

## Pipe roughness calibration in water distribution systems using grey numbers

Stefano Alvisi and Marco Franchini

### ABSTRACT

This paper presents a procedure based on the use of grey numbers for the calibration (with uncertainty) of pipe roughness in water distribution systems. The pipe roughness uncertainty is represented through the grey number amplitude (or interval). The procedure is of a wholly general nature and can be applied for the calibration (with uncertainty) of other parameters or quantities, such as nodal demands. In this paper, for the purpose of roughness calibration, a certain number of nodal head measurements made under different demand conditions is assumed to be available at different locations (nodes); all other topological and geometric characteristics of the system are considered to be known exactly.

The general approach to pipe roughness calibration (taking account of uncertainty) focuses on identifying the grey roughness values which produce grey head values at the measuring nodes such as to encompass the observed values grouped on the basis of the different demand scenarios and, at the same time, have as small an 'amplitude' as possible.

The proposed procedure was applied to two synthetic case studies and to one real network. The tests on the synthetic case studies show that the proposed procedure is able to correctly solve the inverse problem, i.e. it can identify the known grey roughness numbers even when they overlap; the same applies when the known grey roughness numbers collapse into known white roughness numbers. The test on the real case offers the possibility of highlighting the potentials of the procedure when applied within a context where measurement errors and other uncertainties are present. The procedure entails computing times that may become lengthy. However, it is possible to reduce these computing times considerably by replacing the hydraulic simulator—to which a number of calls must be made during the calibration procedure (for objective function evaluation)—with an approximation based on a first-order Taylor series expansion. This approach introduces acceptable approximations within the context of the problem considered.

**Key words** | calibration, grey numbers, hydraulic model, uncertainty, water distribution systems

**Stefano Alvisi** (corresponding author)  
**Marco Franchini**  
Dipartimento di Ingegneria,  
Università degli Studi di Ferrara,  
Via Saragat, 1,  
44100 Ferrara,  
Italy  
E-mail: [stefano.alvisi@unife.it](mailto:stefano.alvisi@unife.it)

### INTRODUCTION

Hydraulic simulation models of a water distribution system represent a tool that water utilities can usefully employ in a variety of activities including system design, rehabilitation, real-time management, etc. In fact, once system characteristics such as topological structure, pipe diameters, lengths and roughness and nodal demands have been fixed,

the utility company can assess the effects of different design or management decisions on the variables of interest, e.g. nodal heads, discharges and velocities in pipes.

However, the simulation model inputs may be affected by uncertainty and/or imprecision (Bargiela & Hainsworth 1989; Xu & Goulter 1996). For example, water demand

is characterized by a natural variability that leads to uncertainty in its quantification (e.g. Gargano & Pianese 2000; Kapelan *et al.* 2005a,b). Roughness can be quantified, but only with a certain degree of imprecision due to the impossibility of measuring it directly in each pipe (e.g. Walski 1983; Revelli & Ridolfi 2002).

For the utility company, or more specifically for the user of the simulation model, it is generally important to know how the uncertainty/imprecision affecting input variables will be reflected in the variables of interest that influence design and management decisions.

In the scientific literature there exist various methods for representing the uncertainty/imprecision present in input variables as well as suitable techniques for transferring it to the output variables of a prefixed model. For instance, the input variables can be interpreted as random variables having specific probability distributions. Starting from these probability distributions, it is possible to reconstruct those of the output variables using either an analytic or numerical approach (Cullen & Frey 1999). The analytic approach is possible only in a very limited number of cases (very simple ones) so that numerical techniques are increasingly relied upon to address real problems (Maskey & Guinot 2003). Among them, it is worth mentioning the Monte Carlo method (e.g. Kottogoda & Rosso 1998) or the Latin Hypercube method (McKay *et al.* 1979). Rather than considering the entire probability distribution, if we concentrate only on first- and second-order statistics of the output variables we can use e.g. First Order Second Moment Analysis (Dettinger & Wilson 1981).

An alternative to the random variable concept is provided by the fuzzy number concept (Zadeh 1965) by means of which a generic input variable can have a 'possibility distribution' associated with it. A fuzzy number can in fact be seen as a generalization of the confidence interval concept applied on a number of levels between 0 and 1, consistent with the 'membership function' concept (Xu & Goulter 1996). Where fuzzy numbers are taken to represent the input variables, fuzzy mathematics can be applied to transfer the uncertainty/imprecision to the output variables, thus arriving at a definition of the corresponding possibility distributions (or corresponding fuzzy numbers) (Kaufmann & Gupta 1985).

Finally, the uncertainty/imprecision associated with the simulation model inputs can be represented using *grey numbers* (Deng 1982). With this technique, uncertainty/imprecision is characterized simply by defining *one* interval and, in this respect, grey numbers can be considered like a fuzzy number confined to only one 'membership function' level (Karmakar & Mujumdar 2006), even though their conceptual background differs (Huang *et al.* 1995). As in the case of fuzzy numbers, it is possible to employ specific mathematics or 'grey mathematics' (Wang & Wu 1998) in order to arrive at a definition of the grey numbers representing the output variables.

With specific reference to the case of water distribution systems, Xu & Goulter (1996, 1998) use the random variables technique to deal with the uncertainty tied to input variables and transfer this uncertainty to the output variables by means of the Monte Carlo method. However, Kapelan *et al.* (2005a,b) use the Latin Hypercube method and Xu & Goulter (1996), Revelli & Ridolfi (2002) and Branisavljević & Ivetić (2006) use the fuzzy number technique. Similarly, Branisavljević *et al.* (2009) apply this technique under specific constraints to reduce the uncertainty on the nodal demands, whereas Xu & Goulter (1997) and Tolson *et al.* (2004) use first-order analysis.

The applications mentioned above all focus on the direct problem, in which the uncertainty/imprecision present in input variables (usually concentrated in nodal demand and/or pipe roughness), irrespective of the method of representation used, is quantified *a priori* on the basis of different information drawn either from manuals or operators' experience or, alternatively, from specific experimental campaigns and then transferred to the output variables using appropriate mathematical techniques.

However, when field measurements of output variables (nodal heads, pipe discharges, etc.) are available, the uncertainty associable with the input variables of the model (which, as previously observed, is generally concentrated in nodal demand and pipe roughness) can be quantified by means of an appropriately formulated calibration process.

In the realm of water distribution systems, examples of calibration—i.e. of the inverse problem, in which uncertainty is also considered—are to be found in e.g. Bush & Uber (1998) who analyze a D-optimal sampling design problem for water networks calibration and evaluate the covariance

matrix of the parameter estimates through the first-order second-moment method. A similar statistical approach is used by [Lansley \*et al.\* \(2001\)](#), while [Todini \(1999\)](#) proposes a new approach based on the Kalman filtering technique to solve the standard least-square type calibration problem: calibration parameters considered are pipe roughness coefficients alone for which the 'best' least-square estimate is found together with the relative variance. [Di Cristo & Leopardi \(2003\)](#) instead propose a procedure based on the application of a Monte Carlo type technique which, by starting from prefixed nodal demand probability distributions and the observed pressure values at several nodes, makes it possible to calibrate the pipe roughness coefficients and derive the corresponding probability distributions. Similarly, [Kapelan \*et al.\* \(2007\)](#) propose a calibration procedure for pipe roughnesses based on the Metropolis algorithm ([Metropolis \*et al.\* 1953](#)) which allows for the definition of the probability distributions of pipe roughnesses.

Parameter uncertainty was not considered, however, in the first studies relative to the calibration of the water distribution systems. Starting from the 1980s (e.g. [Walski 1983](#); [Ormsbee & Wood 1986](#); [Bhave 1988](#); [Lansley & Basnet 1991](#); [Datta & Sridharan 1994](#); [Reddy \*et al.\* 1996](#)), numerous calibration procedures were developed all sharing a common feature, namely that of seeking a set of values for parameters (typically roughness) which can best reproduce (in terms of least squares) the observed values.

Overall, for a complete and very updated review of the state of the art in calibration of water distribution systems, the reader can refer to [Savic \*et al.\* \(2009\)](#). For the calibration within the framework of the water resources, the reader can refer to [Beven \(2001\)](#) or to [Wagener \*et al.\* \(2004\)](#) where the problems of the uncertainty linked to the 'behavioural sets' ([Beven & Binley 1992](#)) and the relative identifiability are widely discussed. Incidentally, very recent applications of these concepts can be found in [Fang & Ball \(2007\)](#) where a procedure searching for behavioural sets and the relative ranges is described with reference to a hydrologic model (Storm Water Management Model - SWMM); another example is given by [van Griensven & Meixner \(2007\)](#) who propose a procedure based on a multi-objective optimization applied to a catchment water quality model where the parameter uncertainty is strictly related to their identifiability as a consequence of the quantity and quality of the

data. As a last example, [Cullmann & Wriedt \(2008\)](#) propose a procedure to identify the factors affecting both the event-based optimal parameter sets of a rainfall-runoff model and their temporal evolution. The procedure used is based on the combination of an algorithm for an event-based Parameter ESTimation (PEST, [Skahill & Doherty 2006](#)) and an algorithm for the DYNamic Identifiability Analysis (DYNIA, [Wagener \*et al.\* 2003](#)).

Furthermore, in contrast with the water distribution systems field, the water resources field offers examples of calibration studies based on fuzzy or grey mathematics. For example, [Ozelkan & Duckstein \(2001\)](#) present a conceptual rainfall-runoff model in which the uncertainty of the parameters is represented by means of fuzzy numbers and propose a procedure for calibrating these numbers based on fuzzy regression ([Tanaka \*et al.\* 1982](#); [Tanaka 1987](#); [Diamond 1988](#)). [Bardossy \*et al.\* \(2006\)](#) describe a process for calibrating the parameters of the Nash unit hydrograph, assuming them to be expressed as fuzzy numbers. [Wu \*et al.\* \(2005\)](#) propose a model for simulating runoff in an aquifer where the uncertainty associated with the parameters is represented by means of grey numbers calibrated on the basis of hydraulic head values observed in the field.

In light of the previously discussed points, this study aims to address the problem of calibration in the context of water distribution systems taking uncertainty into account through the use of grey numbers. Here we shall focus our attention specifically on the calibration of pipe roughness based solely on nodal head measurements; however, the procedure described below can also be extended to the calibration of other parameters or quantities e.g. water demand coefficients (useful for characterizing water demand over a 24-hour period) and other types of field measurements can be used, such as pipe discharges and tank and reservoir levels.

The reason for using grey numbers is their simple mathematics and their intuitive way of dealing with uncertainty which is simply represented through an interval without involving any probabilistic or possibility concept. In other terms, they offer practitioner engineers a robust tool to quantify the intuitive idea that uncertainty implies an interval/range without making any assumption about the value distribution within it while some simple mathematics allows such an uncertainty to be transferred to the

quantities of interest (nodal heads and pipe discharges). On the other hand, grey mathematics (as shown in this paper) allows the uncertainty related to pipe roughnesses (i.e. the grey roughness numbers) to be estimated starting from the interval covered, for instance, by the nodal heads under different demand conditions without making any assumption on such variability or dispersion.

For these reasons, we believe that the approach we are about to present in this paper is not a competitor of a probabilistic approach or an approach based on fuzzy mathematics but, instead, can be considered as a different way for dealing with uncertainty. Indeed, each of the three approaches previously mentioned (probabilistic, fuzzy and grey) has its own advantage. As observed by Xu & Goulter (1996), when adequate statistical data on the uncertainty are available (from which mean, variance and covariance of the uncertainty can be estimated/are of interest), then the approach based on probability analysis is the most appropriate. On the other hand, when uncertainty is related to an incomplete knowledge as opposed to a truly random variability, then the fuzzy mathematics is appropriate. When interest is only on the overall general dispersion observed without making any assumption on the behaviour of the quantity considered within such dispersion, then grey mathematics becomes appropriate.

We provide an introduction to grey numbers below and show how they can be used to quantify the uncertainty tied to the variables of interest (nodal heads, pipe discharges), based on the uncertainty in pipe roughness, once water demands have been assigned to the network nodes (direct problem). Thereafter we describe a procedure for estimating the grey numbers representing pipe roughness based on hydraulic head measurements at prefixed network nodes (inverse problem), after water demands have been assigned (e.g. demands are represented by their mean/expected values). We then present and discuss the results obtained by applying the procedure to two synthetic case studies and to one real case, and finally draw our conclusions.

## GREY NUMBERS

A grey number (Deng 1982, 1987; Huang *et al.* 1995; Karmakar & Mujumdar 2006) is defined as a number

whose exact value is unknown, but which falls within an interval that is known. A grey number thus enables us to represent the uncertainty associated with a given parameter by means of an interval whose upper and lower limits are known, although we have no information about the parameter's distribution within the interval itself (Liu & Lin 1998).

Formally, where  $x$  indicates a closed, limited set of real numbers, a grey number  $x^\pm$  can be defined as follows (Cheng *et al.* 2002; Yang *et al.* 2004):

$$x^\pm = [x^-, x^+] = \{x \in x^\pm | x^- \leq x \leq x^+\} \quad (1)$$

where  $x^-$  and  $x^+$  are the lower and upper limits of the interval, respectively. The grey number  $x^\pm$  becomes a 'white number' when  $x^\pm = x^- = x^+$ , whereas it is referred to as a 'black number' when  $x^\pm = [x^-, x^+] = [-\infty, +\infty]$  (Karmakar & Mujumdar 2006).

Since a grey number represents the interval of possible variation of a given parameter, to some extent it takes on the same meaning as an interval value with the same upper and lower limits (Alefeld & Herzberger 1983). Practically speaking, the arithmetic of grey numbers coincides with that of interval values (Liu *et al.* 2000), it being possible to define the following operations between grey numbers (Wang & Wu 1998):

$$\text{addition : } [x^-, x^+] + [y^-, y^+] \triangleq [x^- + y^-, x^+ + y^+] \quad (2)$$

$$\text{subtraction : } [x^-, x^+] - [y^-, y^+] \triangleq [x^- - y^+, x^+ - y^-] \quad (3)$$

$$\text{product : } [x^-, x^+] \times [y^-, y^+] \triangleq [\min\{x^-y^-, x^-y^+, x^+y^-, x^+y^+\}, \max\{x^-y^-, x^-y^+, x^+y^-, x^+y^+\}] \quad (4)$$

$$\text{division : } [x^-, x^+] \div [y^-, y^+] \triangleq [x^-, x^+] \times \left[ \frac{1}{y^+}, \frac{1}{y^-} \right] \quad (5)$$

where  $x^\pm = [x^-, x^+]$  and  $y^\pm = [y^-, y^+]$  are two grey numbers,  $x^-, x^+, y^-$  and  $y^+$  represent real numbers,  $\triangleq$  means 'being defined as' and, in the case of division, we must have  $0 \notin [y^-, y^+]$ .

We can also define a function  $f$  of  $n$  grey numbers  $x_1^\pm, \dots, x_i^\pm, \dots, x_n^\pm$  as (Yang *et al.* 2004):

$$[f(x_1^\pm, \dots, x_i^\pm, \dots, x_n^\pm)]^\pm = [[f(x_1^\pm, \dots, x_i^\pm, \dots, x_n^\pm)]^- , [f(x_1^\pm, \dots, x_i^\pm, \dots, x_n^\pm)]^+] \quad (6)$$

where  $[f(x_1^\pm, \dots, x_i^\pm, \dots, x_n^\pm)]^\pm$  is the grey function and  $[f(x_1^\pm, \dots, x_i^\pm, \dots, x_n^\pm)]^-$  and  $[f(x_1^\pm, \dots, x_i^\pm, \dots, x_n^\pm)]^+$  are the minimum and maximum values of the function, respectively. From a practical viewpoint, therefore, calculating the function  $f$  of  $n$  grey numbers  $x_1^\pm, \dots, x_i^\pm, \dots, x_n^\pm$  specifically means looking for (a) the set of real values  $x_1, \dots, x_i, \dots, x_n$  which minimizes the function where  $x_i^- \leq x_i \leq x_i^+$  and (b) the set of real values  $x_1, \dots, x_i, \dots, x_n$  (different from the previous one) which maximizes it, still contained in the domain  $x_i^- \leq x_i \leq x_i^+$ . In other words, it is necessary to solve two distinct problems of minimization and maximization of a function of  $n$  variables, each variable  $x_i$  being defined over an assigned interval  $x_i^\pm$ .

## THE DIRECT PROBLEM

We shall now address the procedure for solving the direct problem, which involves calculating the (grey numbers representing the) output variables (pipe discharges and nodal heads) once the input variables have been assigned. More precisely, network topology is assumed known and lengths and diameters are given (i.e. white numbers). Nodal demands are likewise treated as white numbers, whereas roughnesses are treated as grey numbers. A distinction is made between two cases: in the first case, the equations of energy and mass balance are solved simultaneously whereas, in the second case, the 'hydraulic solver' is replaced by a Taylor series expansion around the solution relative to the mid-value of the grey number representing the roughness of pipes. This latter case is considered because it enables computing times to be greatly reduced, as shown below.

### Solution of the direct problem by means of a hydraulic simulator

Let us consider a generic water distribution network made up of  $N_p$  pipes and  $N_n$  nodes, of which  $N_a$  nodes have a known head and  $N_b$  an unknown head ( $N_n = N_a + N_b$ ). Let  $L_i$  be the length,  $D_i$  the diameter and  $ks_i$  the roughness (according to Gauckler-Strickler (Gauckler 1868)) of each pipe (with  $i = 1:N_p$ ) and  $q_j$  the water demand at each node of unknown head (with  $j = 1:N_b$ ). Moreover, let the head

$HO_k$  be known at nodes with an imposed head (where  $k = 1:N_a$ ). Let the unknowns be represented by the discharges  $Q_i$  in  $N_p$  pipes (with  $i = 1:N_p$ ) and the heads  $H_j$  at the  $N_b$  nodes of unknown head (with  $j = 1:N_b$ ).

If each of the inputs ( $L_i$ ,  $D_i$ ,  $ks_i$ ,  $q_j$ , and  $HO_k$ ) were exactly known, i.e. could be represented by white numbers, the outputs (namely, the unknowns  $Q_i$  and  $H_j$ ) would also be white numbers and could be calculated by simultaneously solving the  $N_b$  equations of mass balance at the nodes and  $N_p$  equations of energy balance in the pipes, whose formulation can be found in a number of textbooks and manuals (e.g. Mays 2005, pp. 459–469). The solution of such a system can be obtained using one of the various algorithms that have been developed over the years (e.g. Cross 1936; Shamir & Howard 1968). Of these, it is worth mentioning the global gradient algorithm proposed by Todini & Pilati (1987), widely used in various software applications, e.g. EPANET (Rossman 2000), and also used in the hydraulic simulator implemented in the MATLAB<sup>TM</sup> environment for the procedure proposed here (see the application section).

If, on the other hand, one or more input variables are known with a certain degree of uncertainty, i.e. can be represented by grey numbers, grey mathematics can be used to transfer this uncertainty to the output variables. Let us assume, for example, that the roughness of each pipe in the network is affected by uncertainty and therefore may be represented with a grey number  $ks_i^\pm$  (with  $i = 1:N_p$ ). Given that pipe discharges and nodal heads are a function of roughness, among other things, they will similarly be represented by means of grey numbers  $Q_i^\pm$  (with  $i = 1:N_p$ ) and  $H_j^\pm$  (with  $j = 1:N_b$ ). Specifically, with reference to the head at the  $j$ th node, using the notation introduced in Equation (6) we can write:

$$\begin{aligned} H_j^\pm &= [f_j(ks_1^\pm, \dots, ks_i^\pm, \dots, ks_{N_p}^\pm)]^\pm \\ &= \left[ [f_j(ks_1^\pm, \dots, ks_i^\pm, \dots, ks_{N_p}^\pm)]^- \right. \\ &\quad \left. \times [f_j(ks_1^\pm, \dots, ks_i^\pm, \dots, ks_{N_p}^\pm)]^+ \right] \end{aligned} \quad (7)$$

where the lower limit (a)  $H_j^- = [f_j(ks_1^\pm, \dots, ks_i^\pm, \dots, ks_{N_p}^\pm)]^-$  and upper limit (b)  $H_j^+ = [f_j(ks_1^\pm, \dots, ks_i^\pm, \dots, ks_{N_p}^\pm)]^+$  of the grey number representing the head at the  $j$ th node can be



calculated by looking for (a) the set of real values  $ks_1, \dots, ks_i, \dots, ks_{Np}$  with  $ks_i^- \leq ks_i \leq ks_i^+, i = 1:Np$ , which, when input into the system constituted by the equations of energy and mass balance, will provide us with the minimum value for the head at the  $j$ th node and (b) the set (different from the previous one) that will provide us with the maximum value for the head at the  $j$ th node, respectively. In other words, in order to be able to calculate the grey number representing the head at the  $j$ th node from the grey numbers representing pipe roughnesses it is necessary to solve two distinct non-linear problems (given the nature of the equations of energy and mass balance): one of minimization and one of maximization. Each of the two problems can be solved by means of the procedure schematically illustrated in Figure 1, where we assume to have at our disposal an optimizer which solves for each node  $j$  the minimization and maximization problems previously mentioned; the maximum number of optimizations is then  $2 \times (Np + Nb)$ . Fewer optimization problems are possible if the interest is in some node or pipe alone.

**Solution of the direct problem by means of a Taylor series expansion**

As shown in the previous section, within each optimization problem it is necessary to call the hydraulic simulator a repeated number of times in order to compute the upper (or lower) limit of the grey number of the variable of interest from the grey numbers representing the inputs. From an operational standpoint, this may imply long computing times, especially in the case of real networks characterized by a large number of nodes and pipes. A possible way to reduce computing times might be to express the link that exists between the input variables affected by uncertainty

and the variable of interest in an approximate form, through a first-order Taylor series expansion. For example, with reference to the case considered in the previous section, the head  $H_j$  at the generic node  $j$  can be expressed as a function of pipe roughness through the following first-order Taylor series expansion:

$$H_j \cong H_j^* + \sum_{i=1}^{Np} \frac{\partial H_j}{\partial ks_i} (ks_i - ks_i^*) \tag{8}$$

where  $H_j^*$  represents the (white) head at the  $j$ th node, calculated by means of the hydraulic simulator on the basis of prefixed reference roughness values ('white' values)  $ks_1^*, \dots, ks_i^*, \dots, ks_{Np}^*$  (normally assumed to be equal to the mid-value of the grey number representing the roughness of each pipe) and a pre-selected or reference demand scenario.

If reference to all Nb nodes is made, Equation (8) can be written as follows:

$$\begin{pmatrix} H_1 \\ \vdots \\ H_j \\ \vdots \\ H_{Nb} \end{pmatrix} = \begin{pmatrix} H_1^* \\ \vdots \\ H_j^* \\ \vdots \\ H_{Nb}^* \end{pmatrix} + \begin{pmatrix} \frac{\partial H_1}{\partial ks_1} & \dots & \dots & \frac{\partial H_1}{\partial ks_{Np}} \\ \vdots & & & \vdots \\ \frac{\partial H_j}{\partial ks_1} & \dots & \dots & \frac{\partial H_j}{\partial ks_{Np}} \\ \vdots & & & \vdots \\ \frac{\partial H_{Nb}}{\partial ks_1} & \dots & \dots & \frac{\partial H_{Nb}}{\partial ks_{Np}} \end{pmatrix} \begin{pmatrix} (ks_1 - ks_1^*) \\ \vdots \\ (ks_i - ks_i^*) \\ \vdots \\ (ks_{Np} - ks_{Np}^*) \end{pmatrix} \tag{9}$$

or, in matrix form,

$$H = H^* + A(ks - ks^*) \tag{10}$$

where the generic term  $a_{ji}$  of matrix  $A$  can be calculated by means of a central difference approximation written in  $ks_i^*$

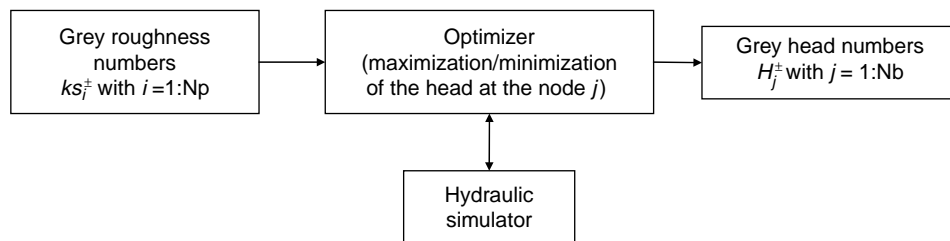


Figure 1 | Procedure for solving the direct problem based on the hydraulic simulator.

with an interval amplitude  $2\Delta ks_i$ :

$$a_{ji} = \frac{\partial H_j}{\partial ks_i} = \frac{H_j(ks_i^* + \Delta ks_i) - H_j(ks_i^* - \Delta ks_i)}{2\Delta ks_i} \quad (11)$$

where  $H_j(ks_i^* + \Delta ks_i)$  represents the head at the  $j$ th node, calculated by hydraulic simulation (with the reference water demands at the nodes) on the basis of the altered roughness value  $ks_i^* + \Delta ks_i$  in the  $i$ th pipe, maintaining a constant roughness in all other pipes.

In reference to Figure 1, the computational times of the process for estimating the grey head numbers can be significantly reduced by replacing the hydraulic model based on the equations of energy and mass balance with the Taylor series of Equation (10), taking into account that the heads  $H_j^*$  and terms  $a_{ji}$  of matrix  $\mathbf{A}$  are calculated only once outside the maximization/minimization cycle. Of course, if different demand conditions are considered, then different matrices  $\mathbf{A}$  have to be calculated (one for each demand condition).

Finally, it is worth observing that when the first-order Taylor series is considered, calculating the upper (or lower) limit of the variable of interest  $H_j^+$  (or  $H_j^-$ ) (with  $j = 1:Nb$ ) from the grey roughness numbers  $ks_i^\pm$  (with  $i = 1:Np$ ) proves to be very simple, since the solution is analytical because the model is perfectly linear. This implies that the upper limit of the variable of interest  $H_j^+$  is obtained by assuming, for each roughness value  $ks_i$  (with  $i = 1:Np$ ), the upper limit  $ks_i^+$  of the corresponding grey number if the coefficient  $a_{ji}$  is positive, and the lower limit  $ks_i^-$  if the coefficient  $a_{ji}$  is negative. The inverse applies for the calculation of the lower limit  $H_j^-$ . This simple rule cannot be applied with reference to the hydraulic simulator since this latter term is not linear. This implies that  $H_j^+$  (or  $H_j^-$ ) may be related neither to  $ks_i^+$  nor to  $ks_i^-$  (with  $i = 1:Np$ ), but instead to generic roughness values  $ks_i$  (with  $i = 1:Np$ ), each of them contained within the interval defined by  $ks_i^\pm$  (with  $i = 1:Np$ ).

In conclusion, by replacing the non-linear system of equations of energy and mass balance with an approximated relation represented by a first-order Taylor series expansion, it is possible to significantly reduce the computing burden. Of course, this implies a certain degree of approximation/error in the value of the variable of interest.

The effects of the approximation introduced using the first-order Taylor series expansion is thoroughly addressed in the discussion on the results of the numerical application.

## THE INVERSE PROBLEM TREATED USING GREY NUMBERS

In contrast to the direct problem previously described, the so-called inverse problem involves estimating the value of one or more inputs (e.g. pipe roughnesses) and the associated uncertainty from the observed values of one or more outputs (e.g. nodal heads), possibly taking account of their (natural) variability.

Let us again consider a generic water distribution system for which the following data are available: the topological characteristics of the network, the length  $L_i$  and diameter  $D_i$  of each pipe, the head  $HO_k$  at the nodes with imposed head and the water demand at the nodes. In other terms, all these latter quantities are represented through white numbers in the procedure described below. Furthermore, let us assume that the head is measured several times at  $N_{mn}$  measuring nodes and, at each measuring node  $mn$ , under  $N_{dc}$  demand conditions (e.g. for each of the 24 hours of the day and/or for different scenarios relative to the opening of several hydrants). Thus, let us assume that for a given measuring node  $mn$  ( $mn = 1:N_{mn}$ ) and demand condition/scenario  $dc$  ( $dc = 1:N_{dc}$ ), a set  $\Omega_{mn,dc}$  of observed heads is available (e.g. the set contains the heads observed for several days at a specific node and at a specific hour of the day and/or the heads observed when some specific hydrants have been opened). In addition, let  $H_{mn,dc,\omega}^{obs}$  be the  $\omega$ th observed head value of this set.

For the purpose of estimating roughness, the  $N_p$  pipes in the network can be grouped into  $N_r$  (with  $1 \leq N_r \leq N_p$ ) homogeneous classes based on their characteristics: for example, pipes of the same material and diameter, installation period, etc. can be grouped together (Mallick *et al.* 2002).

In this study, the calibration problem consists of determining, for each pipe class, the grey roughness number based on the sets of observed head values  $\Omega_{mn,dc}$  (with  $mn = 1:N_{mn}$  and  $dc = 1:N_{dc}$ ). More precisely, the calibration problem can be formulated as a search for the  $N_r$  grey numbers representing the roughness uncertainty of

the  $N_r$  classes of pipes such that, for each measuring node  $mn$  and for each demand condition  $dc$ , the grey number

$$O.F. = \sum_{mn=1}^{N_{mn}} \sum_{dc=1}^{N_{dc,mn}} \frac{H_{mn,dc}^+(ks_1^\pm, \dots, ks_{ic}^\pm, \dots, ks_{N_r}^\pm) - H_{mn,dc}^-(ks_1^\pm, \dots, ks_{ic}^\pm, \dots, ks_{N_r}^\pm)}{H_{mn,dc}^{obs}} \cdot \prod_{v=1}^{N_v} \exp(\delta_{v,mn,dc}) \quad (14)$$

representing the head at the measuring node (resulting from a direct simulation) contains the sets of observed head values  $\Omega_{mn,dc}$  and, at the same time, has the smallest ‘amplitude’ (understood as the difference between the upper and lower extremes). That is, the grey head numbers that best approximate the observed head intervals at each measuring node under the different demand conditions fit as ‘tightly’ as possible around them.

Formally, where  $ks_{ic}^\pm$  ( $ic = 1:N_r$ ) indicates one of the  $N_r$  grey numbers which represent the roughness to be calibrated and  $H_{mn,dc}^\pm(ks_1^\pm, \dots, ks_{ic}^\pm, \dots, ks_{N_r}^\pm)$  the grey number representing the head at the node  $mn$ , for the demand condition  $dc$  expressed as a function of the  $N_r$  grey roughness numbers, the solution of the calibration problem is given by  $ks_1^\pm, \dots, ks_{ic}^\pm, \dots, ks_{N_r}^\pm$  so that:

$$\sum_{mn=1}^{N_{mn}} \sum_{dc=1}^{N_{dc,mn}} [H_{mn,dc}^+(ks_1^\pm, \dots, ks_{ic}^\pm, \dots, ks_{N_r}^\pm) - H_{mn,dc}^-(ks_1^\pm, \dots, ks_{ic}^\pm, \dots, ks_{N_r}^\pm)] \text{ is minimum} \quad (12)$$

subject to

$$H_{mn,dc}^-(ks_1^\pm, \dots, ks_{ic}^\pm, \dots, ks_{N_r}^\pm) \leq H_{mn,dc,\omega}^{obs} \leq H_{mn,dc}^+(ks_1^\pm, \dots, ks_{ic}^\pm, \dots, ks_{N_r}^\pm) \quad (13)$$

$$\forall \omega \in \Omega_{mn,dc}, \text{ and } mn = 1 : N_{mn}, dc = 1 : N_{dc,mn}$$

The calibration problem can therefore be seen as a problem of minimizing the total amplitude of the grey numbers representing the dependent variable (the head at the nodes of observation) while satisfying the constraint given by the fact that these grey numbers must include all of the observed values at each node  $mn$  for each demand condition  $dc$ . Analogously, it may be seen as a problem of minimizing the objective function (OF) (Equation (14)), in which satisfaction of the constraint is guaranteed by penalizing the solutions that violate it, i.e. those for which

at least one observed value falls outside the ‘interval’ of the simulated grey head:

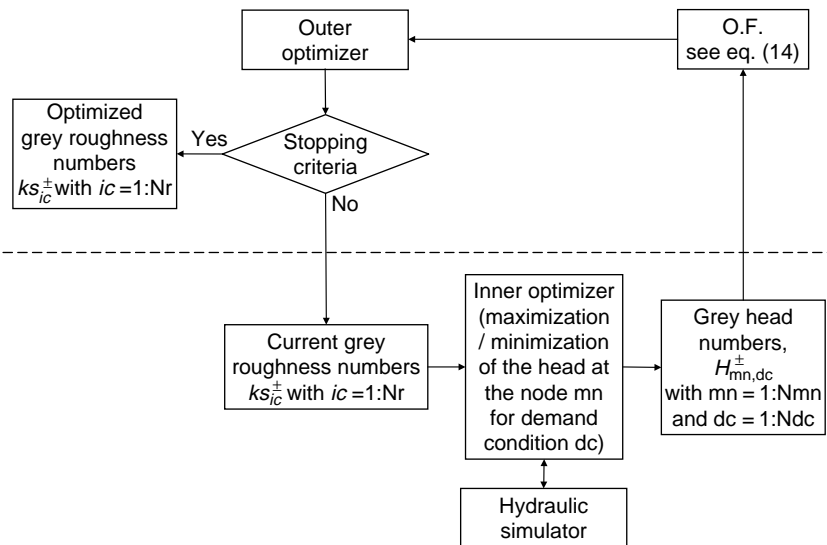
where the first term represents the sum (over  $N_{mn}$  measuring nodes and over  $N_{dc,mn}$  demand conditions) of the amplitudes of the grey heads, rendered dimensionless by normalizing them with respect to the mean value of the observed heads at the corresponding nodes and at the corresponding demand conditions. The second term represents the penalty to be attributed for  $N_v$  violations  $\delta_{v,mn,dc}$  at the  $mn$ th measuring node for the  $dc$ th demand condition:

$$\delta_{v,mn,dc} = \begin{cases} \frac{H_{mn,dc,v}^{obs} - H_{mn,dc}^+(ks_1^\pm, \dots, ks_{ic}^\pm, \dots, ks_{N_r}^\pm)}{H_{mn,dc,v}^{obs}} \cdot 100 & \text{if } H_{mn,dc,v}^{obs} > H_{mn,dc}^+ \\ \frac{H_{mn,dc}^-(ks_1^\pm, \dots, ks_{ic}^\pm, \dots, ks_{N_r}^\pm) - H_{mn,dc,v}^{obs}}{H_{mn,dc,v}^{obs}} \cdot 100 & \text{if } H_{mn,dc,v}^{obs} < H_{mn,dc}^- \end{cases} \quad (15)$$

It is worthwhile observing that the calibration problem now formulated is closely connected to the direct problem. In order to evaluate the objective function it is necessary to estimate the grey number representing the dependent variable (the grey head numbers at the  $N_{mn}$  nodes of observation relative to the different demand conditions  $N_{dc}$ ) based on the grey roughness numbers, even if grouped together into  $N_r$  classes.

The procedure can be schematized as a cycle (see Figure 2). The main/external optimizer produces generic solutions of the calibration problem consisting of the grey numbers  $ks_{ic}^\pm$  (with  $ic = 1:N_r$ ) which represent the roughnesses of the  $N_r$  classes of pipes making up the network. The method used to generate the generic grey number representing the roughness (i.e. the decision variables of the main/external optimization process) is described in the appendix. These  $N_r$  grey numbers are then assigned to each pipe according to their membership, thus producing  $N_p$  grey roughness numbers. The second/inner optimizer, starting from these  $N_p$  grey roughness numbers, produces the head grey numbers at the measuring nodes ( $mn = 1:N_{mn}$ ) for the different demand conditions





**Figure 2** | Procedure for solving the calibration problem based on the hydraulic simulator: the part below the dashed line represents the inner/secondary optimization problem which coincides with the procedure for solving the direct problem.

( $dc = 1:Ndc$ ) by using the hydraulic simulator. These grey head numbers are used to evaluate the OF (see Equation (14)). The procedure is stopped when one of the main/external optimizer termination criteria is satisfied, thus providing the optimal solution, i.e. the (optimal)  $Nr$  grey roughness numbers and, at the same time, the  $Nmn$  grey head numbers at the measuring nodes.

It is worth noting that the procedure illustrated here is based on the repetition of the direct problem each time that a tentative solution  $ks_{ic}^{\pm}$  ( $ic = 1:Nr$ ) is evaluated. In the context of the direct problem, it is possible to replace the hydraulic simulator with a first-order Taylor series expansion to achieve a considerable reduction in computing times at the cost of a modest degree of error/approximation in the results obtained.

## APPLICATIONS

The calibration procedure was applied to two case studies, one synthetic and the other relative to a real system. The first case considers a simple network made up of 9 nodes and 12 pipes for a total of 4 loops (see Figure 3). The layout of the network is derived from an example given by Marchi & Rubatta (1981) in relation to an inspection of parts of the Azienda Municipalizzata Gas e Acqua (AMGA) water distribution system in Genoa; it has also been used by

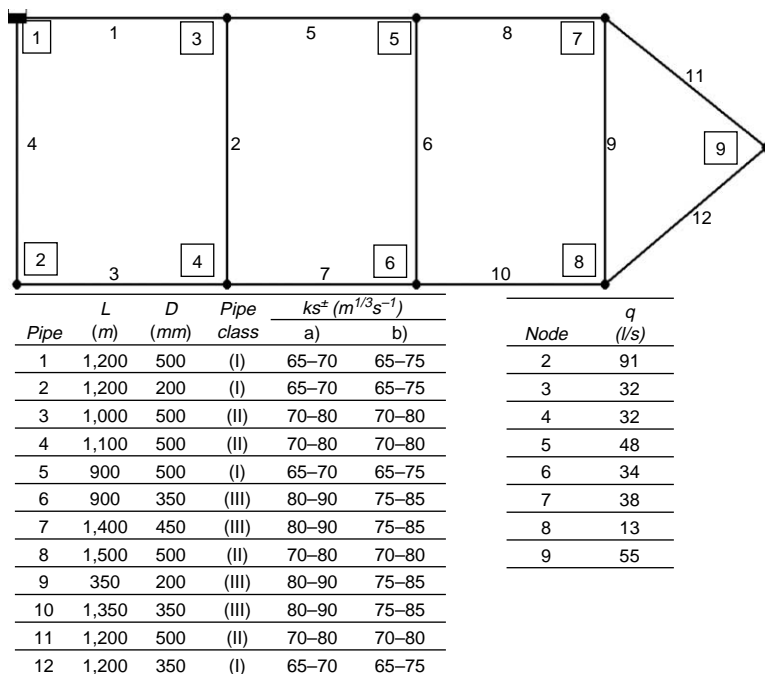
Revelli & Ridolfi (2002) in an application for analyzing water distribution systems by means of fuzzy numbers within the framework of a direct problem. The second case study is represented by a real water distribution network made up of around 400 pipes (see Figure 4). This network is part of the water distribution system of Ferrara (Italy).

The aim of the first case study was to show the ability of the procedure to identify known solutions, i.e. to solve the inverse problem correctly. For this purpose, two tests were performed. In the first test we tested the capability of the calibration procedure to identify known grey roughness numbers, while in the second test we tested its capability to collapse into known white roughness numbers.

The aim of the second case study, i.e. the real water distribution system, was to test the procedure on a fairly complex network where measurements are affected by errors and the nodal water demands can only be estimated. With reference to this real water distribution system, both the calibration procedure based on the hydraulic simulator and that based on the first-order Taylor series expansion were applied.

### Synthetic case study: identification of known grey roughness numbers

In this example, the length and diameter of each pipe were assumed to be known and the pipes were grouped into three



**Figure 3** | Synthetic case study: layout of the network and associated characteristics (L: pipe length, D: pipe diameter, ks: Gauckler-Strickler roughness coefficients, q: mean water demand).

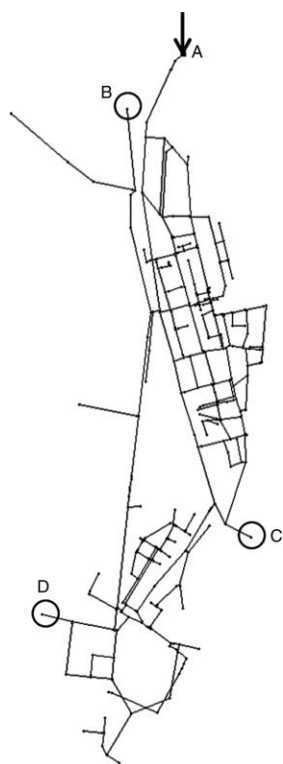
classes ( $N_r = 3$ ); for each class a grey roughness number  $ks_{ic}^{\pm}$  was assigned. The roughness was expressed according to the Gauckler-Strickler coefficient (it may be worth recalling that the Gauckler-Strickler roughness coefficient  $ks$  is equal to  $1/n$  where  $n$  is the Manning roughness coefficient). Furthermore, two sets of grey roughness numbers were considered (see Figure 3). In the set  $a$ , the grey roughness numbers do not overlap ( $ks_1^{\pm} = [65,70]$ ;  $ks_2^{\pm} = [70,80]$ ;  $ks_3^{\pm} = [80,90]$ ) whereas in the set  $b$  the grey roughness numbers do overlap ( $ks_1^{\pm} = [65,75]$ ;  $ks_2^{\pm} = [70,80]$ ;  $ks_3^{\pm} = [75,85]$ ). For each of the two sets of grey roughnesses, the corresponding grey heads at all the nodes of the water distribution systems were computed by solving the direct problem. In order to do this, the water demands shown in Figure 3 (i.e. one demand condition alone) were considered. Incidentally, the hydraulic simulator implemented in the MATLAB<sup>TM</sup> environment, necessary to perform the direct problem (see Figure 1) was based on the global gradient method (Todini & Pilati 1987).

Once the direct problem had been solved, the inverse problem was then considered. It was assumed that the head values observed at different measuring nodes were represented by the two values corresponding to the two

extremes of the grey heads previously produced. The calibration procedure, based on the use of the hydraulic simulator, was applied considering different sub-cases, i.e. assuming that different numbers  $N_{mn}$  of measuring nodes were available ( $N_{mn} = 1, 2, 3, 4$  and  $5$ ) and, for each number  $N_{mn}$  of measuring nodes, two distinct sets of nodes were considered.

Regarding the secondary/inner and the main/outer optimizers (see Figure 2), use was made of the MATLAB<sup>TM</sup> *fmincon* function based on Sequential Quadratic Programming (Powell 1978, 1983; Schittowski 1985) (incidentally, of the many options offered by the *fmincon* function, only that relevant to the definition of the lower and upper bounds of the independent/decisional variables was activated, thus searching for the optimum of the objective function of interest over a closed domain). For the purpose of verifying the effectiveness of the proposed procedure, the search for the grey roughness number was carried out over a wide domain  $D$  defined between  $d_1 = 50 m^{1/3} s^{-1}$  and  $d_2 = 120 m^{1/3} s^{-1}$  (for the meaning of the symbols  $d_1$  and  $d_2$  see the appendix).

Table 1 shows the values of the grey roughness numbers produced by the calibration procedure for the different



**Figure 4** | Real case study: layout of the water distribution system. Node A is the feeding point of the network where the head is set and discharge measured whereas nodes B, C and D are the pressure measuring nodes used for the calibration.

sets of measuring nodes when the set *a* of grey roughnesses (i.e. the set of grey roughness numbers which do not overlap) is considered. Similarly, but with reference to the set *b* (i.e. the sets of grey roughness numbers overlap), Table 2 shows the calibrated grey roughness numbers for the same measuring nodes.

Table 1 shows that when three or more measuring nodes are considered, the proposed procedure is capable of identifying the correct grey roughness number for each pipe class. On the contrary, when one or two measuring nodes are considered, independently of the measuring nodes selected, the calibration procedure provides grey roughness numbers which are inconsistent with the known solutions. This is comprehensible if we consider that, in this test case, just one demand condition was used and that the number of pipe classes (and thus of grey roughness numbers to be calibrated) is  $N_r = 3$ . Thus, since the number  $N_{mn}$  of measuring nodes (one or two) multiplied by the number of demand conditions (just one) is smaller than the number

**Table 1** | Synthetic case study, set (a): not overlapping grey roughness numbers for the different pipe classes. Grey roughness numbers produced by the calibration procedure in the cases where observed head values are available at different sets of nodes

Nmn	Measuring nodes	$ks \pm (m^{1/3} s^{-1})$		
		Pipe class (I)	Pipe class (II)	Pipe class (III)
		[65, 70]	[70, 80]	[80, 90]
1	4	[75, 86]	[64, 67]	[69, 103]
1	9	[69, 76]	[67, 73]	[69, 97]
2	4, 9	[71, 72]	[66, 77]	[70, 93]
2	2, 7	[69, 70]	[68, 79]	[68, 96]
3	2, 5, 8	[65, 70]	[70, 80]	[80, 90]
3	7, 8, 9	[65, 70]	[70, 80]	[80, 90]
4	2, 3, 6, 7	[65, 70]	[70, 80]	[80, 90]
4	5, 6, 8, 9	[65, 70]	[70, 80]	[80, 90]
5	2, 3, 4, 5, 6	[65, 70]	[70, 80]	[80, 90]
5	2, 3, 5, 8, 9	[65, 70]	[70, 80]	[80, 90]

$N_r = 3$  of grey roughness numbers to be calibrated, the calibration problem is under-determined. Consequently, the calibration procedure is not able to provide the correct grey roughness numbers. In other words, in order to guarantee identifiability, a minimum number of measuring nodes multiplied by the number of demand conditions has to be considered.

A similar consideration can be made with reference to the results of Table 2. However, in this second case, it can

**Table 2** | Synthetic case study, set (b): overlapping grey roughness numbers for the different pipe classes. Grey roughness numbers produced by the calibration procedure in the cases where observed head values are available at different sets of nodes

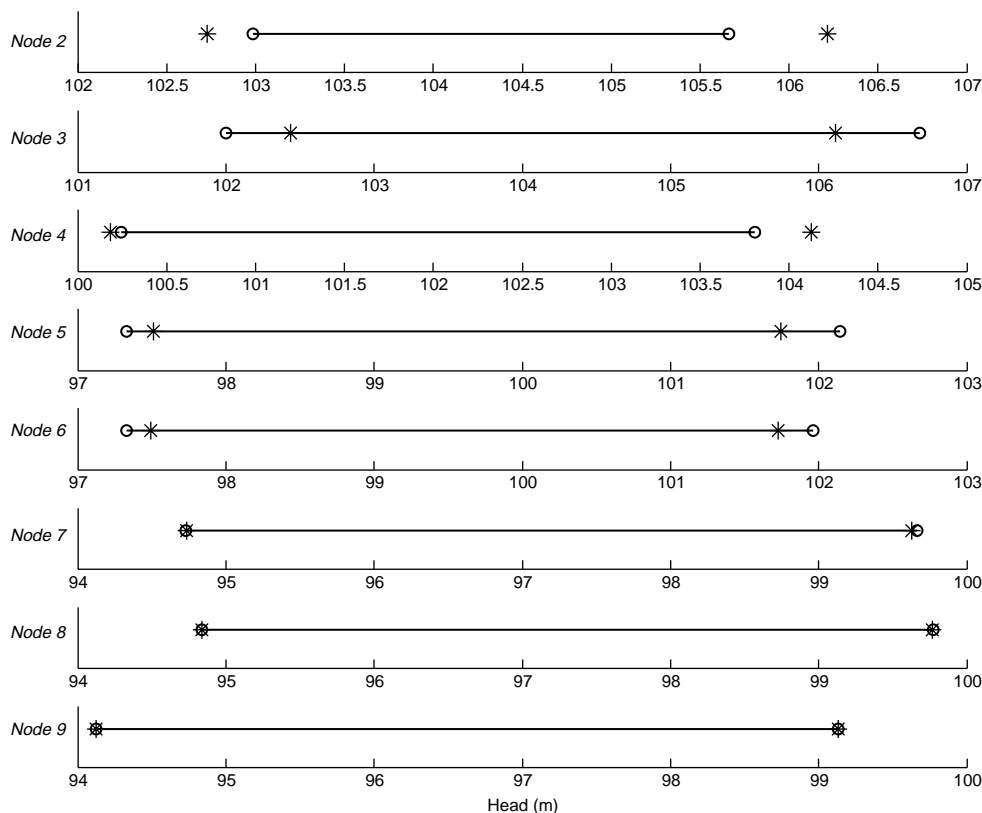
Nmn	Measuring nodes	$ks \pm (m^{1/3} s^{-1})$		
		Pipe class (I)	Pipe class (II)	Pipe class (III)
		[65, 75]	[70, 80]	[75, 85]
1	4	[71, 81]	[66, 76]	[81, 99]
1	9	[72, 83]	[63, 72]	[70, 84]
2	4, 9	[71, 77]	[66, 77]	[63, 90]
2	2, 7	[69, 73]	[68, 80]	[62, 99]
3	2, 5, 8	[65, 75]	[70, 80]	[75, 85]
3	7, 8, 9	[63, 81]	[72, 74]	[79, 80]
4	2, 3, 6, 7	[65, 75]	[70, 80]	[75, 85]
4	5, 6, 8, 9	[65, 75]	[70, 80]	[75, 85]
5	2, 3, 4, 5, 6	[65, 75]	[70, 80]	[75, 85]
5	2, 3, 5, 8, 9	[65, 75]	[70, 80]	[75, 85]

also be observed that even when  $N_{mn} = 3$  (recall that  $N_{dc_{mn}} = 1 \forall mn$ ), for a particular set of measuring nodes (nodes 7, 8 and 9) the calibrated grey roughness numbers are quite different from the real ones. However, Figure 3 shows that the three measuring nodes considered are not distributed over the water distribution system, since they are located near each other in a restricted zone of the pipe network. At least for this case, they are not representative of the entire piezometric surface.

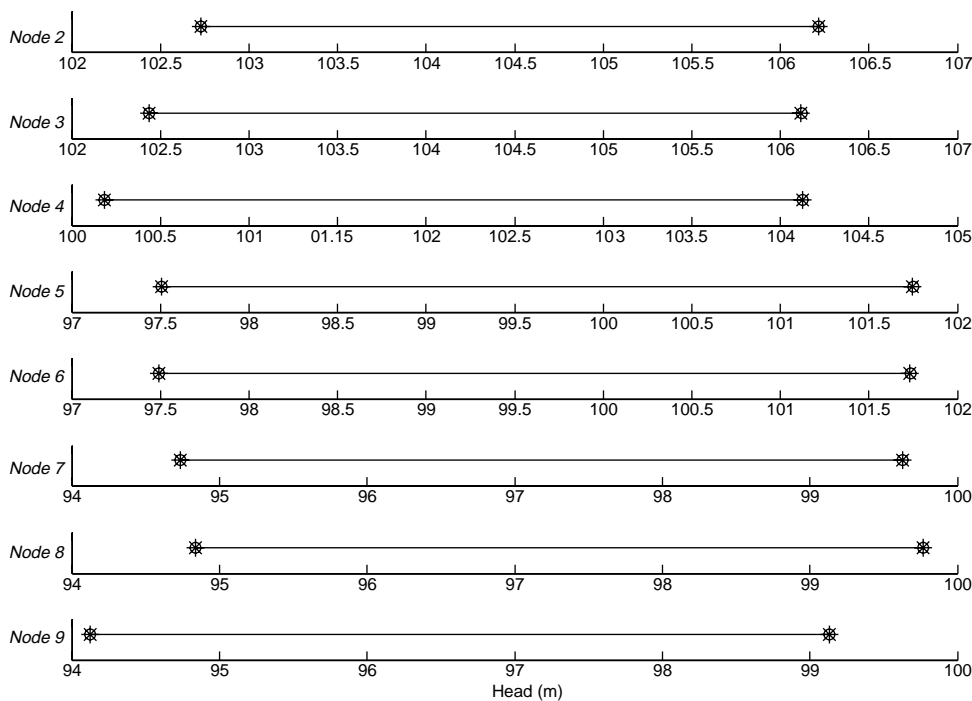
This is confirmed by Figure 5 where, for all the nodes of the water distribution system, the grey head numbers obtained on the basis of the grey roughness numbers calibrated according to the measurements at the nodes 7, 8 and 9 are compared with the available head values (symbol \*) in all the nodes. As can be seen, the calibrated grey roughness numbers, even if not consistent with the real ones, lead to grey head numbers at the calibration nodes (see nodes 7, 8 and 9) which perfectly include the observed head values by fitting as tightly as possible around them,

consistently with the proposed calibration problem formulation (see Equations (14) and (15)). At the same time, this does not happen for all the other nodes, whose head values were not considered during the calibration. In fact, for example at node 2 or 4, 'available/observed' head values fall outside the simulated grey head numbers.

When three nodes more uniformly distributed over the water distribution system are considered (see the case of measuring nodes 2, 5, 8), the grey roughness numbers calibrated are consistent with the real ones. Figure 6 shows the grey head numbers obtained on the basis of these calibrated grey roughness numbers contrasted with the head values occurring at all the nodes of the water distribution system. It is worth noting that the calibrated grey roughness numbers lead to grey head numbers fitting as tightly as possible around the heads used in the calibration phase and, at the same time, correctly reproduce the head 'variability' at the nodes not used in the calibration phase.



**Figure 5** | Synthetic case study: grey head numbers (solid line) deriving from the calibration procedure based on head measurements available at nodes 7, 8 and 9 for the set  $b$  of assigned grey roughnesses (see also Table 2). The available/measured head values are represented with the symbol \*.



**Figure 6** | Synthetic case study: grey head numbers (solid line) deriving from the calibration procedure based on head measurements available at nodes 2, 5 and 8 for set b of assigned grey roughnesses (see also Table 2). The available/measured head values are represented with the symbol \*.

Thus, these results highlight that not only a minimum number of measuring nodes (plus demand conditions) is necessary but these nodes also have to be well located along the water distribution system. In other terms, the sampling design represents a fundamental preliminary step that has to be faced in order to set up the calibration problem correctly, as highlighted by several authors (e.g. Kapelan *et al.* 2003).

Finally, it is worth observing that when  $N_{mn} > 3$  and  $N_{dc_{mn}} = 1 \forall mn$ , the calibration procedure was able to correctly reproduce the real grey roughness number of each pipe class in all the test cases considered.

### Synthetic case study: identification of known white roughness numbers

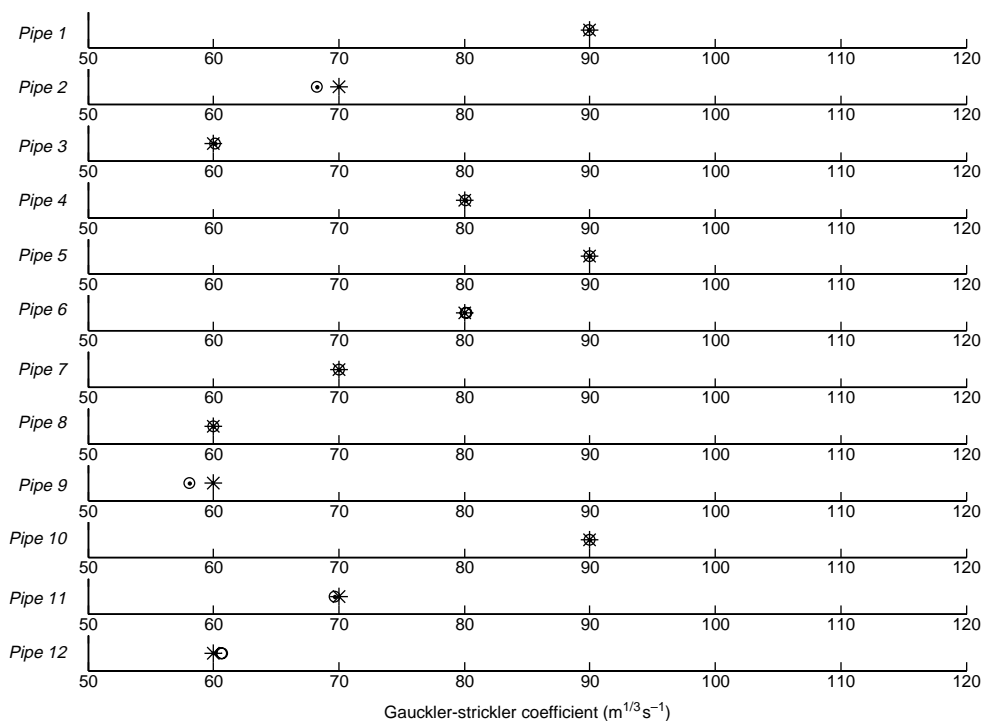
A second experiment was performed on the water distribution system shown in Figure 3 to highlight the capability of the proposed calibration procedure to identify the pipe roughnesses properly, even when they are represented by white numbers i.e. when grey roughness numbers collapse into white roughness numbers.

For this purpose, a white roughness number was assigned to each pipe of the network (see Figure 7 where

the assigned pipe roughnesses are represented by the symbol \*) and two demand scenarios/conditions were considered. The former was represented by the mean nodal water demand (see Figure 3) and the latter by the peak water demand obtained by multiplying the mean water demand by the hourly peak coefficient 1.2. For each demand condition, heads at all the nodes of the water distribution systems were therefore computed through the hydraulic simulator (direct problem).

These head values were then considered as measurements at all the network nodes with reference to the two demand conditions (i.e.  $N_{dc_{mn}} = 2 \forall mn$  and  $N_{mn} = N_b$ ) and the inverse problem was solved by using the calibration procedure based on use of the hydraulic simulator. It is worth noting that, in this test, the pipes were not grouped into classes and thus for each pipe of the network a grey number (expected to collapse into a white number) was initially assumed and calibrated. Incidentally, the calibration is in this case over-determined since the number of grey roughness numbers to be calibrated is  $N_r = N_p = 12$ , whereas  $N_{mn} = N_b = 9$  measuring nodes and  $N_{dc_{mn}} = 2 \forall mn$  were considered, which implies a total of 18 available data (a similar situation is considered in Kapelan *et al.* 2007).





**Figure 7** | Synthetic case study: white roughness numbers assigned to each pipe (symbol \*) and the grey roughness numbers obtained by means of the calibration procedure (solid line with circles at the ends): each calibrated grey roughness number tends to collapse into a white number and thus the circles at the ends coincide.

As in the previous test, for the purpose of verifying the potential of the proposed procedure, the search for the grey roughness numbers was carried out over a wide domain  $D$  defined between  $d_1 = 50 \text{ m}^{1/3} \text{ s}^{-1}$  and  $d_2 = 120 \text{ m}^{1/3} \text{ s}^{-1}$  (for the meaning of the symbols  $d_1$  and  $d_2$  see the appendix). In this test case, use was made of the SCE-UA (Shuffled Complex Evolution–University of Arizona; Duan *et al.* 1992) algorithm to provide a first near optimal solution which was subsequently refined using the MATLAB<sup>TM</sup> *fmincon* function based on Sequential Quadratic Programming. The method of coupling a global with a local search optimizer is frequently adopted when several parameters/coefficients have to be optimized, as happens in this case (see Duan *et al.* 2003 for a general review of this approach). In fact, the global search method has the ability to avoid local optima but is usually weak in refining the search when the region of the global optimum is identified. The refinement ‘work’ can then be carried out by the local search method.

The calibrated grey roughness numbers for each pipe are shown in Figure 7 and compared to the assigned white

roughness numbers. In particular, a solid line with two circles at the extremes is used, as in the previous figures, to represent the grey numbers. However, in this case, each calibrated grey roughness number collapses into a white number and thus the two extremes coincide. In more detail, it can be observed that not only do the calibrated grey roughness numbers collapse into white numbers but that they also coincide (or at least approximate very well) with the assigned white roughness numbers. Small errors can be observed only for pipes 2 and 9, but it is worth noting that these two pipes are also characterized by very low velocities (around  $0.2 \text{ m s}^{-1}$ ) and small head losses (nearly  $0.45 \text{ m}$  on a length of  $1200 \text{ m}$ ). Moreover, the grey head numbers obtained (through the direct problem) on the basis of the calibrated grey (almost white) roughness numbers always contain the observed head values (symbol \*) for all the nodes of the water distribution system and for both the demand conditions and, at the same time, are very narrow (the amplitude is about  $1.5 \text{ mm}$ ). This means that the grey nodal heads for each demand condition also tend to collapse into white nodal heads which coincide with those observed.

### A real case study: calibration of the grey roughness values using hydraulic simulator

The real case study is represented by a completely separable part of the water distribution system of Ferrara (Italy). The total length of the pipe network is nearly 30 km and covers an area of approximately 5 km<sup>2</sup>. The population served is around 5,000 inhabitants and the demand type is mainly domestic. With respect to pipe materials, the pipe network consists mainly of asbestos cement pipes and, to a lesser degree, cast iron and steel pipes. With reference to Figure 4, node A is the feeding point of the network where the pressure and the inlet discharge are measured. In the hydraulic simulation model, this is the node where the head is set; the inlet discharge is used to characterize the nodal demand pattern. Nodes B, C and D are the pressure measuring nodes used for calibration purposes (see Figure 4). The available pressure and discharge data are relative to a time window of two days (48 hours).

The network is located near the main pumping station of Pontelagoscuro, a station equipped with variable speed pumps. The pressure patterns observed at the measuring

nodes (see Figure 8) clearly reflect the typical operational rules of the pumping station. According to these rules, the head is increased during the hours of maximum demand and reduced at night when demand, and consequently the head losses in the pipe network, is very small. Moreover, Figure 8 shows that the head pressure at nodes B and C oscillates around 30 m, whereas at node D the head pressure is around 24 m and the variations are smaller than those at the nodes B and C. This is due to a partially closed valve located near node D, but whose setting was not known exactly even after consulting the technicians of the agency managing the water distribution system.

The pattern of the nodal water demands is shown in Figure 9. This pattern was defined according to the discharges measured at node A. The water demands were then allocated at each node according to the number of users served. It is worth noting that this demand allocation procedure leads to nodal demands which are an estimate and not the real (unknown) nodal water demands, as always happens when real water distribution systems are considered.

To calibrate roughness, the pipes were grouped into three classes (Nr = 3): one for each pipe material. For each class

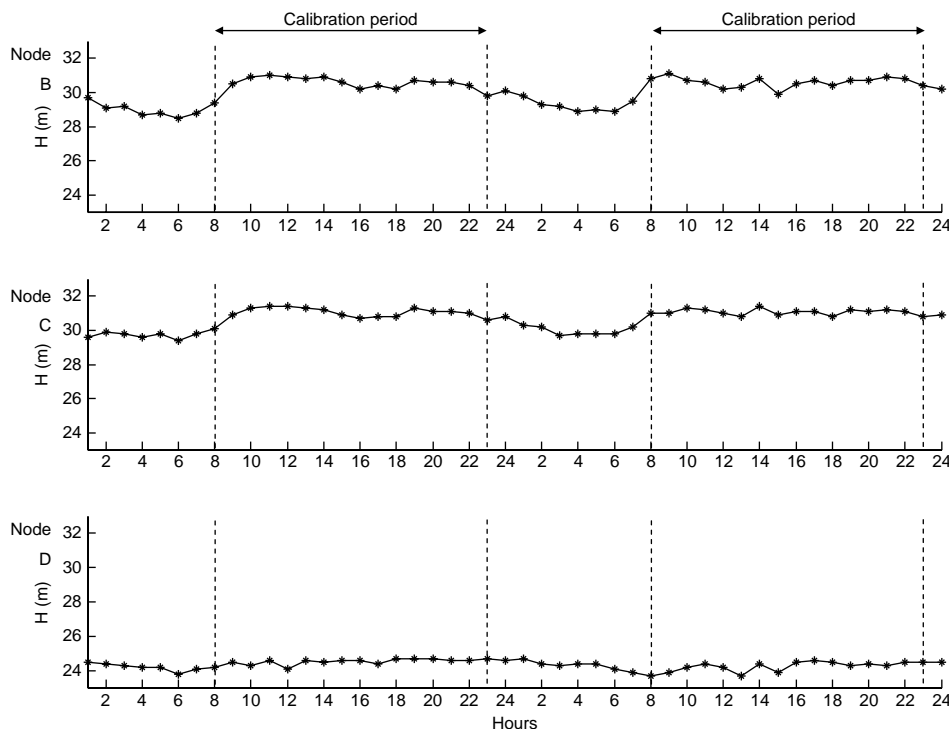


Figure 8 | Real case study: head pressure values observed at nodes B, C and D.

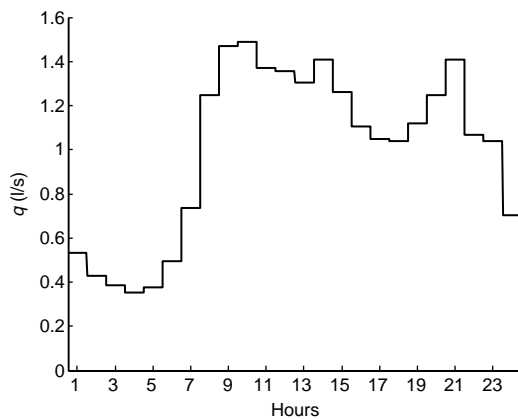


Figure 9 | Real case study: nodal demand pattern.

of pipes, the search for the grey roughness number was carried out over a wide domain defined between  $d_1 = 50 \text{ m}^{1/3} \text{ s}^{-1}$  and  $d_2 = 120 \text{ m}^{1/3} \text{ s}^{-1}$ . The heads observed at nodes B, C and D used for the calibrations are relative to the diurnal hours (08:00–23:00) of each of the two days (see Figure 8). This is the time window during which the nodal demands are higher than the daily average nodal demand (i.e. the hourly coefficients are higher than 1, see Figure 9). In other words,  $N_{dc} = 15$  demand conditions, one for each hour of the day between 08:00 and 23:00, were considered for each measuring node  $nm$  with  $mn = 1:(N_{mn} = 3)$ . Furthermore, for each demand condition, two measures were available: one for each of the two days considered and thus each set  $\Omega_{mn,dc}$  contains 2 observed heads.

The assumption of considering in the calibration phase only the head observed during the diurnal hours is due to the fact that, during the night hours, the pipe discharges, and consequently, the head losses, are very small and thus the nodal heads are almost completely insensitive to pipe roughnesses. Unfortunately, data relative to fire flow loading conditions, which are usually characterized by high pipe velocities, were not available.

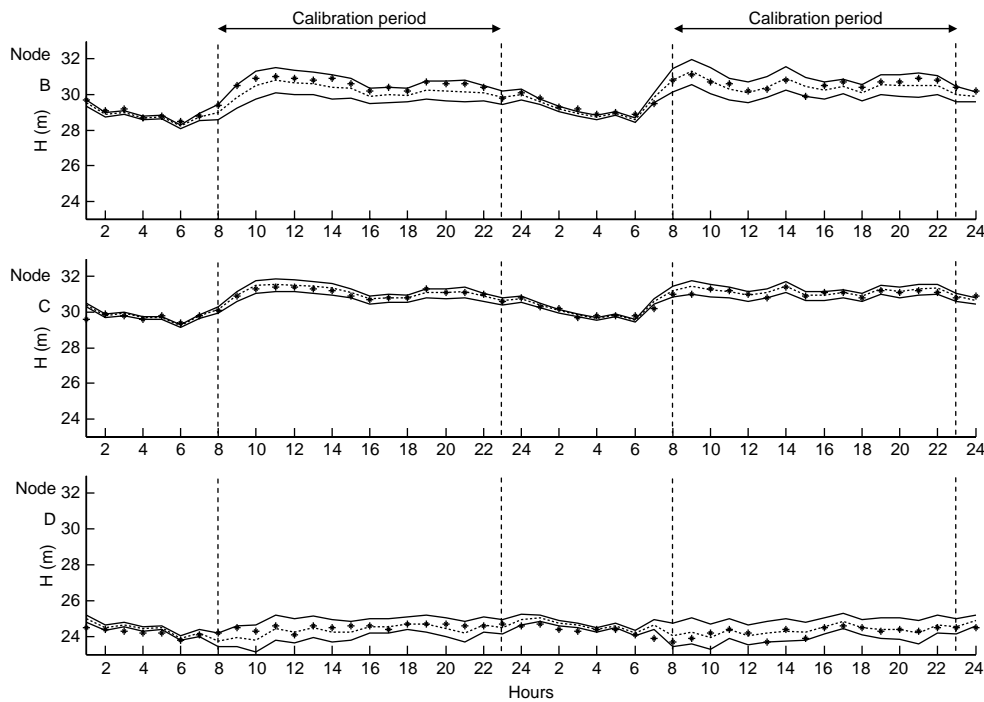
For the purposes of calibration, the procedure based on the use of the hydraulic simulator was adopted. The SCE-UA algorithm was used as optimizer. The grey roughness numbers obtained from the procedure are  $ks^\pm = [86,105]$  for asbestos cement pipes,  $ks^\pm = [78,87]$  for steel pipes and  $ks^\pm = [75,84]$  for cast iron pipes. These intervals are in agreement with those indicated in technical manuals for the corresponding materials (see e.g. Mays 2005), although it is

worth noting that their amplitude is also partially affected (amplified) by the uncertainty characterizing the nodal demand allocation and the setting of the valve located near to the node D.

Figure 10 shows the corresponding grey head numbers for nodes B, C and D. Consistently with the objective function used (see Equation (14)), the head values observed at nodes B, C and D and used during the calibration are always included in the band obtained by solving the direct problem according to the above-mentioned calibrated grey roughness numbers. On the contrary, the heads measured during the night hours (not used in the calibration phase) are sometimes not included in the band, even if the gap is only of a few centimetres (and in some cases even a few millimetres). However, the overall pattern of the heads during the night hours is well represented. Moreover, it can be observed that the amplitude of the band increases during the diurnal hours and reduces during the night hours. This is perfectly consistent with the different sensitivity of the nodal heads to the pipe roughness due to the different pipe velocities which can be observed during diurnal and night hours. It is worth noting that numerical tests (not reported in this paper) show that the amplitude of the band during the night hours would remain very similar to that plotted in Figure 10 even if very broad grey roughness numbers (e.g.  $ks^\pm = 50\text{--}120 \text{ m}^{1/3} \text{ s}^{-1}$ ) were considered: this confirms *a posteriori* the assumption of using, for the calibration purpose, only the heads observed during the diurnal hours.

In Figure 10 a dotted line is also shown for each node. This line represents the white head values obtained through a direct simulation using, for each pipe roughness class, the central point of the corresponding calibrated grey roughness number. With respect to these simulated white heads, the root mean square error is equal to 26 cm for node B, 19 cm for node C and 28 cm for node D. This aspect is important from a practical point of view. In fact, the grey roughness number allows an uncertainty interval to be defined around the nodal heads, whereas the corresponding central values lead to ‘central/reference’ head values which approximate the observed heads well, given the structure of the objective function used (see Equation (14)).

Finally, it is worth highlighting that the calibration procedure entails a fairly long computing time, requiring nearly 190,000 seconds to complete the calibration of



**Figure 10** | Real case study: grey head values (head band: solid lines) deriving from the calibration procedure based on heads observed between 08:00 and 23:00 at nodes B, C and D. The dotted line represents the white head values obtained through a direct simulation when reference to the central point of the calibrated grey roughness numbers is made. The available/measured head values are represented with the symbol \*.

this real water distribution system model. However, the computational time can be reduced by using a first-order Taylor series expansion in place of the hydraulic simulator, at the cost of a moderate degree of approximation/error in the grey roughness values (as shown in the following section).

### A real case study: calibration of the grey roughness values using the first-order Taylor series expansion

The calibration procedure based on use of the first-order Taylor series expansion was applied to the real case study described in the previous section. The same sets of observed data and pipe classes were considered, and the search for the grey roughness numbers was carried out over the domain defined between  $d_1 = 50 \text{ m}^{1/3} \text{ s}^{-1}$  and  $d_2 = 120 \text{ m}^{1/3} \text{ s}^{-1}$ . In order to apply the first-order Taylor series expansion, the coefficients  $a_{ji}$  (see Equation (11)) of the matrix  $A$  (see Equation (10)) were calculated on the basis of a reference value  $ks^*$  equal to  $85 \text{ m}^{1/3} \text{ s}^{-1}$  for all the pipe classes (the central value of the search domain defined between  $d_1 = 50 \text{ m}^{1/3} \text{ s}^{-1}$  and  $d_2 = 120 \text{ m}^{1/3} \text{ s}^{-1}$ ) and an amplitude  $\Delta ks = 10 \text{ m}^{1/3} \text{ s}^{-1}$ .

It is worth noting that the matrix  $A$  has to be calculated for each demand condition considered in the calibration process; its estimation was therefore repeated for each hour within the interval 08:00–23:00. The results of the calibration procedure based on the first-order Taylor series expansion are presented in Table 3 where it is shown that the grey roughness numbers for the three pipe classes ( $ks_{ic}^{\pm, \text{Taylor}}$  with  $ic = 1:(Nr = 3)$ , see column (a)) are similar to those obtained by using the procedure based on the hydraulic simulator ( $ks_{ic}^{\pm}$  with  $ic = 1:(Nr = 3)$ , see column (b)): the differences are of some percent units alone (between 0 and

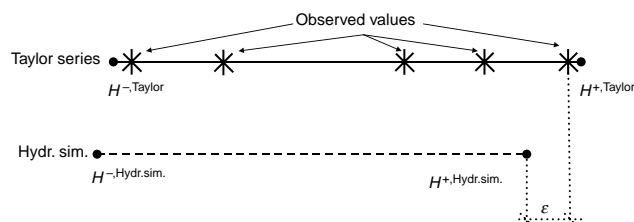
**Table 3** | Real case study: grey roughness numbers produced by the calibration procedure based on the first-order Taylor series expansion (column (a)) compared with those produced by the procedure based on the hydraulic simulator (column (b)) and the absolute percentage error on the lower (column (c)) and the upper (column (d)) extreme with respect to the grey numbers produced by the procedure based on the hydraulic simulator

	(a)	(b)	(c)	(d)
Pipe class	$ks_{ic}^{\pm, \text{Taylor}} (\text{m}^{1/3} \text{ s}^{-1})$	$ks_{ic}^{\pm} (\text{m}^{1/3} \text{ s}^{-1})$	(%)	(%)
Asbestos cement	[83, 100]	[86, 105]	3.5	4.8
Steel	[78, 94]	[78, 87]	0	8.0
Cast iron	[74, 83]	[75, 84]	1.3	1.2

8%, see columns (c) and (d)) with respect to the upper/lower limits of the grey roughness numbers obtained by using the procedure based on the hydraulic simulator (see column (b)).

It was then investigated what happens when the direct problem is solved with the hydraulic simulator coupled with the grey roughness numbers  $ks_{ic}^{\pm, \text{Taylor}}$  with  $ic = 1:(Nr = 3)$ . It is evident that the grey head numbers at the nodes of interest are slightly different from those obtained in the calibration phase, thus leading to the possibility that several observed heads are no longer contained within the grey numbers of the (hydraulically) simulated heads. In order to quantify this eventuality, we introduced the error  $\varepsilon$ . With reference to Figure 11, the solid line represents the grey number of the head at a generic node obtained by means of the calibration procedure based on the Taylor series expansion. This grey number obviously includes the observed head values, given the type of objective function used in the calibration phase. The broken line indicates the grey head number computed *a posteriori* through the direct simulation based on the hydraulic simulator and using the grey roughness numbers  $ks_{ic}^{\pm, \text{Taylor}}$  ( $ic = 1:Nr$ ). The error  $\varepsilon$  is defined as the maximum absolute value of the difference between the observed value (larger/smaller) not included within the grey (head) number and the extreme of the grey (head) number nearest to it. The latter is obtained, as already mentioned, by means of the hydraulic simulator coupled with  $ks_{ic}^{\pm, \text{Taylor}}$  ( $ic = 1:Nr$ ) (see Figure 11).

On the basis of this definition, Table 4 shows the errors  $\varepsilon$  across all the nodes and demand conditions. The errors (in metres) are extremely small, with the mean error equal to 8 cm and the maximum error equal to 16 cm (see node C at 19:00).



**Figure 11** | The solid line represents the grey head number of the generic node of observation following the calibration of the grey roughness numbers, where the calibration was performed using the Taylor series expansion in place of the hydraulic simulation model. The symbol \* represents the observed head values. The broken line represents the grey head number at the same node of observation obtained through direct simulation using the hydraulic simulator and the grey roughness numbers obtained by calibration performed using the Taylor series expansion.

**Table 4** | Real case study: approximation errors  $\varepsilon$  (m) for the procedures based on the first-order Taylor series expansion for the different nodes at different hours of the day (i.e. for the different demand conditions dc); the approximation error  $\varepsilon$  is zero for the hours not shown in the table

Node	dc (hour)	$\varepsilon$ (m)
B	08:00	0.03
B	09:00	0.05
B	15:00	0.12
B	17:00	0.06
B	19:00	0.04
C	00:00	0.03
C	16:00	0.11
C	19:00	0.16
C	22:00	0.08
D	08:00	0.12
D	09:00	0.10
Mean error		0.08

To summarize, the calibration procedure based on the first-order Taylor series expansion leads to grey roughness numbers which differ from those obtained by using the hydraulic simulator of some percent units and, at the same time, leads to errors on the heads which are extremely small. There is, moreover, a significant difference in the computing time, which is greatly reduced when the Taylor series expansion is used. In fact, the use of the Taylor series expansion allows for the reduction of the computing time from nearly 190,000 s to nearly 1,000 s.

Note that (because of the computational benefits produced by the first-order Taylor series expansion) even when the calibration problem is solved using the hydraulic simulator, the process can be accelerated by initializing it with the approximated solution obtained through the use of the first-order Taylor series expansion. We have in fact verified that when this initialization is used, the computational time reduces from nearly 190,000 s to nearly 44,000 s while the results are exactly the same.

## FINAL REMARKS AND CONCLUSIONS

Grey numbers are a mathematical tool for representing the uncertainty associated with the roughness of pipes, which varies according to material and date of installation and cannot actually be measured in the context of a real water



distribution system. The calibration procedure proposed, based on grey numbers, is designed to provide an estimate of pipe roughness taking into account the factor of uncertainty.

The two synthetic cases described show that the calibration procedure can correctly solve the inverse problem. In fact, the first synthetic case shows that known grey roughness numbers, even when they overlap, are always correctly identified when the calibration problem is over-determined. When the number of measuring nodes is smaller than the number of the grey roughnesses to be calibrated, the solution of the calibration problem respects (as expected) the constraints related to the objective function but the values of the grey roughness numbers do not reflect the true values in the network (under-determined problem). When the grey roughness numbers overlap (which is certainly a more complex problem than when they do not overlap), the selection of the nodes (the global number of nodes selected being constant) is very important since it can negatively affect the solution of the calibration problem, thus suggesting that a proper sampling design is necessary to correctly set up the procedure.

The second synthetic case represents further proof of the procedure to correctly solve the inverse problem since it is shown that even (known) white roughness numbers can be identified, i.e. the calibration procedure is able to identify grey roughness numbers which can collapse into single values representing the (known) white roughness numbers.

The numerical example relative to the real case allows for the identification of the grey roughness numbers relative to different materials used in the network. These grey roughness numbers produce grey head numbers that contain all the head values considered in the calibration process and, at the same time, the band around them is as narrow as possible. Furthermore, when reference to the central point of these grey roughness numbers is made, the white head values that can be obtained by means of a direct simulation well respect the observed values: this result certainly has an interesting impact on practical applications since it is then possible with this procedure to define a white roughness “reference set” around which it is then possible to perform a model uncertainty prediction through the use of the grey roughness numbers and relative mathematics.

In this numerical example a marked insensitivity of the night heads to the roughness, due to the very low water

velocity, is also shown. The combination of this circumstance with the imperfect knowledge of the nodal demand produces a situation where it is impossible to identify grey roughness numbers of reasonable amplitude producing grey head numbers containing the heads measured during the night period (and in fact, calibration was performed by using heads observed during a period of higher water velocity in the pipes). In other terms, when the demand scenarios are not selected properly in a real system, no (reasonable) solutions may exist in terms of grey roughness numbers that allow for the calculation of grey head numbers containing the observed values. Indeed, this is not the only case. Such situations can also occur when either the model is very poor (bad description of the network layout, badly defined diameters and/or lengths), information on pump and/or valve setting is incomplete or wrong or measurements are wrong or of poor quality. However, all these situations are expected to be solved (when possible) before applying a calibration procedure, independently of the technique used. Otherwise, the calibration process would tend to compensate errors in data and problem description by forcing the parameters out of any feasible range in order to reproduce observed data.

The analyses conducted have moreover shown that, within the framework of the calibration and simulation procedure, the Taylor series expansion represents a valid alternative to the hydraulic simulator. It speeds up computing times significantly at the cost of only very modest approximation errors. Because of the computational benefits that can be obtained by using the first-order Taylor series expansion (even when the calibration problem is solved by means of the hydraulic simulator), the process can be accelerated by initializing it with the approximated solution obtained with the use of the first-order Taylor series expansion as shown in the numerical example relative to the real case.

Finally, it should be highlighted that the procedure illustrated here with reference to the calibration of pipe roughness based solely on nodal head measurements can be extended to the calibration of other parameters, e.g. water demand coefficients. Other types of field measurements such as pipe discharges can also be considered. These aspects are the subject of ongoing research and results will be presented in due course.

## ACKNOWLEDGEMENTS

The current research is funded by the project ENVIREN. The authors are grateful to the three anonymous reviewers for their helpful comments and criticisms which greatly helped us to improve the quality of the paper.

## REFERENCES

- Alefeld, G. & Herzberger, J. 1983 *Introduction to Interval Computations*. Academic Press, London.
- Bargiela, A. & Hainsworth, G. D. 1989 *Pressure and flow uncertainty in water systems*. *J. Water Resour. Plann. Manage.* **115** (2), 212–229.
- Bardossy, A., Mascellani, G. & Franchini, M. 2006 Fuzzy unit hydrograph. *Water Resour. Res.* **42** (2), W02401.
- Beven, K. 2001 *Rainfall–Runoff Modelling, The Primer*. John Wiley & Sons, Ltd, Chichester, UK.
- Beven, K. & Binley, A. 1992 *The future of distributed models: model calibration and uncertainty prediction*. *Hydrol. Process.* **6**, 279–298.
- Bhave, P. R. 1988 *Calibrating water distribution network models*. *J. Environ. Eng.* **114** (1), 120–136.
- Branisavljević, N. & Ivetić, M. 2006 *Fuzzy approach in the uncertainty analysis of the water distribution system of Becej*. *Civil Eng. Environ. Syst.* **23** (3), 221–236.
- Branisavljević, N., Prodanović, D. & Ivetić, M. 2009 *Uncertainty reduction in water distribution network modelling using inflow data*. *Urban Water J.* **6** (1), 69–79.
- Bush, C. A. & Uber, J. G. 1998 *Sampling design methods for water distribution model calibration*. *J. Water Resour. Plann. Manage.* **124** (6), 334–344.
- Cheng, F. J., Hui, S. H. & Chen, Y. C. 2002 *Reservoir operation using grey fuzzy stochastic dynamic programming*. *Hydrol. Process.* **16**, 2395–2408.
- Cross, H. 1936 *Analysis of flow in networks of conduits or conductor*. *Eng. Exp. Stat. Bull.* **286**, 1–32.
- Cullen, A. C. & Frey, H. C. 1999 *Probabilistic Techniques in Exposure Assessment, A Handbook for Dealing with Variability and Uncertainty in Models and Inputs*. Society of Risk Analysis, Plenum Press, New York, USA.
- Cullmann, J. & Wriedt, G. 2008 *Joint application of event-based calibration and dynamic identifiability analysis in rainfall-runoff modelling: implications for model parametrisation*. *J. Hydroinformatics* **10** (4), 301–316.
- Datta, R. S. N. & Sridharan, K. 1994 *Parameter estimation in water-distribution systems by least-squares*. *J. Water Resour. Plann. Manage.* **120** (4), 405–422.
- Deng, J. L. 1982 *Control problems of grey systems*. *Syst. Control Lett.* **1** (5), 288–294.
- Deng, J. L. 1987 *Basic Methodology of Grey System*. Publishing House of Huazhong, University of Science and Technology, Wuhan.
- Dettinger, M. D. & Wilson, J. L. 1981 *First order analysis of uncertainty in numerical models for groundwater flows*. *Water Resour. Res.* **17** (1), 149–161.
- Di Cristo, C. & Leopardi, A. 2003 *Influenza dell'incertezza della domanda sulla calibrazione e simulazione delle reti idriche*. *Atti del XXIX Convegno Nazionale di Idraulica e Costruzioni Idrauliche*, Trento, Ed. Bios (in Italian).
- Diamond, P. 1988 *Fuzzy least squares*. *Inform. Sci.* **46**, 141–157.
- Duan, Q., Sorooshian, S. & Gupta, V. K. 1992 *Effective and efficient global optimization for conceptual rainfall runoff models*. *Water Resour. Res.* **24** (7), 1163–1173.
- Duan, Q., Gupta, V. K., Sorooshian, S., Rousseau, A. N. & Turcotte, R. (eds.) 2003 *Calibration of watershed models*. *Water Science and Application 6*, American Geophysical Union, Washington, USA.
- Fang, T. & Ball, J. E. 2007 *Evaluation of spatially variable control parameters in a complex catchment modelling system: a genetic algorithm application*. *J. Hydroinformatics* **9** (3), 163–174.
- Gargano, R. & Pianese, D. 2000 *Reliability as tool for hydraulic network planning*. *J. Hydraul. Eng.* **126** (5), 354–364.
- Gauckler, P. G. 1868 *Du mouvement de l'eau dans les conduits*. *Annales de Ponts et Chaussées* **15** (4), 229–281.
- Huang, G. H., Baetz, B. W. & Patry, G. G. 1995 *Grey integer programming: an application to waste management planning under uncertainty*. *Eur. J. Operat. Res.* **83**, 594–620.
- Kapelan, Z., Savic, D. A. & Walters, G. A. 2003 *Multiobjective sampling design for water distribution model calibration*. *J. Water Resour. Plann. Manage.* **129** (6), 466–479.
- Kapelan, Z., Savic, D. A. & Walters, G. A. 2005a *Multiobjective rehabilitation of water distribution systems under uncertainty*. *Water Resour. Res.* **41** (11), W11407.
- Kapelan, Z., Savic, D. A., Walters, G. A. & Babayan, A. V. 2005b *Risk and robustness based solutions to a multiobjective water distribution system rehabilitation problem under uncertainty*. *Water Sci. Technol.* **53** (1), 61–75.
- Kapelan, Z., Savic, D. A. & Walters, G. A. 2007 *Calibration of water distribution hydraulic models using a Bayesian-type procedure*. *J. Hydraul. Eng.* **133** (8), 927–936.
- Karmakar, S. & Mujumdar, P. P. 2006 *Grey fuzzy optimization for water quality management of a river system*. *Adv. Water Resour.* **29**, 1088–1105.
- Kaufmann, A. & Gupta, M. M. 1985 *Introduction to Fuzzy Arithmetic: Theory and Applications*. Van Nostrand, Reinhold, New York.
- Kottogoda, N. T. & Rosso, R. 1998 *Statistics, Probability and Reliability for Civil and Environmental Engineers*. McGraw-Hill, New York.
- Lansley, K. E. & Basnet, C. 1991 *Parameter estimation for water distribution networks*. *J. Water Resour. Plann. Manage.* **117** (1), 126–144.
- Lansley, K. E., El-Shorbagy, W., Ahmed, I., Araujo, J. & Haan, C. T. 2001 *Calibration assessment and data collection for water distribution networks*. *J. Hydraul. Eng.* **127** (4), 270–279.

- Liu, S. & Lin, Y. 1998 *An Introduction to Grey Systems: Foundation, Methodology and Applications*. IIGSS Academic Publisher, PA.
- Liu, S., Gao, T. & Dang, Y. 2000 *Grey Systems Theory and its Applications*. The Science Press of China, Beijing.
- Mallick, K. N., Ahmed, I., Tickle, K. S. & Lansey, K. E. 2002 **Determining pipe grouping for water distribution networks**. *J. Water Resour. Plann. Manage.* **128** (2), 130–139.
- Marchi, E. & Rubatta, A. 1981 *Fluids Mechanics*. UTET, Torino, Italy (in Italian).
- Maskey, S. & Guinot, V. 2003 **Improved first-order second moment method for uncertainty estimation in flood forecasting**. *Hydrol. Sci. J.* **48** (2), 183–196.
- Mays, L. W. 2005 *Water Resources Engineering*. John Wiley & Sons, Inc., NJ.
- McKay, M. D., Conover, W. J. & Beckman, R. J. 1979 **A comparison of three methods for selecting values of input variables in the analysis of output from a computer code**. *Technometrics* **211**, 239–245.
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N. & Teller, A. H. 1953 **Equations of state calculations by fast computing machines**. *J. Chem. Phys.* **21**, 1087–1091.
- Ormsbee, L. E. & Wood, D. J. 1986 **Explicit pipe network calibration**. *J. Water Resour. Plann. Manage.* **112** (1), 166–182.
- Ozelkan, E. C. & Duckstein, L. 2001 **Fuzzy conceptual rainfall-runoff models**. *J. Hydrol.* **253**, 41–68.
- Powell, M. J. D. 1978 **A fast algorithm for nonlinearly constrained optimization calculations**. In *Numerical Analysis* (ed. G. A. Watson). *Lecture Notes in Mathematics*, (Vol. 630). Springer Verlag, Germany.
- Powell, M. J. D. 1983 *Variable Metric Methods for Constrained Optimization, Mathematical Programming: The State of the Art* (ed. A. Bechem, M. Grotschel & B. Korte), pp. 288–311. Springer Verlag, Germany.
- Reddy, P. V. N., Sridharan, K. & Rao, P. V. 1996 **WLS method for parameter estimation in water distribution networks**. *J. Water Resour. Plann. Manage.* **122** (3), 157–164.
- Revelli, R. & Ridolfi, L. 2002 **Fuzzy approach for analysis of pipe networks**. *J. Hydraul. Eng.* **128** (1), 93–101.
- Rossman, L. A. 2000 *EPANET 2 Users Manual*. National Risk Management Research Laboratory, Office of Research and Development. U.S. Environmental Protection Agency, Cincinnati, Ohio, USA.
- Savic, D. A., Kapelan, Z. & Jonkergouw, P. 2009 **Quo vadis water distribution model calibration?** *Urban Water J.* **6** (1), 3–22.
- Schittowski, K. 1985 **NLQPL: a Fortran-subroutine solving constrained nonlinear programming problems**. *Operat. Res.* **5**, 485–500.
- Shamir, U. & Howard, C. 1968 **Water distribution systems analysis**. *J. Hydraul. Div.* **94**, 219–234.
- Skahill, B. E. & Doherty, J. 2006 **Efficient accommodation of local minima in watershed model calibration**. *J. Hydrol.* **329**, 122–139.
- Tanaka, H. 1987 **Fuzzy data analysis by possibilistic linear models**. *Fuzzy Sets Syst.* **24**, 363–375.
- Tanaka, H., Uejima, S. & Assai, K. 1982 **Linear regression analysis with fuzzy model**. *IEEE Trans. Syst. Man Cybern.* **12** (6), 903–907.
- Todini, E. 1999 **Using a Kalman filter approach for looped water distribution network calibration**. In *Proceedings of Water Industry Systems: Modelling and Optimization Applications* (ed. D. A. Savic & G. A. Walters), (Vol. 1), pp. 327–336 Exeter, UK.
- Todini, E. & Pilati, S. 1987 **A gradient algorithm for the analysis of pipe network**. In *International conference on Computer Applications for Water Supply and Distribution*. Leicester Polytechnic, UK.
- Tolson, B. A., Maier, H. R., Simpson, A. R. & Lence, B. J. 2004 **Genetic algorithms for reliability-based optimization of water distribution systems**. *J. Water Resour. Plann. Manage.* **130** (1), 63–72.
- van Griensven, A. & Meixner, T. A. 2007 **Global and efficient multi-objective auto-calibration and uncertainty estimation method for water quality catchment models**. *J. Hydroinformatics* **9** (4), 277–291.
- Wagener, T., McIntyre, N., Lees, M. J., Wheater, H. S. & Gupta, H. V. 2003 **Towards reduced uncertainty in conceptual rainfall-runoff modelling: dynamic identifiability analysis**. *Hydrol. Process.* **17**, 455–476.
- Wagener, T., Wheater, H. S. & Gupta, H. V. 2004 *Rainfall-Runoff Modelling in Gauged and Ungauged Catchments*. Imperial College Press, London.
- Walski, T. M. 1983 **Technique for calibrating networks models**. *J. Water Resour. Plann. Manage.* **109** (4), 360–372.
- Wang, Q. & Wu, H. 1998 **The concept of grey number and its property**. In *Proceedings of NAFIPS*, pp. 45–49 USA.
- Wu, Q., Zhou, W., Li, S. & Wu, X. 2005 **Application of grey numerical model to groundwater resource evaluation**. *Environ. Geol.* **47**, 991–999.
- Xu, C. & Goulter, I. 1996 **Uncertainty analysis of water distribution networks**. In *Proceedings of Stochastic Hydraulics '96* (ed. K. S. Tickle, I. C. Goulter, C. Xu, S. A. Wasimi & F. Bouchart), pp. 609–616. Balkema, Rotterdam.
- Xu, C. & Goulter, I. 1997 **Reliability assessment of water distribution networks using the first order reliability method**. In *Managing Water: Coping with Scarcity and Abundance, Proc. 27th IAHR Congr.* (ed. M. English & A. Szollosi-Nagy), pp. 429–434. ASCE.
- Xu, C. & Goulter, I. C. 1998 **Probabilistic model for water distribution reliability**. *J. Water Resour. Plann. Manage.* **124** (4), 218–228.
- Yang, Y., John, R. & Chiclana, F. 2004 **Grey sets: a unified model for fuzzy sets and rough sets**. In *Proceedings of the 2004 UK Workshop on Computational Intelligence*, pp. 239–246 Loughborough, UK.
- Zadeh, L. A. 1965 **Fuzzy sets**. *Inform. Control* **8**, 338–353.

## APPENDIX

In the calibration procedure shown in Figure 2, the decision variables of the external optimization process are the  $N_r$  grey numbers  $ks_{ic}^{\pm}$  (with  $ic = 1:N_r$ ) representing the roughnesses of the  $N_r$  classes of pipes. With reference to Figure A1, the generic decision variable i.e. the generic grey number  $ks_{ic}^{\pm}$  is generated in the following manner. The lower extreme  $d_1$  and upper extreme  $d_2$  of the domain  $D$  over which the grey number will be generated are assumed to be assigned *a priori*. Since the generic grey number  $ks_{ic}^{\pm}$  that we want to generate is characterized by two values (the lower extreme  $ks_{ic}^{-}$  and upper extreme  $ks_{ic}^{+}$ ), two real numbers are generated,  $p_1$  and  $p_2$ , falling between 0 and 1. The lower extreme  $ks_{ic}^{-}$  is given by (see Figure A1):

$$ks_{ic}^{-} = d_1 + p_1 \cdot (d_2 - d_1) \quad (16)$$

In other words,  $p_1$  provides the position of the lower extreme of the grey number being generated in percentage terms, relative to the amplitude of the domain  $D$ .

The upper extreme  $ks_{ic}^{+}$  is given by (see Figure A1):

$$ks_{ic}^{-} = d_1 + p_1 \cdot (d_2 - d_1) \quad ks_{ic}^{+} = ks_{ic}^{-} + p_2 \cdot (d_2 - ks_{ic}^{-})$$

**Figure A1** | External optimization process: generation of the grey roughness number within a closed interval defined by  $d_1$  and  $d_2$ .

$$ks_{ic}^{+} = ks_{ic}^{-} + p_2 \cdot (d_2 - ks_{ic}^{-}) \quad (17)$$

It should however be pointed out that, since the real numbers  $p_1$  and  $p_2$  are generated between 0 and 1, it is possible that for certain values of  $p_1$  or  $p_2$  the grey number may collapse into a white number. In fact, if  $p_1$  were to assume a value of 1,  $ks_{ic}^{-}$  would coincide with the upper extreme  $d_2$  of the domain  $D$  over which the grey number is being generated (see Equation (16)); consequently, as  $d_2 - ks_{ic}^{-} = 0$ ,  $ks_{ic}^{-}$  would also coincide with  $ks_{ic}^{+}$  irrespective of the value of  $p_2$  (see Equation (17)). Similarly, irrespective of the value of  $p_1$ , if  $p_2$  were to assume a value of 0,  $ks_{ic}^{+}$  would coincide with  $ks_{ic}^{-}$  (see Equation (17)).

This is not a problem in itself since, when several grey roughness numbers are calibrated simultaneously, the fact that one (or more than one) collapses into a white number does not affect the grey nature of the heads simulated at the nodes which is guaranteed by the presence of at least one grey roughness number. At the same time, the white roughness number(s) obtained through the calibration would indicate that, for a given class of pipes, the roughness coefficient(s) has (have) to be restricted on a definitely narrow interval. However, if we would like to ensure in any event that a minimum amplitude for the grey number is generated, the real number  $p_1$  could be generated between 0 and an upper extreme  $p_1^{\max} < 1$ , whereas  $p_2$  could be generated between a lower extreme  $p_2^{\min} > 0$  and 1 (for instance the upper extreme  $p_1^{\max}$  for  $p_1$  could be set equal to 0.95 while a lower extreme  $p_2^{\min} = 0.05$  could be set for  $p_2$ ). Incidentally, these latter positions would not reduce the amplitude of the search domain since  $ks_{ic}^{-}$  can coincide with  $d_1$  where  $p_1 = 0$  and  $ks_{ic}^{+}$  can coincide with  $d_2$  where  $p_2 = 1$ .