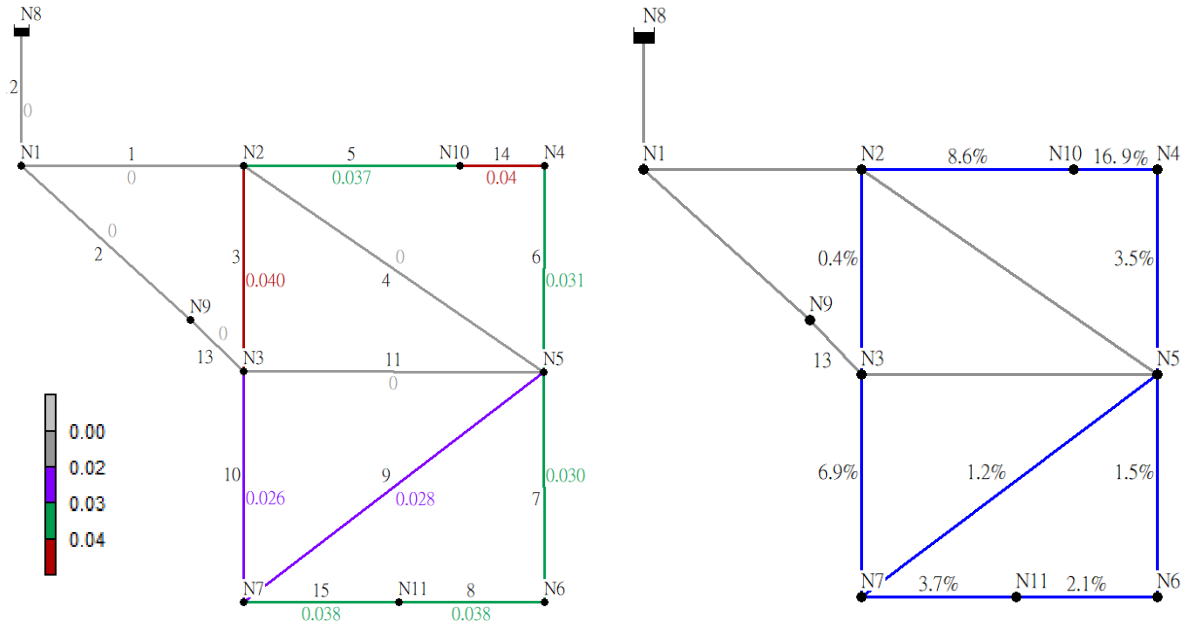


Figure 4: Estimated friction factors (left) and their error in comparison with the true values (right); see Table 5 in Appendix D for further analysis details. Note that the grey-coloured pipes are undetectable given the incorporated transient source and sensor arrangement.

It is seen from Figure 3 that the results from the proposed calibration strategy yield a relatively good match between estimated and actual wave speeds. The maximum relative error is 2.31% for link L8 (between nodes 6 and 11) while the minimum relative error is 0.13% for link L14 (see Table 4 in Appendix D for further details). Specifically, Table 4 also provides the relative error between the prior and the true values to illustrate the improvement of the posterior comparing to the prior. The highest improvement is observed at link L6, in which the relative error reduces from 2.83% to 0.47%, while the smallest improvement is at link L14 (between nodes 4 and 10) as the relative error increase from 0.25% to 0.13%.

The calibrated results for friction factor are shown in Figure 4, which exhibit a significant improvement given our assumed non-informative prior (see Table 5 in Appendix D for further details). The highest improvement is observed at link L9 (between nodes 5 and 7), in which the relative error reduces from 42.86% to 1.25%, while the smallest improvement is at link L14, again, as the relative error increase from 17.65% to 16.95%. It should be noted that for link 3 (between nodes 2 and 3), the estimated friction factor seems to be worse than the prior. However, the estimated result is only 0.001 different to the prior, which results in 0.39% of the

relative error. This is trivial in this case as the measurement error is considered and it is unlikely that the prior equals to the actual value.

In terms of computation time, it takes only 0.58 s to complete the calibration process using an Intel® Core™i7-3.1GHz (16.0 GB) computer.

Improving estimated results by the number of experiments

As discussed in Eqs. (13) and (14), the accuracy of the estimated results can be enhanced by repeating experiments. Figure 5 depicts the box whisker of the estimated results over 20 runs. It is seen that single runs of the proposed method may result in outliers (e.g. estimated wave speed at links L3 and L6 in Figure 5a, or friction factor at links L9 and L10). However, the mean posteriors are much closer to the actual values. In addition, the uncertainty range of estimated parameter can also be quantified, in which the box whiskers cover almost all true values, except for estimated wave speeds at links L5 and L14 and for estimated friction factors at links L14 and L15. This issue will be discussed in the following subsection.

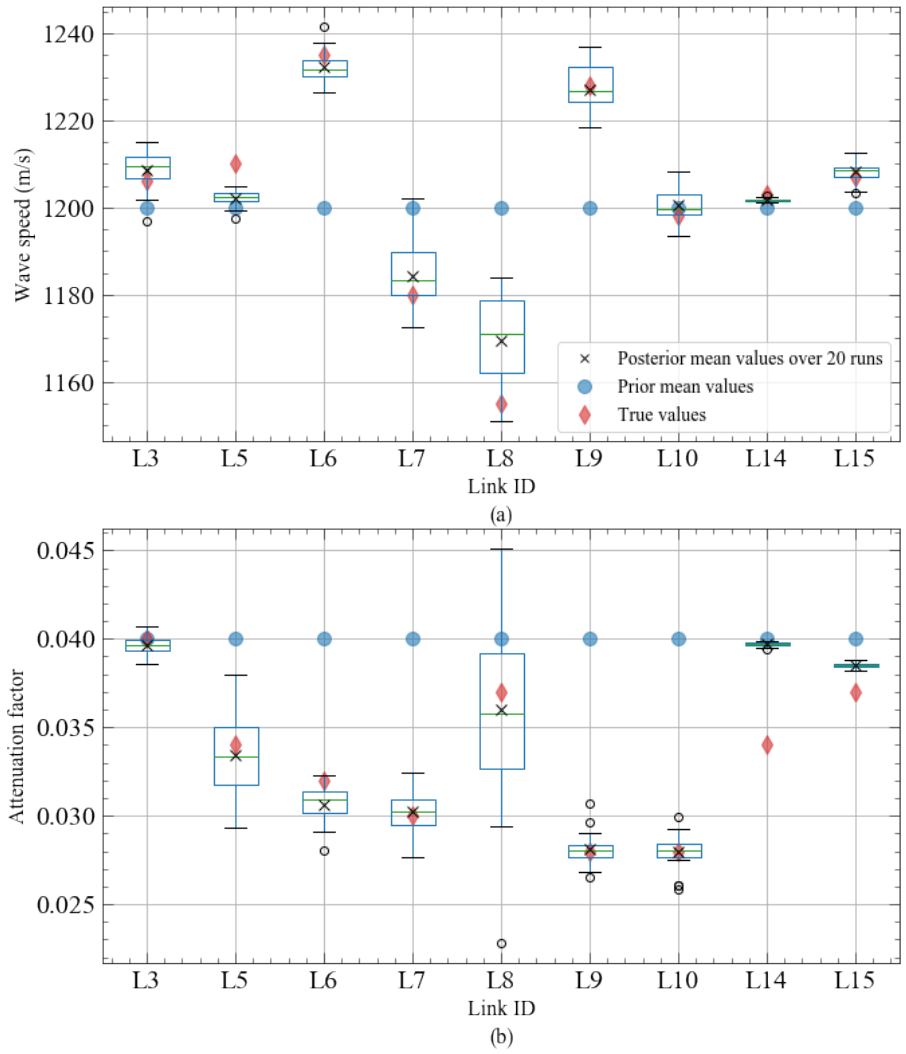


Figure 5: Estimated results over multiple experiments: (a) wave speed and (b) attenuation factor

Effects of short pipes to the estimated results

It is seen in Figure 5 that the actual values of the wave speed in links L5 and L14 as well as of the friction factor in links L14 and L15 stay outside the box whiskers of the estimated results. These inaccurate estimations are the consequence of having short pipes in the fastest path. Let us closely examine the fastest path between the transient source N10 and sensor N3, where the wave speeds of links L3 and L5 have been estimated. The associated lengths for these links are 1500 (m) and 600 (m), respectively. The time delay distribution of link L3, as a result, dominates the distribution of link L5 (the wave speed mean and standard deviation of link L5 is approximately 2.5 times larger than that of link L3) as shown in Table 2. As the path delay is the summation of the link delays, the effect of link L3 is also approximately 2.5 times

larger than link L5. A small perturbation of the wave speed in link 3 can cause a large variation in the time delay of the path and vice versa, a large perturbation of the wave speed in link 5 may only cause a small change in the path time delay. The proposed methods is, therefore, focuses on estimating the wave speed of link L3, while keeping the wave speed of link 5 close to the prior.

For the estimation of the friction factor (which is based on estimating the attenuation factor), this issue is even more problematic as the estimation of the attenuation factor is also dependent on the estimation of the time delay in each link along the fastest path (see equation 10).

Effects of measurement error and prior information of wave speeds to selection of the fastest paths.

The prior information of the to-be-calibrated parameters and measurement errors from measurement devices not only directly affect the estimated results (which can easily be found elsewhere in the literature, e.g. Shang et al., 2006, Do et al., 2017b), but also affect the selection of the fastest paths. In order to evaluate these influences, different levels of uncertainty are assigned to either prior distributions or measurement errors to assess the results.

Specifically, two tests are implemented in this study. For the first test, the prior of wave speeds are assume to vary ±10, ±20 ±50, ±100, and ±200 around their means (i.e. the relative errors between the prior and the actual wave speed are varied from 0.83% to 16.67%), while no measurement error is considered. Note that for this test, the mean wave speed prior (i.e. 1200 m/s) is fixed, the actual wave speeds of the pipes are randomly generated within the uncertainty range. For the second test, the prior of wave speeds is assumed a fixed boundary of ±50 around their means, while the standard deviation of the measurement error is set at 0.0025, 0.005, 0.01, 0.025, and 0.05. Figure 6 shows the likelihood scores of the 2-fastest paths between the transient source N_{11} and sensor N_5 as an example. Figure 6a shows the

influence of wave speed prior on the likelihood while Figure 6b shows the influence of the measurement error. It is seen that only if the prior boundary reaches ±200, the likelihood of the second-fastest path L_7 to L_8 is higher than the first-fastest path (i.e. the actual path L_9 to L_{15}). Similarly, if the standard deviation of the measurement error reaches 0.05, the second-fastest path will be selected as the fastest path. As a result, link L9, in which the actual transient wave travels through, was not selected for estimation in both cases.

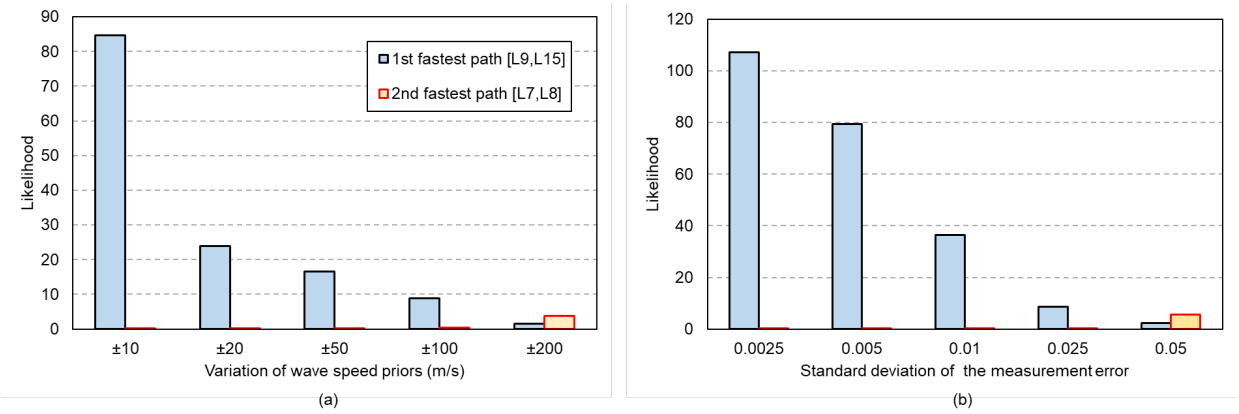


Figure 6: Effects of prior information and measurement error on selection of the most likely path from transient source N_{11} to sensor node N_5. (a): Influence of the wave speed prior on fastest path likelihood; (b): Influence of the measurement error on fastest path likelihood

Case study 2

The second numerical case study aims at evaluating the proposed methodology for a realistic network. The Adelaide central business district (CBD) gravity-fed WDS with more than 125 km (Stephen et al 2018), therefore, is adopted for this case study. The section of interest is highlighted in Figure 7, which consists of 178 nodes and 252 pipes with the diameter ranging from 80 to 750 (mm). The parameter of interest in this case study is the wave speeds of the pipes, therefore the estimation results of the attenuation factors will not be discussed here.

This network is predominantly cast-iron pipes with the mean of the wave speed is 1250 (m/s). Only 6 pipes are asbestos cement (AC), polyethylene (PE), and PVC pipes. The mean wave speed of these pipes are 800, 600 and 400 (m/s), respectively. It is assumed that 4 transient sources and 15 sensor locations (randomly selected) are placed across the network as seen in Figure 7, which results in a total of 60 time delay observations.



Figure 7: Adelaide CBD water distribution system and its highlighted partial network of interest for wave speed estimation.

As the pipe lengths in this network section are relatively short (ranging from 10 m to 260 m), the variation of the wave speed in this case study is assumed to be within ± 20 of their means, while the measurement error in this case is assumed $N(0, \sigma_\tau = 0.0004)$. These assumptions are made considering prior empirical assessments of the network. to ensure the fastest paths selected by the proposed method are the actual fastest paths in the network.

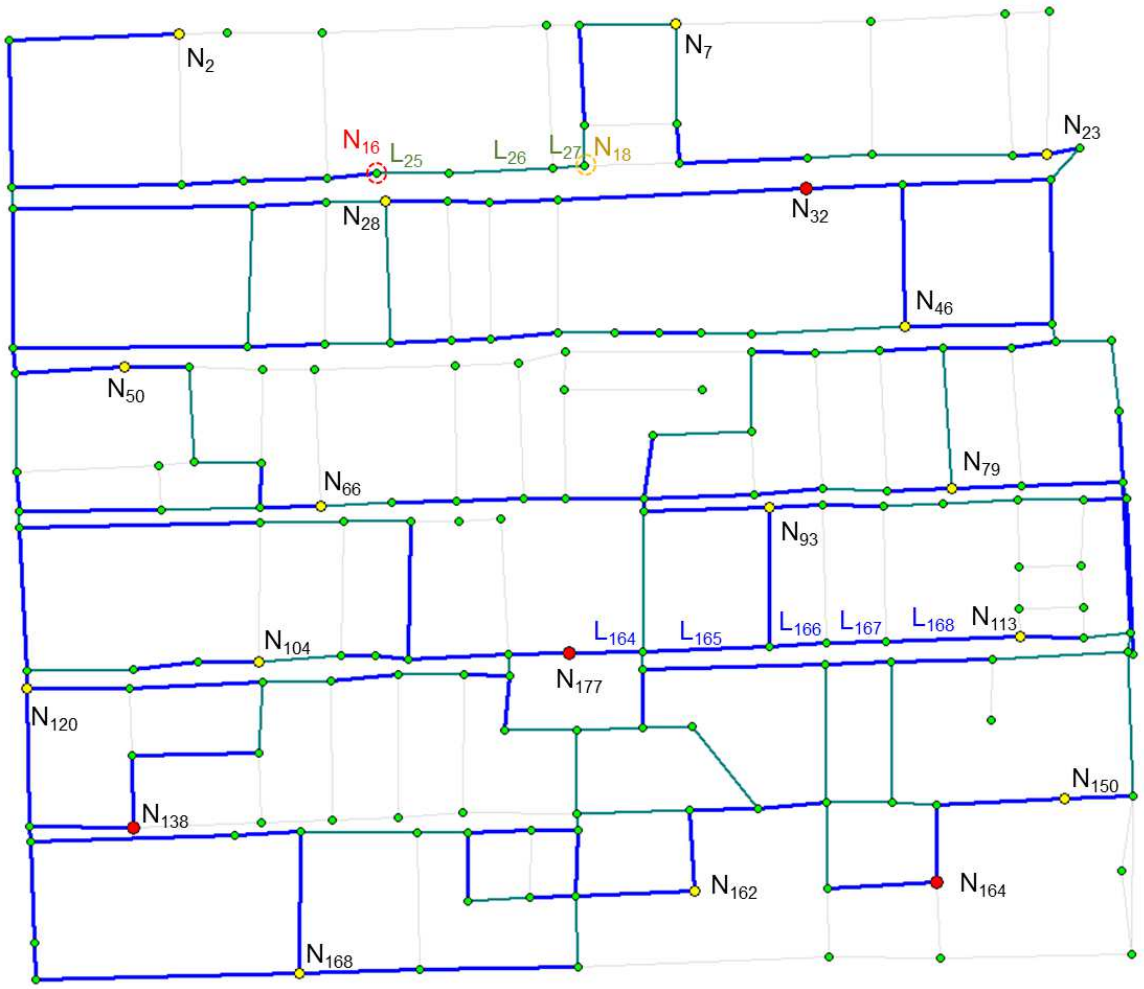


Figure 8: Transient source N16 and sensor N18 (highlighted by a dashed circle) are added to improve the calibration of wave speed along these sections.

Figure 8 plots the results from the proposed method. The pipes in light grey indicate that they do not belong to any fastest paths. Therefore, the wave speeds associated with these pipes could not be estimated. Out of 252 pipes, 174 pipes (highlighted in blue and dark green colours) belong to one of the fastest paths between the transient sources to the sensors, in which their wave speeds are estimated by the proposed model. It is important to note that with 60 observations (derived from 4 transient sources and 15 sensors), the attempt to estimate 174 unknown wave speed parameters constitutes an overdetermined problem, leading to non-unique estimates. In other words, the determinant of the coefficient matrix $\left(\boldsymbol{E}_{\hat{\boldsymbol{P}}}^{\boldsymbol{T}}\Sigma_{\tau}^{-1}\boldsymbol{E}_{\hat{\boldsymbol{P}}}+\Sigma_{T}^{-1}\right)$ in Eq. (9) is small, which suggests only marginal improvements in some pipes.

The results from the proposed method show that the estimated wave speeds in 109 pipes (highlighted in blue colour) significantly improve from their initial values. Estimate results of 65 pipes (highlighted in dark green colour) do not show improvement. It is seen that the improvement is often associated with pipe sections where the number of pipes within the fastest paths is equal or less than 5 pipes, for example from transient source N32 to sensor N28 (i.e. 4 pipes), from N177 to N113 (i.e. 5 pipes), or N138 to N168 (i.e. 4 pipes). In addition, the pipes that are closer to transient sources are more likely to have better estimation. This is because these pipes are often associated with multiple fastest-paths from the source to the sensors. The accuracy of the estimation is reduced if the paths consist of a relatively large number of pipes. Figure 9a plots an example of a section of 5 pipes (from L164 to L168) in which their wave speeds are improved by the proposed method, while Figure 9b shows an example of a section consist of three pipes (L25, L26 and L27) where the wave speed estimates do not improve. Improvement of the calibration can be carried by adding more sensors or by having additional transient tests around these places. For instance, if an extra transient source is added at node N16 and a sensor is placed at node N18 (marked by red and yellow circles in Figure 8), the estimated results of these three pipes are significantly improved as shown Figure 9c.

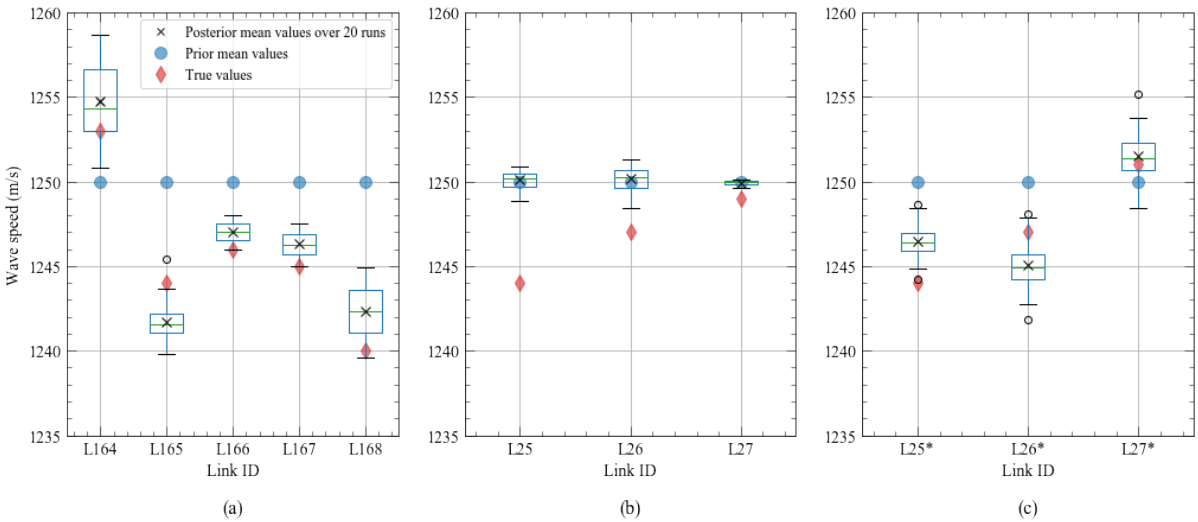


Figure 9: Example of calibrated wave speeds and improvement by adding additional transient source and sensor.

CONCLUSIONS

This paper developed a wave propagation path–based method for parameter estimation in pipeline networks. For the first time, it exploits only the coarse features of transient responses including wavefront delay and attenuation, thereby bypassing the computational and modelling challenges of traditional transient calibration techniques. The proposed framework incorporates probabilistic formulations through maximum likelihood and maximum a posteriori estimation, enabling it to handle both parametric uncertainty and measurement error.

Two case studies highlighted the effectiveness of the proposed method. The case study in the small water network illustrated the theoretical underpinnings and showed how the approach yields accurate estimates of wave speeds and friction factors despite imperfect prior information. The second case demonstrated the applicability of the method at the scale of a real urban water distribution system. The results confirmed that estimation reliability will be influenced by factors such as existence of short pipes, sparse sensors, and high uncertainty in priors. These limitations can be mitigated through multiple experiments and strategic placement of sources and sensors.

Overall, the wave propagation path strategy provides a computationally efficient, analytically tractable, and scalable approach for water network dynamic calibration. Its ability to produce reliable estimates without solving full hydraulic models makes it well suited for relatively large scale networks. Future work should explore integration with field experiments, extension to networks with mixed materials and operational conditions, and adaptive sensor placement strategies to maximise identifiability of key parameters.

CONFLICT OF INTEREST

The authors declare no conflicts of interest relevant to this study.

FUNDING

This work was funded by the Australian Research Council (DP230101513).

AUTHOR CONTRIBUTIONS

Aaron C. Zecchin: Conceptualization; Methodology; Resources; Writing – original draft; Writing – review & editing; Funding acquisition; Project administration. Nhu Do: Software; Data Curation; Formal analysis; Investigation; Visualisation; Writing – original draft; Writing – review & editing. Wei Zeng: Resources; Writing – review & editing; Funding acquisition. Alireza Keramat: Formal analysis; Investigation; Validation; Visualisation; Writing – review & editing. Martin F. Lambert: Resources; Writing – review & editing; Funding acquisition; Project administration.

DECLARATIONS

The authors declare that they have no competing interests.

AVAILABILITY STATEMENT

The data availability package includes the numerical data underlying all figures, the associated Python scripts used to generate these figures, and the network model files for the case-study networks, which are publicly available at https://doi.org/10.5281/zenodo.19562404.

REFERENCES

Alvisi, S., & Franchini, M. (2010). Pipe roughness calibration in water distribution systems using grey numbers. Journal of Hydroinformatics, 12(4), 424-445.

Asadzadeh, M., Tolson, B. A., & McKillop, R. (2010). A two stage optimization approach for calibrating water distribution systems. In Water Distribution Systems Analysis 2010 (pp. 1682-1694).

Blocher, C., Pecci, F., & Stoianov, I. (2021). Prior assumptions for leak localisation in water distribution networks with uncertainties. Water Resources Management, 35, 5105–5118.

Bohorquez, J., Alexander, B., Simpson, A. R., & Lambert, M. F. (2020). Leak detection and topology identification in pipelines using fluid transients and artificial neural networks. Journal of Water Resources Planning and Management, 146(6), 04020040.

Capponi, C., Meniconi, S., Lee, P. J., Brunone, B., & Cifrodelli, M. (2020). Time-domain analysis of laboratory experiments on the transient pressure damping in a leaky polymeric pipe. Water Resources Management, 34, 501–514.

Chen, R., Cui, S., Khoo, D. W. Y., & Khoo, B. C. (2022). Modelling and quantification of measurement uncertainty for leak localisation in a gas pipeline. Measurement: Sensors, 24, 100443.

Covas, D., Stoianov, I., Ramos, H., Graham, N., & Maksimovic, C. (2004). The dynamic effect of pipe-wall viscoelasticity in hydraulic transients. Part I—Experimental analysis and creep characterization. Journal of Hydraulic Research, 42(5), 517-532.

Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. Numerische Mathematik. 1 (1): 269–271. https://doi.org/10.1007/BF01386390.

Do, N. C., Simpson, A. R., Deuerlein, J. W., & Piller, O. (2018). Locating inadvertently partially closed valves in water distribution systems. Journal of Water Resources Planning and Management, 144(8), 04018039.

Do, N. C., Simpson, A., Deuerlein, J., & Piller, O. (2017, September). Online Demand Estimation of Geographical and Non-Geographical Distributed Demand Pattern in Water Distribution Networks. In CCWI 2017-Computing and Control for the Water Industry (pp. 8-p).

Do, N. C., Zecchin, A. C., Lambert, M. F., Zeng, W., & Gong, J. (2024). Real-Time Localization of Hydraulic Transient Events in Water Distribution Systems Using Paired Transient Wave Arrival Time Strategy. Journal of Water Resources Planning and Management, 150(11), 04024053.

Duan, H. F., & Keramat, A. (2022). Uncertainty quantification of transient-based leakage identification: A frequency domain framework. Water Resources Research, 58(12), e2022WR032512.

Duan, H.-F. (2015). Uncertainty analysis of transient flow modeling and transient-based leak detection in elastic water pipeline systems. Water Resources Management, 29, 5413–5427.

Ferrante, M., & Zecchin, A. (2023). Separation of pressure signals caused by waves traveling in opposite directions. Journal of Hydroinformatics, 25(5), 1747-1765.

Garthwaite, P., I. Jolliffe, and B. Jones. 2002. Statistical inference. 2nd ed. London: Oxford Science Publications.

Hutton, C. J., Kapelan, Z., Vamvakeridou-Lyroudia, L., & Savić, D. A. (2014). Dealing with uncertainty in water distribution system models: A framework for real-time modeling and data assimilation. Journal of Water Resources Planning and Management, 140(2), 169-183.

Imani, M., Zecchin, A., Zeng, W., & Lambert, M. F. (2025). Generalization and Analysis of Elastic Water Column Model for Hydraulic Transient Analysis of Water Distribution Systems. Journal of Water Resources Planning and Management, 151(10), 04025054.

Keramat, A., Duan, H. F., Pan, B., & Hou, Q. (2022). Gradient-based optimization for spectral-based multiple-leak identification. Mechanical Systems and Signal Processing, 171, 108840.

Khedr, A., Tolson, B., & Ziemann, S. (2015). Water distribution system calibration: Manual versus optimization-based approach. Procedia Engineering, 119, 725-733.

Kim, S. (2016). Impedance method for abnormality detection of a branched pipeline system. Water resources management, 30(3), 1101-1115.

Lansey, K. E., El-Shorbagy, W., Ahmed, I., Araujo, J., & Haan, C. T. (2001). Calibration assessment and data collection for water distribution networks. Journal of hydraulic engineering, 127(4), 270-279.

Lee, J. S., Zeng, W., Lambert, M. F., & Gong, J. (2024). Calibration of frequency-dependent wave speed and attenuation in water pipes using a dual-Sensor and paired-IRF approach. Journal of Water Resources Planning and Management, 150(11), 04024051.

Liggett, J. A., & Chen, L. C. (1994). Inverse transient analysis in pipe networks. Journal of hydraulic engineering, 120(8), 934-955.

Massari, C., Yeh, T. C., Ferrante, M., Brunone, B., & Meniconi, S. (2014). A stochastic tool for determining the presence of partial blockages in viscoelastic pipelines: first experimental results. Procedia Engineering, 70, 1112-1120.

Massari, C., Yeh, T.-C. J., Brunone, B., Ferrante, M., & Meniconi, S. (2013). Diagnosis of pipe systems by means of a stochastic successive linear estimator. Water Resources Management, 27, 4637–4654.

Meniconi, S., Brunone, B., Ferrante, M., & Massari, C. (2011). Small amplitude sharp pressure waves to diagnose pipe systems. Water Resources Management, 25, 79–96.

Meniconi, S., Capponi, C., Frisinghelli, M., & Brunone, B. (2021). Leak detection in a real transmission main through transient tests: Deeds and misdeeds. Water Resources Research, 57(3), e2020WR027838.

Ormsbee, L. E., & Wood, D. J. (1986). Explicit pipe network calibration. Journal of Water Resources Planning and Management, 112(2), 166-182.

Piller, O., & van Zyl, J. E. (2014). Modeling control valves in water distribution systems using a continuous state formulation. Journal of Hydraulic Engineering, 140(11), 04014052.

Preis, A., Whittle, A., & Ostfeld, A. (2009). Online hydraulic state prediction for water distribution systems. In World Environmental and Water Resources Congress 2009: Great Rivers (pp. 1-23).

Sarkamaryan, S., Haghighi, A., & Adib, A. (2018). Leakage detection and calibration of pipe networks by the inverse transient analysis modified by Gaussian functions for leakage simulation. Journal of Water Supply: Research and Technology—AQUA, 67(4), 404-413.

Savic, D. A., Kapelan, Z. S., & Jonkergouw, P. M. (2009). Quo vadis water distribution model calibration?. Urban Water Journal, 6(1), 3-22.

Shang, F., Uber, J. G., van Bloemen Waanders, B. G., Boccelli, D., & Janke, R. (2008, March). Real time water demand estimation in water distribution system. In Water Distribution Systems Analysis Symposium 2006 (pp. 1-14).

Vrachimis, S. G., Eliades, D. G., Taormina, R., Kapelan, Z., Ostfeld, A., Liu, S., ... & Polycarpou, M. M. (2022). Battle of the leakage detection and isolation methods. Journal of Water Resources Planning and Management, 148(12), 04022068.

Wang, X. J., Lambert, M. F., Simpson, A. R., Liggett, J. A., & V ı´ tkovský, J. P. (2002). Leak detection in pipelines using the damping of fluid transients. Journal of Hydraulic Engineering, 128(7), 697-711.

Wu, Z. Y. (2009). Unified parameter optimisation approach for leakage detection and extended-period simulation model calibration. Urban Water Journal, 6(1), 53-67.

Zecchin, A. C., Do, N., Gong, J., Leonard, M., Lambert, M. F., & Stephens, M. L. (2022). Optimal pipe network sensor layout design for hydraulic transient event detection and localization. Journal of Water Resources Planning and Management, 148(8), 04022041.

Zecchin, A. C., Lambert, M. F., & Simpson, A. R. (2012). Inverse Laplace transform for transient-state fluid line network simulation. Journal of engineering mechanics, 138(1), 101-115.

Zecchin, A. C., Simpson, A. R., Lambert, M. F., White, L. B., & Vítkovský, J. P. (2009). Transient modeling of arbitrary pipe networks by a Laplace-domain admittance matrix. Journal of Engineering Mechanics, 135(6), 538-547.

Zecchin, A. C., White, L. B., Lambert, M. F., & Simpson, A. R. (2013). Parameter identification of fluid line networks by frequency-domain maximum likelihood estimation. Mechanical Systems and Signal Processing, 37(1-2), 370-387.

Zecchin, A. C., White, L. B., Lambert, M. F., & Simpson, A. R. (2013). Parameter identification of fluid line networks by frequency-domain maximum likelihood estimation. Mechanical Systems and Signal Processing, 37(1-2), 370-387.

Zeng, W., Gong, J., Zecchin, A. C., Lambert, M. F., Simpson, A. R., & Cazzolato, B. S. (2018). Condition assessment of water pipelines using a modified layer-peeling method. Journal of Hydraulic Engineering, 144(12), 04018076.

Zeng, W., Zecchin, A. C., Gong, J., Lambert, M. F., Cazzolato, B. S., & Simpson, A. R. (2025). Anomaly Recognition in Water Pipe Systems Using a Characterization Framework and Persistent Hydraulic Transient Waves. Journal of Water Resources Planning and Management, 151(2), 06024002.

Zhang, C., Lambert, M. F., Gong, J., Zecchin, A. C., Simpson, A. R., & Stephens, M. L. (2020). Bayesian inverse transient analysis for pipeline condition assessment: parameter estimation and uncertainty quantification. Water Resources Management, 34, 2807-2820.

Zhang, C., Lambert, M. F., Gong, J., Zecchin, A. C., Simpson, A. R., & Stephens, M. L. (2020). Bayesian inverse transient analysis for pipeline condition assessment: Parameter estimation and uncertainty quantification. Water Resources Management, 34, 2807–2820.

Zhou, X., Xu, W., Xin, K., Yan, H., & Tao, T. (2018). Self-adaptive calibration of real-time demand and roughness of water distribution systems. Water Resources Research, 54(8), 5536-5550.

APPENDIX A: Transient model in pipe systems

The water hammer model govern the wave propagation in a pipe network via the following hyperbolic differential equations (Wylie93)

$$\begin{aligned} &\frac{\partial h}{\partial x} + L\frac{\partial q}{\partial t} + Rq + R_{NL}q^2 = 0 \\ &\frac{\partial q}{\partial x} + C\frac{\partial h}{\partial t} + C_R * \frac{\partial h}{\partial t} = 0 \end{aligned} \tag{15}$$

where $h = h(x,t)$ is the piezometric head (x and t are the space and temporal coordinates), $q = q(x,t)$ is the volumetric flow rate, $L = 1/gA$ is the hydraulic inertance (g is gravitational acceleration, A is cross-sectional area), $R = fv_0/gAD$ is the linear resistance (f is the Darcy-Weisbach friction factor, v_0 is the steady-state mean fluid velocity , and D is the diameter), $R_{NL} = f/2DA^2$ is the non-linear resistance, $C = gA/a^2$ is the hydraulic capacitance, and $C_R = C_R(t)$ is the retarded capacitance function (the $*$ denotes the convolution operator), typically associated with the visco-elastic effects[2] (e.g., Covas 2004). The wave speed is given by

$$a = \sqrt{\frac{K/\rho}{1 + \left(\frac{K}{E_0}\right)\left(\frac{D}{e}\right)c_1}} \tag{16}$$

where K is the fluid elastic bulk modulus, ρ is the fluid density, e is the pipe wall thickness, E_0 is Young's modulus of elasticity of the pipe material, and c_1 is the longitudinal pipeline restraint parameter.

[2] With reference to standard notation, $C_R * \partial h/\partial t = 2A\partial \varepsilon_R/\partial t$ where ε_R is the retarded circumferential strain.

APPENDIX B: Analytical derivation of the propagating wavefront

Within this Appendix, the form of the initial incident wavefront is determined. Considering a semi-infinite pipeline on domain $x \in [0,\infty)$ with boundary condition $h(0,t) = u(t)$, the Laplace-domain solution to Eq. (15) after omitting the nonlinear is given by (Zecchin et al. 2012)

$$\hat{h}(x,s) = e^{-x\sqrt{(Ls+R)(Cs+s\hat{C}_R(s))}}\hat{u}(s) \tag{17}$$

where $\hat{g}$ is the Laplace transform of g and s is the Laplace variable. As demonstrated by the method of characteristics solution, the pressure wave incident from $x = 0$ will propagate along the characteristic line, defined by $dx/dt = a$, and will arrive some point $x = l$ at time $t = l/a$. Defining $T = a/l$, of interest is the initial waveform magnitude as it arrives at $x = l$, which is given by $h(l,T^+)$. Noting that $\lim_{t\to T^+} h(l,t) = \lim_{t\to 0^+} h(l,T+t)$, a modified version of Laplace's initial value theorem leads to the relationship

$$\lim_{t\to T^+} h(l,t) = \lim_{|s|\to\infty} se^{sT}e^{-l\sqrt{(Ls+R)(Cs+s\hat{C}_R(s))}}\hat{u}(s) \tag{18}$$

where the first exponential term is introduced to account for the time-shift. The resulting exponent in Eq. (18) can be manipulated to yield

$$sT - l\sqrt{(Ls+R)(Cs+s\hat{C}_R)} = sT\left(1-\sqrt{1+\frac{R}{Ls}+\frac{\hat{C}_R(s)}{C}+\frac{R\hat{C}_R(s)}{LCs}}\right) \tag{19}$$

where, using the Binomial relationship $\sqrt{1+z} = 1+z/2-z^2/8+\cdots$, this can be expressed as

$$-\frac{sT}{2}\left(\frac{R}{Ls}+\frac{\hat{C}_R(s)}{C}+\frac{R\hat{C}_R(s)}{LCs}\right)+\cdots \tag{20}$$

which holds for large s, as the leading order of the Laplace-domain terms under the square root in the right side of Eq. (19) is $1/s$. Multiplying through by sT and considering Eq. (20) in the limit $s \to \infty$, the higher order terms are neglected (these are not represented in Eq. (20), but they are of order $1/s$ once multiplied by sT), the third term in the brackets approaches zero as $\lim_{|s|\to\infty} \widehat{C}_R(s) = 0$, leads to the following representation of Eq. (18) as

$$\lim_{t\to T^+} h(l,t) = \exp\left(-\frac{1}{2}\left(\frac{R}{L}+\frac{C_R(0)}{C}\right)T\right)\lim_{|s|\to\infty} s\widehat{u}(s) \tag{21}$$

where the relationship $\lim_{|s|\to\infty} s\widehat{C}_R(s) = C_R(0)$ is used. Finally, the first arrival of the wave front has the form

$$h\left(l,T^+\right) = e^{-\frac{1}{2}\left(\frac{R}{L}+\frac{C_R(0)}{C}\right)T} u\left(0^+\right) \tag{22}$$

APPENDIX C: Choices of prior probabilities

An algorithm to compute the prior probability in Eq. (5) is proposed here for a specific experiment and sensor. The delay times $\{\tau(P^{[k]})\}_{m=0}^{M-1}$ are first computed from the theoretical (yet inaccurate) wave-speed formula, and using $\sum_{\lambda\in P^{[k]}} \tilde{T}_\lambda$, which yields one delay value for each path $P^{[k]}$. Based on these computed delays, we construct prior probabilities π_m that favour paths with smaller delay magnitudes. Since shorter delays are more likely to represent physically plausible or dominant propagation paths, the prior for hypothesis m should increase when $\tau(P^{[m]}) < \tau(P^{[k]}), k \neq m$. To map the given discrete delays to a valid probability distribution, we assign each path a non-negative weight $w_m = \exp\left[-\beta\,\tau(P^{[m]})\right], \beta > 0$, which decreases monotonically with delay. The prior probabilities are then obtained by normalisation, $\pi_m = \frac{w_m}{\sum_{k=1}^{M} w_k}$, ensuring that $\sum_{m=1}^{M} \pi_m = 1$ and that paths with smaller $\tau(P^{[m]})$ receive higher prior probability.

Alternatively, to incorporate the principle that shorter propagation delays correspond to more plausible paths, we can assign prior probabilities based on the relative ordering of the discrete delay estimates. A physically reasonable assumption is that a path is more credible when its delay is strictly smaller than those of all competing paths. Accordingly, we adopt a rank-based prior that reflects the relative position of each path within the ordered set of delays. Since a smaller delay indicates a more physically plausible or dominant propagation path, we quantify the relative credibility of path m by counting the number of competing paths with larger delays. This score is defined as $s_m = |\{k \neq m : \tau(P^{[m]}) < \tau(P^{[k]})\}|$, where $|\,|$ indicates the number of elements in the set. Paths with shorter delays therefore obtain larger scores. The prior probability is then obtained by normalising these scores, $\pi_m = \frac{s_m}{\sum_{k=1}^{M} s_k}$, which ensures $\sum_{m=1}^{M} \pi_m = 1$.

APPENDIX D: Analysis details for case study 1

The first case study aimed to demonstrate a detailed, worked example of the proposed methodology for the calibration of wave speeds and friction factors. In this appendix, the detailed analysis results are provided in Tables 5 and 6, while the key results have been presented earlier in the relevant figures.

Table 4: Estimated wave speeds and their comparison with the wave speed prior

Link ID	Identifiable	Magnitude			Relative error (%)	
		$\bar{a}_{prior}$	a_{est}	a_{true}	$\bar{a}_{prior}$ vs. a_{true}	a_{est} vs. a_{true}
L1	No	1200	1200	1209	-	-
L2	No	1200	1200	1216	-	-
L3	**Yes**	**1200**	**1204.30**	**1206**	**0.50%**	**0.14%**
L4	No	1200	1200	1164	-	-
L5	**Yes**	**1200**	**1200.46**	**1210**	**0.83%**	**0.79%**
L6	**Yes**	**1200**	**1229.24**	**1235**	**2.83%**	**0.47%**
L7	**Yes**	**1200**	**1191.77**	**1180**	**1.69%**	**1.00%**
L8	**Yes**	**1200**	**1181.74**	**1155**	**3.90%**	**2.31%**
L9	**Yes**	**1200**	**1232.96**	**1228**	**2.28%**	**0.40%**
L10	**Yes**	**1200**	**1197.04**	**1198**	**0.17%**	**0.08%**
L11	No	1200	1200	1243	-	-
L12	No	1200	1200	1206	-	-
L13	No	1200	1200	1162	-	-

L14	**Yes**	**1200**	**1201.46**	**1203**	**0.25%**	**0.13%**
L15	**Yes**	**1200**	**1209.02**	**1207**	**0.58%**	**0.17%**

Table 5: Estimated friction factors and their comparison with the prior

Link ID	Observed	Magnitude			Relative error (%)	
		$\bar{f}_{prior}$	f_{est}	f_{true}	f_{prior} vs. f_{true}	f_{est} vs. f_{true}
L1	No	0.04	-	0.025	-	-
L2	No	0.04	-	0.024	-	
L3	**Yes**	**0.04**	**0.0401**	**0.04**	**0.00%**	**0.39%**
L4	No	0.04	-	0.031	-	-
L5	**Yes**	**0.04**	**0.037**	**0.034**	**17.65%**	**8.63%**
L6	**Yes**	**0.04**	**0.031**	**0.032**	**25.00%**	**3.54%**
L7	**Yes**	**0.04**	**0.030**	**0.03**	**33.33%**	**1.55%**
L8	**Yes**	**0.04**	**0.038**	**0.037**	**8.11%**	**2.15%**
L9	**Yes**	**0.04**	**0.028**	**0.028**	**42.86%**	**1.25%**
L10	**Yes**	**0.04**	**0.026**	**0.028**	**42.86%**	**6.93%**
L11	No	0.04	-	0.03	-	-
L12	No	0.04	-	0.024	-	-
L13	No	0.04	-	0.024	-	-
L14	**Yes**	**0.04**	**0.0397**	**0.034**	**17.65%**	**16.95%**
L15	**Yes**	**0.04**	**0.038**	**0.037**	**8.11%**	**3.66%**