# Crisp discharge forecasts and grey uncertainty bands using data-driven models

S. Alvisi, E. Creaco and M. Franchini

## ABSTRACT

A data-driven artificial neural network (ANN) model and a data-driven evolutionary polynomial regression (EPR) model are here used to set up two real-time crisp discharge forecasting models whose crisp parameters are estimated through the least-square criterion. In order to represent the total uncertainty of each model in performing the forecast, their parameters are then considered as grey numbers. Comparison of the results obtained through the application of the two models to a real case study shows that the crisp models based on ANN and EPR provide similar accuracy for short forecasting lead times; for long forecasting lead times, the performance of the EPR model deteriorates with respect to that of the ANN model. As regards the uncertainty bands produced by the grey formulation of the two data-driven models, it is shown that, in the ANN case, these bands are on average narrower than those obtained by using a standard technique such as the Box–Cox transformation of the errors; in the EPR case, these bands are on average larger. These results therefore suggest that the performance of a grey data-driven model depends on its inner structure and that, for the specific models here considered, the ANN is to be preferred.

**Key words** | artificial neural networks, discharge forecasting, evolutionary polynomial regression, grey numbers, uncertainty

**S. Alvisi** (corresponding author)
**E. Creaco**
**M. Franchini**
Department of Engineering,
University of Ferrara,
Via Saragat, 1,
44122, Ferrara,
Italy
E-mail: *stefano.alvisi@unife.it*

## INTRODUCTION

Hydrologic models used for real-time forecasting of discharges (and/or levels) at the basin outlet or the end of a river reach typically have a deterministic/crisp nature (Krzysztofowicz 1999). Their structure may contain different levels of physical information (e.g. Todini 1988) or be set up on the basis of the so-called data-driven techniques (e.g. ASCE Task Committee on the Application of Artificial Neural Networks in Hydrology 2000). However, the lack of any quantification of the uncertainty associated with the forecast values may constitute a factor that limits the use of the forecasting model in practical applications aimed at managing water resources and flood protection (Todini 2004, 2007). In actual fact, a forecasting model that does not take into account the uncertainty associated with the forecast may cause the user of the model, or the decision maker, to have excessive confidence in the forecast value; should the latter prove to be incorrect, this may lead to even worse consequences than if no forecasting model had been available (Montanari & Grossi 2008). For this reason, the quantification of uncertainty in hydrological models has recently been investigated in numerous research studies. Several methods aimed at estimating the uncertainty band to be linked to a generic hydrologic forecasting model have been proposed (e.g. Krzysztofowicz 1999; Todini 2007, 2008). In general, these methods can be grouped into different categories (see Solomatine & Shrestha 2009 for a general classification) such as analytical methods (e.g. Tung 1996), simulation and sampling-based methods (e.g. Kuczera & Parent 1998), Bayesian methods (e.g. Krzysztofowicz 1999) or methods based on the analysis of model errors (e.g. Montanari & Grossi 2008).

Some of these methods (e.g. simulation and sampling-based) require assumptions to be made on the parameter probability distributions, while most of the approaches

doi: 10.2166/nh.2012.121

based on the analysis of model errors require assumptions to be made regarding the residuals (e.g. normality and homoscedasticity). More recently, a new category of models for the estimation of the uncertainty band of a generic forecasting model is gaining popularity: the category of models based on data-driven techniques (Shrestha & Solomatine 2008). Within this context, Shrestha & Solomatine (2008) propose the 'uncertainty estimation based on local errors and clustering' (UNEEC) method based on the use of clustering and machine-learning techniques (e.g. M5 model trees) to estimate the uncertainty band of a generic forecasting model. Alvisi & Franchini (2011) propose an alternative method based on fuzzy neural networks to set up a hydrologic forecasting model which can produce crisp/deterministic forecasts and uncertainty bands, both summarized through fuzzy numbers.

Taking our cue from the work of Alvisi & Franchini (2011), in this paper we propose a further and different method for estimating the uncertainty band of two different data-driven forecasting models. This method is based on the grey number theory (Deng 1982) which enables the representation of the uncertainty through an interval. This interval is indeed a grey number and is expected to represent the total uncertainty of the model in making a forecast with a given lead time.

Grey theory is used here since, like fuzzy theory, it represents an appropriate tool for modelling uncertainties that do not originate from randomness but are caused by imprecise (or incomplete) knowledge about a real system (Liu & Lin 2006; Jacquin & Shamseldin 2007). When developing a discharge forecasting model, the representation of the physical process (the rainfall–runoff transformation and/or the flood propagation) is generally incomplete, particularly when a data-driven approach is used. It is this imprecise representation of the process which is the major source of uncertainty in the forecasting phase; grey number theory can therefore be used as a valid tool for modelling the related uncertainty/imprecision.

The proposed method is applied to two data-driven forecasting models characterized by completely different structures in order to highlight how the data-driven model structure can affect the results of the 'greyification' process. The first model is based on artificial neural networks (ANNs), which represent one of the most widely used data-driven techniques to develop hydrological forecasting models (ASCE Task Committee on the Application of Artificial Neural Networks in Hydrology 2000; Maier & Dandy 2000). The benefits derived from the use of neural networks to develop hydrological forecasting models are mainly connected to their intrinsic ability to reproduce properly the generally strongly non-linear relationships among hydrological variables, also in cases where such relationships are unknown or cannot be made explicit *a priori*.

The second model is based on the evolutionary polynomial regression (EPR) technique (Giustolisi & Savic 2006). EPR enables model identification and parameterization to be performed automatically and simultaneously. This technique has been developed significantly in recent years (Giustolisi & Savic 2009) and has been used in different areas of civil and environmental engineering (e.g. Giustolisi *et al.* 2008; El-Baroudy *et al.* 2010). Berardi *et al.* (2010) stress that the main features that make EPR appealing for practical data-driven modelling can be summarized as: (1) intuitive introduction of existing prior knowledge on the phenomenon considered, if any; and (2) possibility of selecting the most effective model according to final modelling purpose.

In the sections that follow, the set-up of the two hydrologic forecasting models at deterministic/crisp level is initially presented. Some basic information about grey numbers is then given, and finally the technique for parameterizing the ANN and the EPR models in grey mode is described. The numerical example refers to the problem of discharge forecasting at the downstream end of a watercourse on the basis of the knowledge of the discharges measured in a section further upstream and assuming that detailed surveys of the river reach and information about lateral inflows are not available. Analysis of the crisp forecasts and the uncertainty bands, produced by the crisp and grey formulation of the two models, respectively, is performed; a comparison with a standard Box–Cox technique (Box & Cox 1974) for defining the uncertainty bands is also developed.

## THE GREY ANN AND EPR FORECASTING MODELS

A description of the two data-driven hydrological forecasting models based on ANNs and EPR is now given. The subsections below first provide a brief description of their

structure and, subsequently, a clear presentation of the grey number theory-based method to characterize the total forecast uncertainty of the models.

## The artificial neural network (ANN)

Different neural network structures exist but one of the most common and widely used is the multilayer feed-forward perceptron (MFFP) network (Rosenblatt 1958) adopted as a reference in this study. A network with a three-layer architecture is considered in particular, comprising an input, a hidden and an output layer made up respectively of $n_p$, $n_h$ and $n_o$ neurons (see Figure 1).

In a neural network of this type, the generic output $o_i$ ($i = 1{:}n_o$) is given by:

$$o_i = f_{out}\left(\sum_{j=1}^{n_h} w_{i,j}\left\{f_{hid}\left[\sum_{k=1}^{n_p}(u_{j,k}\,p_k) + a_j\right]\right\} + b_i\right) \quad i = 1, \ldots, n_o \tag{1}$$

where $p_k$ ($k = 1{:}n_p$) represents the $k$th element of the input vector $\boldsymbol{p}$; $w_{i,j}$ ($i = 1{:}n_o$ and $j = 1{:}n_h$) represents the element of the weight matrix $\boldsymbol{W}$ $[n_o \times n_h]$ that links the $i$th output to the $j$th neuron of the hidden layer; $u_{j,k}$ ($j = 1{:}n_h$ and $k = 1{:}n_p$) represents the element of the weight matrix $\boldsymbol{U}$ $[n_h \times n_p]$ that links the $j$th neuron of the hidden layer to the $k$th input; $a_j$ ($j = 1{:}n_h$) represents the $j$th element of the bias vector $\boldsymbol{a}$; $b_i$ ($i = 1{:}n_o$) represents the $i$th element of the bias vector $\boldsymbol{b}$; and $f_{hid}$ and $f_{out}$ represent the transfer functions in the hidden layer and in the output layer, respectively. The transfer functions $f_{hid}$ and $f_{out}$ used in this study are the hyperbolic tangent sigmoid transfer function:

$$f_{hid}(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{2}$$

and the linear function:

$$f_{out}(x) = x \tag{3}$$

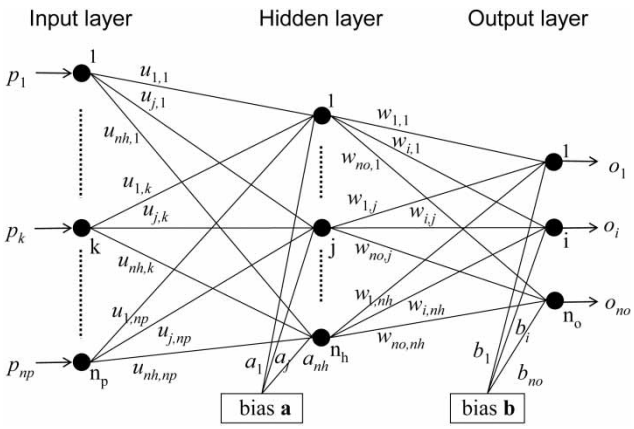


**Figure 1** | Architecture of a multilayer feed-forward neural network.

## The evolutionary polynomial regression (EPR)

EPR is a data-driven modelling technique that integrates the best features of numerical regression with genetic programming and symbolic regression (Giustolisi & Savic 2006). The general expression of the EPR formula is:

$$o = \sum_{i=1}^{n_t} F(\mathbf{p}, f(\mathbf{p}), c_i) + c_0 \tag{4}$$

where $o$ is the estimated output of the system/process, $c_i$ is a constant value, $F$ is a function constructed by the process, $\mathbf{p}$ is the vector of input variables (made up of $n_p$ elements), $f$ is a function defined by the user (such as natural logarithm, exponential, etc.), and $n_t$ is the length (number of terms) of the polynomial (bias $c_0$ excluded, if any; Giustolisi & Savic 2006).

Given this general formula, EPR allows several pseudo-polynomial expressions belonging to the class of Equation (4) to be used (Giustolisi & Savic 2006). The expression adopted in this study is simple and avoids the use of trigonometric functions (since the output is not characterized by periodicity) or the natural logarithm (since the output must always be positive):

$$o = c_0 + \sum_{i=1}^{n_t} c_i p_1^{\gamma_{i,1}} p_2^{\gamma_{i,2}} \ldots p_{np}^{\gamma_{i,np}} \tag{5}$$

The form of Equation (5) actually used in the numerical application (see 'Analysis and discussion of the results') is identified through an evolutionary searching strategy which enables the simultaneous selection (from a predefined set) of the input variables to be inserted in each monomial of the summation and the corresponding

exponents, whereas the coefficients $c_0$ and $c_i$ (with $i = 1{:}n_t$) are computed through a numerical regression based on the least-square method (Giustolisi & Savic 2006).

## The grey numbers

A grey number (Deng 1982, 1987; Liu & Lin 2006) is defined as a number whose exact value is unknown, but which falls within a known interval. A grey number thus enables us to represent the uncertainty associated with a given variable by means of an interval whose upper and lower limits are known, *although we have no information about the variable distribution within the interval itself* (Liu & Lin 2006; Alvisi & Franchini 2010). In other words, there are no probability assumptions on its distribution within such an interval.

Formally, where $s$ indicates a closed, limited set of real numbers with $s^-$ and $s^+$ being the lower and upper interval limits, respectively, a grey number $s^\pm$ can be defined as follows (Yang *et al.* 2004):

$$s^\pm = [s^-, s^+] = \{s \in s^\pm | s^- \le s \le s^+\} \tag{6}$$

It is possible to define the four operations between grey numbers as well as to compute a function $f$ of $n$ grey numbers $s_1^\pm, \cdots, s_i^\pm, \cdots, s_n^\pm$ as (Yang *et al.* 2004):

$$\left[f\left(s_1^\pm, \cdots, s_i^\pm, \cdots, s_n^\pm\right)\right]^\pm = \left[\left[f\left(s_1^\pm, \cdots, s_i^\pm, \cdots, s_n^\pm\right)\right]^-, \left[f\left(s_1^\pm, \cdots, s_i^\pm, \cdots, s_n^\pm\right)\right]^+\right] \tag{7}$$

where $\left[f\left(s_1^\pm, \cdots, s_i^\pm, \cdots, s_n^\pm\right)\right]^\pm$ is the grey function and $\left[f\left(s_1^\pm, \cdots, s_i^\pm, \cdots, s_n^\pm\right)\right]$ and $\left[f\left(s_1^\pm, \cdots, s_i^\pm, \cdots, s_n^\pm\right)\right]^+$ are the minimum and maximum function values, respectively. From a practical viewpoint, therefore, calculating the function $f$ of $n$ grey numbers $s_1^\pm, \cdots, s_i^\pm, \cdots, s_n^\pm$ specifically means looking for: (a) the set of real values $s_1, \cdots, s_i, \cdots, s_n$ which minimizes the function, where $s_i^- \le s_i \le s_i^+$; and (b) the set of real values $s_1, \cdots, s_i, \cdots, s_n$ (different from the previous one) which maximizes it, also contained in the domain $s_i^- \le s_i \le s_i^+$ In other words, it is necessary to solve two distinct problems of minimization and maximization of a function of $n$ variables, each variable $s_i$ being defined over an assigned interval $s_i^\pm$.

## The artificial neural network with grey parameters

Let us now consider a neural network whose structure coincides with the one previously described (see 'The ANN') but whose parameters are represented by grey numbers. Given the definition of function of grey numbers and since the output of a neural network is a function of weights and biases, if even only one of them is represented by a grey number, the network output will likewise be represented by a grey number. In particular, using the procedure for calculating a function of grey numbers previously described, the grey number characterizing the output of a neural network with grey weights and biases can be computed by looking for: (a) the particular set of weights and biases (where each value must be searched for within the corresponding interval) that minimizes the output of the neural network, thus producing the lower limit of the grey output variable; and (b) the set that maximizes it, thus producing the upper limit of the grey output variable. According to this procedure, for each output it should therefore be necessary to solve two optimization problems: first minimization and then maximization. Indeed, the procedure can be simplified by using the increasing monotonic property of the transfer functions typically used in neural networks, such as tan-sigmoid or log-sigmoid (Hagan *et al.* 1996).

In fact, since these functions are monotonic-increasing, the lower (upper) extreme of the output of these functions is obtained by considering the minimum (maximum) of their argument. Since the argument is the sum of several intervals derived from the product of the inputs and corresponding weights, its minimum (maximum) is obtained by adding up the lower (upper) extremes of each of these intervals (e.g. Alvisi & Franchini 2011).

In short, it is possible to characterize the grey number of the output variable without performing any optimization procedure, given the monotonic characteristic of the transfer functions.

## The evolutionary polynomial regression with grey parameters

As has been done for the ANN case, let us now consider an EPR model whose structure coincides with that previously described in the section 'The EPR', but whose parameters

are represented by grey numbers. The general procedure for calculating a function of grey numbers can also be initially considered in this case. Thus, the grey number characterizing the output of an EPR model with grey exponents and coefficients can be computed by looking for: (a) the particular set of exponents and coefficients (where each value must be searched for within the corresponding interval) that minimizes the output of the polynomial, thus producing the lower limit of the grey output variable; and (b) the set that maximizes it, thus producing the upper limit of the grey output variable. However, given its structure, it is also possible for the EPR model to directly compute the lower and upper limits of the grey output variable without performing any optimization. In fact, the polynomial is a sum of several terms; since each term of the polynomial is given by a coefficient multiplied by the product of $n_p$ power terms, its minimum (maximum) can be obtained by simply setting the coefficient and exponents of each polynomial term at their interval minimum (maximum) value. This of course applies to always positive inputs as in the case considered herein: the inputs are in fact discharges. Finally, once the minimum and maximum values of each term of the polynomial have been obtained, the lower (upper) extreme of the output of the whole polynomial is simply obtained by adding up the minimum (maximum) values of the polynomial terms.

## Grey formulation of the data-driven models to produce uncertainty bands

The ANN and the EPR models with grey parameters described above can be used to represent the total model uncertainty in performing forecasts. More precisely, with reference to the problem of discharge forecasting at the downstream end of a watercourse on the basis of the knowledge of the discharges measured in an upstream section, let us assume that the deterministic/crisp forms of the forecasting ANN and EPR models have been completely identified and parameterized by using the procedures proposed in the scientific literature (see the numerical example). Thus, the inputs to the models are known (for instance, some discharge values observed at the downstream end of the watercourse and in the upstream section up to the forecasting time $t$) and the crisp outputs $\hat{Q}^c(t+n\Delta t)$ can be calculated (i.e. the crisp discharges forecast at the downstream end of the watercourse $n$ $\Delta t$ time intervals ahead). Incidentally, the subscript d (for downstream) is not used in the notation $\hat{Q}^c(t+n\Delta t)$ since it is obvious within the present context that the discharge forecast is made in the downstream section.

Let us now consider the grey formulations of the previously identified ANN and EPR models. With reference to the ANN model for example, its grey formulation has the same structure as the crisp formulation but the weights and biases are grey numbers. Similarly for the EPR model, the coefficients and exponents are now grey numbers since the model structure is the same as the crisp formulation.

The output of the grey models is a grey number which represents their total uncertainty. We associate this grey number with the crisp forecast relative to the same time instant produced by the crisp formulation of the same model in order to have a crisp/deterministic forecast (which represents the best possible forecast) and an uncertain/imprecise forecast (which stresses the uncertainty linked to the model used) simultaneously.

The grey models can of course be applied once the grey parameters are known and they therefore have to be calibrated against observed data, as is done in the case of the corresponding crisp formulations. To this end, the grey parameters are searched for in such a way that the band obtained through the envelope of the intervals representing the grey numbers of the output variables (a) always includes the crisp forecast value, and (b) includes *at least* a preselected percentage PI of observed values. Thus, given $n_{obs}$ values of discharge observed at the downstream end of the watercourse $Q^{obs}_{d,1}, Q^{obs}_{d,2}, \ldots, Q^{obs}_{d,m}, \ldots, Q^{obs}_{d,n_{obs}}$ the grey output variable, i.e. the grey numbers $\hat{Q}^{\pm}_m$(where $m = 1{:}n_{obs}$) representing the output of the grey model (ANN or EPR), that is, the grey forecast discharge at the downstream end of the watercourse, must always contain the corresponding crisp forecast values $\hat{Q}^c_m$ (where $m = 1{:}n_{obs}$). If for example PI has been set at 100%, for each of the $n_{obs}$ available observations the grey output variable $\hat{Q}^{\pm}_m$ (where $m = 1{:}n_{obs}$) must also always contain the corresponding observed value. Similarly, if we assume PI = 95%, the grey output variable must always contain the crisp forecast value and at least 95% of the $n_{obs}$ discharges.

However, for a fixed PI, there are infinite different sets of grey parameters of the grey model considered such that the band obtained through the envelope of the grey output variables will always respect these constraints. Among all these sets, we search for the one producing the narrowest band.

Formally speaking, since the grey number of the predicted discharge $\hat{Q}_m^{\pm}$ is a function of the grey numbers of the parameters of the model, we look for the lower and upper extremes of the intervals representing these grey numbers in such a way that:

$$\sum_{m=1}^{n_{\text{obs}}} \left| \hat{Q}_m^{+} - \hat{Q}_{\text{m}}^{-} \right| \text{ is minimum} \tag{8}$$

subject to

$$\hat{Q}_m^{-} \le \hat{Q}_m^{\text{c}} \le \hat{Q}_m^{+} \quad m = 1{:}n_{\text{obs}} \tag{9}$$

and

$$\left( \frac{1}{n_{\text{obs}}} \sum_{m=1}^{n_{\text{obs}}} \delta_m \right) 100 \ge \text{POC} \tag{10}$$

where

$$\delta_m = \begin{cases} 1 & \text{if } \hat{Q}_m^{-} \le Q_{\text{d},m}^{\text{obs}} \le \hat{Q}_m^{+} \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

From a computational point of view, the constraint that the uncertainty band must always include the forecast provided by the crisp model can be ensured by obliging the grey parameters of the *grey model* to include the parameters of the *crisp model*. Furthermore, from an operative point of view, the set of crisp parameters of the *crisp model* can represent, within the calibration process, a good reference point around which to search for the set of grey parameters, thus speeding up the *grey model* calibration process.

Given the nature of the procedure described above, it is worth stressing that the grey numbers representing the model parameters do not quantify the uncertainty related to the parameters alone but instead quantify the uncertainty of the overall model. At the same time, the grey numbers representing the model output, i.e. the intervals that at each time instant quantify the total model uncertainty, define a band which is expected to contain a preset percentage (for instance, PI = 95%) of observed values. This (preselected) percentage is a measure of the uncertainty considered acceptable for the model. These bands can be used to assess the spread of the model prediction around the observed values and can therefore be interpreted as an estimate of the *validation uncertainty* (or 'emulation uncertainty'; Todini 2010) of the selected model and not of the *predictive uncertainty* (Todini 2008, 2010) since the procedure proposed does not develop any probability link between the true future value of the discharge and its corresponding crisp forecast.

## THE CASE STUDY

The crisp and grey formulations of the ANN and EPR models were applied to the case study represented by a reach of the Tiber River (Italy) delimited by the gauging stations of Pierantonio and Monte Molino, whose main hydrographical characteristics are summarized in Table 1. The location of the two gauging stations and the relative basin boundaries are shown in Figure 2. The numerical application involves 14 flood events (eight used for calibration and six for validation), all described with a time-step of half an hour.

The crisp ANN and EPR models were initially set up in order to forecast the discharge at the downstream end of the watercourse $n$ hours ahead, i.e. at $t + n\Delta t$, with $n = 1$, 5 and 9 and $\Delta t$=1 hour. In order to make the comparison as fair as possible, the same inputs were used for both models (see Table 2). In particular, as shown in Table 2, the discharge

**Table 1** | Hydrographical characteristics of the river reach: area subtended by the upstream section $A_u$, area subtended by the downstream section $A_d$, inter-watershed area $\Delta A$, length of reach $L$, mean bottom slope $S_0$, mean section width $B$, mean travel time $T_L$

| Tiber river reach | $A_u$ (km$^2$) | $A_d$ (km$^2$) | $\Delta A$ (km$^2$) | $L$ (km) | $S_0$ | $B$ (m) | $T_L$(h) |
|---|---|---|---|---|---|---|---|
| Pierantonio–Monte Molino | 1,805 | 5,279 | 3,474 | 71.06 | 0.0011 | 49.2 | 9 |

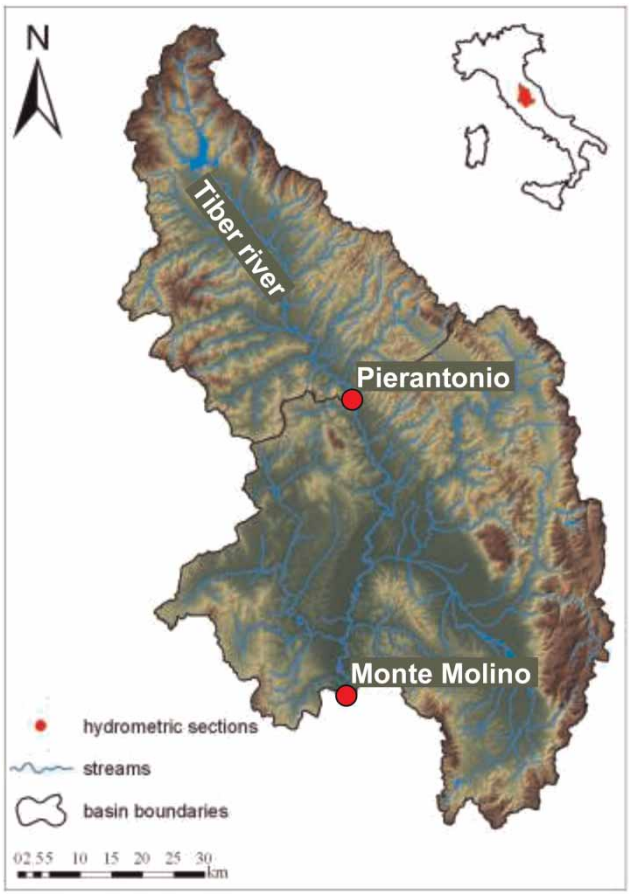


**Figure 2** | Location of the gauging stations within the watershed of the upper-middle reach of the Tiber river.

observed at the downstream end of the watercourse at the forecasting time $t$ and that observed at the time $t-\Delta t$ were used as inputs independently of the forecasting time horizon considered, since they represent the most up-to-date information available in the section where the discharge is forecast. Furthermore, with reference to the section upstream, the following inputs were selected: for the 1 hour ahead forecast, the discharges observed at $t-8\Delta t$ and $t-9\Delta t$; for the 5-hours-ahead forecast, the discharges observed at $t-4\Delta t$ and $t-5\Delta t$; and for the 9-hours-ahead forecast, the discharges observed $t$ and $t-\Delta t$. These inputs were selected since the mean travel time $T_L$ of the reach is of the order 9 hours. As a consequence, for each different forecasting lead time, these inputs represent the discharges observed in the upstream section at $t+n\Delta t-T_L$ and $t+n\Delta t-(T_L+\Delta t)$. Finally, it is worth noting that no information concerning the lateral inflows between the two gauging stations was available.

Regarding the structure of the models with reference to the crisp ANN forecasting model, the number of neurons in the hidden layer was searched for by enumeration in the interval 2–10 and finally set equal to 5 in order to avoid penalizing the model efficiency (Campolo *et al.* 1999; Thirumalaiah & Deo 2000). The crisp parameters, i.e. weights and biases, were calibrated using the Levenberg Marquardt algorithm (Hagan & Menhaj 1994) and an early stopping procedure in order to avoid over-fitting problems (ASCE Task Committee on the Application of Artificial Neural Networks in Hydrology 2000; Demuth & Beale 2000). To this end, the calibration set was split into two subsets: one for training (five events) and one for testing (three events).

The same subsets were used for the parameterization of the crisp EPR forecasting model. In particular, as mentioned above (see section 'The EPR'), the actual form of the crisp EPR model (see Equation (5)) is identified through an 'evolutionary searching strategy' which enables the simultaneous selection of the input variables to be inserted in each monomial of the summation and the corresponding exponents, which, in turn, are searched for within pre-selected intervals uniformly discretized. The coefficients $c_0$ and $c_i$ (with $i=1{:}n_t$) are instead computed through a numerical regression based on the least-square method (Giustolisi & Savic 2006).

**Table 2** | Inputs to the ANN and EPR models used to forecast the downstream discharge at different lead times $t+n\Delta t$, with $n=1$, 5 and 9 hours, $\Delta t=1$ hour

| | Inputs | | | |
|---|---|---|---|---|
| **Lead time** | **Downstream discharge observed at** | | **Upstream discharge observed at** | |
| $n=1$ | $T$ | $t-\Delta t$ | $t-8\Delta t$ | $t-9\Delta t$ |
| $n=5$ | $T$ | $t-\Delta t$ | $t-4\Delta t$ | $t-5\Delta t$ |
| $n=9$ | $T$ | $t-\Delta t$ | $t$ | $t-\Delta t$ |

The 'evolutionary searching strategy' consists of a multi-objective genetic algorithm which maximizes the efficiency of the model (measured through the root-mean-square error (RMSE) computed on the training set) and minimizes the complexity of the structure (measured through the number $n_t$ of monomials). This multi-objective calibration approach leads to a set of near-optimal polynomial models (each one characterized by a different number $n_t$ of monomials and where each monomial may have different input variables and exponents). This set represents a trade-off between efficiency and complexity. The final crisp EPR model was selected from these near-optimal models by analysing their performances (in terms of RMSE) on the testing set *a posteriori*. The final crisp EPR models for the 1-, 5- and 9-hours-ahead forecasts are listed in Table 3.

Regarding the parameterization of the grey formulation of both models, it is worth recalling that their structure is assumed to be the same as the corresponding crisp formulation; only the grey parameters need calibration. To this end, PI was set equal to 95%, i.e. the bands obtained through the envelope of the intervals representing the grey numbers of the output variables are required to contain at least 95% of the observed discharges. Furthermore, in order to ensure the inclusion of the crisp parameters in the corresponding grey parameters (see section 'Grey formulation of the data-driven models to produce uncertainty bands'), for each crisp parameter *par* (bias or weight in the case of the ANN model; coefficient or exponent for the EPR model) the lower extreme was expressed as $par^- = par \times (1 - \Delta^-)$ and the upper extreme $par^+$ of the corresponding grey parameter was expressed as $par^+ = par \times (1 + \Delta^+)$. The optimal values $\Delta^-$ and $\Delta^+$ were searched for in the open interval ]0,1[ by using the SCE-UA algorithm (Shuffled Complex Evolution, University of Arizona; Duan *et al.* 1992) and making a subsequent refinement through the MATLAB™ *fmincon* function based on Sequential Quadratic Programming (Powell 1983; Schitlowski 1985). It is worth noting that searching for the optimal values $\Delta^-$ and $\Delta^+$ in the open interval ]0,1[ ensures that, if the crisp parameter *par* is negative (positive), the corresponding grey parameter is also characterized by negative (positive) lower and upper extremes. Furthermore, a grey parameter is prevented from collapsing into a crisp value. This information, combined with the fact that the inputs considered are all discharges and thus always positive, was used to simplify the computation of the grey output of the EPR model, as described in the section 'The EPR with grey parameters'.

## ANALYSIS AND DISCUSSION OF THE RESULTS

The analysis and the comparison of the ANN and the EPR models are initially performed with reference to their crisp formulation.

**Table 3** | Crisp EPR models for the different forecasting time horizons ($\hat{Q}^c(t+\Delta t)$: crisp discharge forecast downstream at $t+\Delta t$; $Q_d^{obs}(t)$: discharge observed downstream at $t$; $Q_u^{obs}(t)$: discharge observed upstream at $t$)

| | **EPR *model*** |
|---|---|
| 1 hour ahead | $\hat{Q}^c(t+\Delta t) = -2.1994\, Q_d^{obs}(t)^{0.2}\, Q_d^{obs}(t-\Delta t)^{0.7} - 2.7238\, Q_d^{obs}(t)^{0.7}\, Q_d^{obs}(t-\Delta t)^{0.1} + 5.5127\, Q_d^{obs}(t)^{0.9} + 7.0563\times10^{-12}\, Q_d^{obs}(t)^{2.1}\, Q_d^{obs}(t-\Delta t)^{1.5}\, Q_u^{obs}(t-9\Delta t)^{0.6}$ |
| 5 hours ahead | $\hat{Q}^c(t+5\Delta t) = -14.7858\ Q_d^{obs}(t-\Delta t)^{0.7} + 12.6834\, Q_d^{obs}(t)^{0.7}\, Q_u^{obs}(t-4\Delta t)^{0.1} + 3.6663\times10^{-11}\, Q_d^{obs}(t)^{2.6}\, Q_d^{obs}(t-\Delta t)^{0.8}\, Q_u^{obs}(t-4\Delta t)^{0.4}\, Q_u^{obs}(t-5\Delta t)^{0.6} - 0.0414\, Q_d^{obs}(t-\Delta t)^{0.2}\, Q_u^{obs}(t-4\Delta t)^{0.7}\, Q_u^{obs}(t-5\Delta t)^{0.8} + 0.0673\, Q_d^{obs}(t)^{0.1}\, Q_u^{obs}(t-4\Delta t)^{1.5}$ |
| 9 hours ahead | $\hat{Q}^c(t+9\Delta t) = +0.0120\, Q_d^{obs}(t)^{1.6} - 9.542\times10^{-9} Q_d^{obs}(t)^{1.3}\, Q_d^{obs}(t-\Delta t)^{2.6}\, Q_u^{obs}(t-\Delta t)^{0.1} + 2.8846\times10^{-9}\, Q_d^{obs}(t)^{1.5}\, Q_d^{obs}(t-\Delta t)^{2.5}\, Q_u^{obs}(t)^{0.1}\, Q_u^{obs}(t-\Delta t)^{0.1} - 9.2595\times10^{-11}\, Q_d^{obs}(t)^{2.6}\, Q_d^{obs}(t-\Delta t)^{0.8}\, Q_u^{obs}(t)^{0.7}\, Q_u^{obs}(t-\Delta t)^{0.4} + 2.3455\, Q_u^{obs}(t)^{0.9}$ |

**Table 4** | Crisp ANN model. Root mean square error (RMSE) and coefficient of determination ($R^2$) for the different forecasting time horizons

| | Calibration | | Validation | |
|---|---|---|---|---|
| | RMSE ($m^3 s^{-1}$) | $R^2$ | RMSE ($m^3 s^{-1}$) | $R^2$ |
| 1 hour ahead | 5.97 | 0.9995 | 5.55 | 0.999 |
| 5 hours ahead | 36.37 | 0.98 | 32.94 | 0.98 |
| 9 hours ahead | 50.53 | 0.96 | 57.01 | 0.94 |

**Table 5** | Crisp EPR model. Root mean square error (RMSE) and coefficient of determination ($R^2$) for the different forecasting time horizons

| | Calibration | | Validation | |
|---|---|---|---|---|
| | RMSE ($m^3 s^{-1}$) | $R^2$ | RMSE ($m^3 s^{-1}$) | $R^2$ |
| 1 hour ahead | 6.42 | 0.999 | 5.60 | 0.999 |
| 5 hours ahead | 39.56 | 0.98 | 43.98 | 0.96 |
| 9 hours ahead | 61.62 | 0.94 | 78.24 | 0.88 |

The RMSE and the coefficient of determination ($R^2$) of the crisp ANN forecasting models for the different forecast time horizons are listed in Table 4 for both calibration and validation phases. Similarly, the RMSE and the $R^2$ of the crisp EPR forecasting models are given in Table 5. Furthermore, Figures 3(a) and Figures 3(b) show the 9-hours-ahead forecasts provided by the crisp ANN and EPR models, respectively, for one of the events of the calibration set, along with the observed discharges. Similarly, Figures 4(a) and Figures 4(b) show the 9-hours-ahead forecasts provided by the crisp ANN and EPR models, respectively, for an event of the validation set.

As can be observed, the two crisp forecasting models provide similar results: in fact, Figures 3 and 4 show that the trends of the discharges forecast by the two crisp models are very similar; in particular, the validation event of Figures 4(a) and Figures 4(b) highlights that the

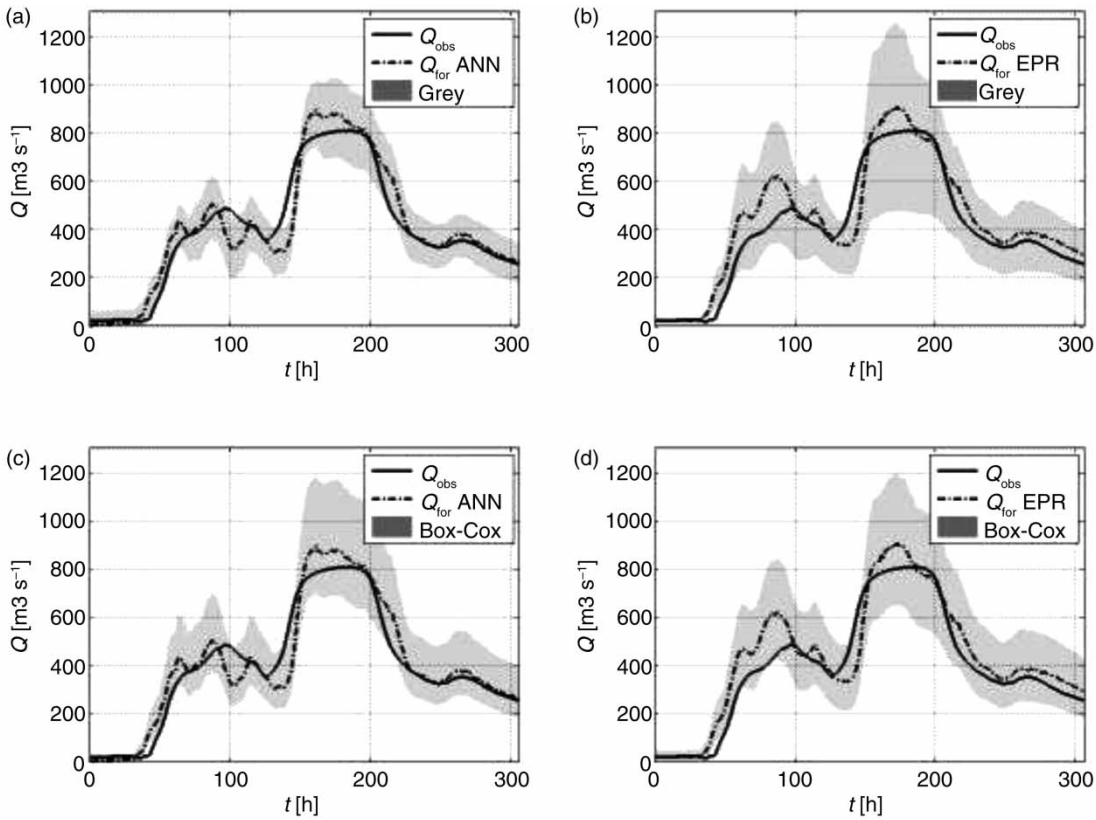


**Figure 3** | Calibration event: lead time 9 hours. Discharges forecast through the (a) crisp ANN model and uncertainty band estimated through the grey formulation; (b) crisp EPR model and uncertainty band estimated through the grey formulation; (c) crisp ANN model and uncertainty band estimated through the Box–Cox method; and (d) crisp EPR model and uncertainty band estimated through the Box–Cox method.

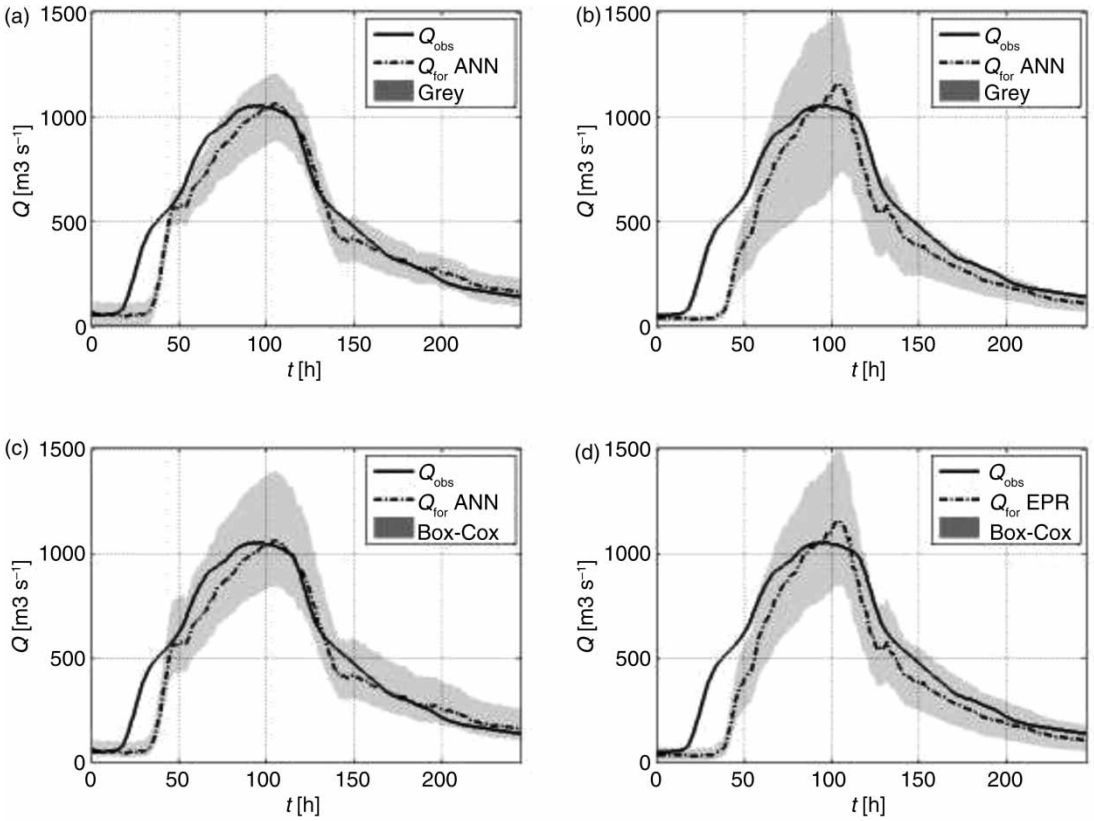


**Figure 4** | Validation event: lead time 9 hours. Discharges forecast through the (a) crisp ANN model and uncertainty band estimated through the grey formulation; (b) crisp EPR model and uncertainty band estimated through the grey formulation; (c) crisp ANN model and uncertainty band estimated through the Box–Cox method; and (d) crisp EPR model and uncertainty band estimated through the Box–Cox method.

forecasts provided by both models tend to underestimate the observed discharges during the rising limb, thus leading to a delay of the forecast; this is comprehensible by considering that no input information is provided concerning the lateral inflow between the two gauging stations considered.

In more detail, the statistics shown in Tables 4 and 5 highlight that the performances of the two models are equivalent up to a forecasting time horizon of 5 hours (both in calibration and validation phases). For a forecasting time horizon of 9 hours, the performances of the crisp ANN forecasting model are slightly better than those of the model based on crisp EPR model (in both calibration and validation phases). For the 9-hours-ahead forecast in the calibration phase the $R^2$ of the crisp ANN model is 0.96 compared to 0.94 of the crisp EPR model, and in validation phase the $R^2$ of the crisp ANN model is 0.94 compared to 0.88 of the crisp EPR model.

As regards the uncertainty bands, Figures 3(a) and Figures 3(b) show the bands obtained through the use of the grey formulation of the ANN and EPR models, respectively, for the event of the calibration set; similarly Figures 4(a) and Figures 4(b) show the bands obtained for the event of the validation set. It is worth noting that in the time interval between the 30th and 50th hour of the validation event, especially in the case of the ANN model (Figure 4(a)), the uncertainty band seems to be very narrow. This is only a visual effect due to the slope of the forecast rising limb, however; indeed, the amplitude of the band is of the order 150 $m^3 s^{-1}$ like that observed in the falling limb between the 150th and 250th hour.

Analysis of these figures shows that the uncertainty bands produced by the grey ANN model are narrower than those produced by the EPR model. This is confirmed by the comparison between Tables 6 and 7 where the Average Width (AW, Shrestha & Solomatine 2008; Zhang

**Table 6** | Grey ANN forecasting model. Average width (AW) and percentage of observed values actually included (POC) evaluated for all the events of the calibration and validation sets and for the different forecasting time horizons. The range of POC when it is evaluated in each event is also provided in brackets

| | Calibration | | Validation | |
|---|---|---|---|---|
| | AW ($m^3 s^{-1}$) | POC (%) | AW ($m^3 s^{-1}$) | POC (%) |
| 1 hour ahead | 26.0 | 95.0 (92.7–99.2) | 22.7 | 95.7 (87.8–98.9) |
| 5 hours ahead | 120.5 | 95.0 (87.3–99.2) | 104.6 | 94.5 (85.7–100) |
| 9 hours ahead | 186.6 | 95.0 (85.8–100) | 159.6 | 94.1 (82.1–100) |

**Table 7** | Grey EPR forecasting model. Average width (AW) and percentage of observed values actually included (POC) evaluated for all the events of the calibration and validation sets and for the different forecasting time horizons. The range of POC when it is evaluated in each event is also provided in brackets

| | Calibration | | Validation | |
|---|---|---|---|---|
| | AW ($m^3 s^{-1}$) | POC (%) | AW ($m^3 s^{-1}$) | POC (%) |
| 1 hour ahead | 44.5 | 95.0 (92.6–100) | 35.6 | 94.6 (87.7–99.4) |
| 5 hours ahead | 220.9 | 95.0 (86.8–100) | 163.6 | 88.7 (73.6–100) |
| 9 hours ahead | 271.9 | 95.0 (83.1–100) | 169.8 | 77.0 (45.5–100) |

*et al.* 2009) defined as:

$$\mathrm{AW} = \frac{1}{n_{\mathrm{obs}}} \sum_{m=1}^{n_{\mathrm{obs}}} \left| \hat{Q}_m^+ - \hat{Q}_m^- \right| \tag{12}$$

and the percentage of observed values actually contained within the band (POC or percentage of coverage; Xiong *et al.* 2009), numerically defined as:

$$\mathrm{POC} = \frac{1}{n_{\mathrm{obs}}} \sum_{m=1}^{n_{\mathrm{obs}}} \delta_m$$
$$\text{where} \quad \delta_m = \begin{cases} 1 & \text{if } \hat{Q}_m^- \le Q_m^{\mathrm{obs}} \le \hat{Q}_m^+ \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

are given.

Table 6 also shows that, in the ANN case, the grey formulation guarantees that the percentages of observed values contained within the different bands (POC) perfectly correspond to the percentages PI imposed in the calibration phase (with reference to the set of all 8 events). This confirms the accuracy of the calibration procedure developed and, at the same time, guarantees values of POC slightly smaller than 95% in the validation phase (i.e. with reference to the set of all six events), thus confirming the robustness of the entire approach in characterizing the vagueness associated with the predicted variable (Shrestha & Solomatine 2008).

Similar considerations do not apply to the grey EPR forecasting model (Table 7). In fact, in the validation phase, the POC index significantly decreases when reference to the time horizon of 5 and 9 hours is made, although the AW is greater than that of the ANN model. Incidentally, it is worth noting that Tables 6 and 7 also show the range of POC when it is evaluated in each event, both in calibration and validation phase. A certain variability is observed for both the models, but the ANN model shows a reduced range for all three lead times.

The grey uncertainty bands of both models were also compared to those drawn on the basis of a methodology well established in the scientific literature (e.g. Feyen *et al.* 2007) which computes the errors between observed and forecast discharges after their Box–Cox transformation (Box & Cox 1974), in order to take into account the heteroscedasticity of the error (Vrugt *et al.* 2003). In more detail, the latter bands were set up in the following way. For each instant in time, the transformed values of the observed and (crisp) forecast discharges were calculated with the parameter $\lambda$ of the transformation estimated through the

method proposed by Box & Cox (1974), which produced $\lambda = 0.3$. The standard deviation $\hat{\sigma}$ of errors between the observed and crisp forecast values in the transformed plane was then calculated. The end points of the 95% uncertainty band were estimated in the transformed plane by adding $\pm 1.96\hat{\sigma}$ to the transformed crisp forecast values, assuming that the transformed errors have a Gaussian distribution. Finally, these end points were transformed back, thus providing the Box–Cox uncertainty band in the natural plane.

Figures 3(c) and Figures 3(d) (for the event of the calibration set) and Figures 4(c) and Figures 4(d) (for the event of the validation set) enable the comparison of the Box–Cox bands and grey bands. It is evident that, in the case of the ANN forecasting model, the grey bands are narrower than those based on the Box–Cox method; in the case of the EPR forecasting model, the widths of the uncertainty bands obtained through the grey approach are greater. These results are confirmed through a comparison of Tables 6 and 8 in the case of the ANN model and Tables 7 and 9 in the case of the EPR model.

The fact that the grey approach used to characterize the total uncertainty of a forecasting model leads to good results when applied to the ANN model and less good results when applied to the EPR model can be explained by considering the different structures of the two models and the very nature of the grey approach which transfers the total model uncertainty to the model parameters. The ANN model is extremely 'flexible' (the Universal Approximation Theorem developed by Hornik *et al.* 1989 states that it is possible to approximate any function through an ANN) and, due to its many and distributed parameters, the 'greyification' process enables the variability/vagueness of the forecast variable to be represented by means of narrow and relatively uniform bands. On the other hand, the EPR model is characterized by a simpler structure in which the power terms play a major role. As a consequence, for given grey exponents, the bandwidth tends to narrow for low values of the input variables (see Table 3) while it broadens significantly in correspondence to high input values which occur in correspondence to the flood peak, as clearly highlighted in Figures 3(b) and Figures 4(b).

All previous comparisons were made with reference to the *entire* events without distinguishing between peaks of different order of magnitude. From a practical point of view, a flood manager could be interested in flood peaks over a certain threshold; the value of the crisp forecast and (even more) of the uncertainty band might therefore be assessed in terms of their capability to capture and contain such flood peaks, assigning less importance to failure when it occurs with reference to floods of smaller entity. This type of assessment is out of the scope of the present study, but it may be useful in practical situations when the reliability of a model has to be quantified.

**Table 8** | Uncertainty bands obtained through the Box–Cox approach applied to the errors of the crisp ANN forecasting model. Average width (AW) and percentage of observed values actually included (POC) for the different forecasting time horizons

| | Calibration | | Validation | |
|---|---|---|---|---|
| | AW ($m^3 s^{-1}$) | POC (%) | AW ($m^3 s^{-1}$) | POC (%) |
| 1 hour-ahead | 33.1 | 94.9 | 27.8 | 94.9 |
| 5 hours-ahead | 178.3 | 95.0 | 146.3 | 96.4 |
| 9 hours-ahead | 241.9 | 95.2 | 198.6 | 96.9 |

**Table 9** | Uncertainty bands obtained through the Box–Cox approach applied to the errors of the crisp EPR forecasting model. Average width (AW) and percentage of observed values actually included (POC) for the different forecasting time horizons

| | Calibration | | Validation | |
|---|---|---|---|---|
| | AW ($m^3 s^{-1}$) | POC (%) | AW ($m^3 s^{-1}$) | POC (%) |
| 1 hour-ahead | 34.4 | 94.9 | 28.9 | 94.5 |
| 5 hours-ahead | 179.2 | 95.0 | 141.2 | 89.8 |
| 9 hours-ahead | 267.7 | 95.0 | 196.1 | 89.2 |

## CONCLUSIONS

Two data-driven models, namely an ANN model and an EPR model, have been applied to perform the discharge forecasting at the downstream end of a river reach on the basis of the discharge values observed at both ends. These two models have different structures: the ANN model is characterized by a complex and highly non-linear interrelated structure, while the EPR model has a general polynomial structure.

Both models have been parameterized in crisp and grey modes. The crisp formulation enables the crisp forecast of the downstream discharges for selected time horizons and both models performed very well with reference to the calibration period; the ANN model slightly outperformed the EPR model in the validation period.

The grey formulation of the two models has been used to characterize the total model uncertainty. In this case, the parameters of the two models are represented by grey numbers which are estimated through a procedure where the envelope of the intervals (band) representing the outputs (grey discharges, calculated at different points in time) must include a preset percentage of observed values and, at the same time, be as narrow as possible. These bands do not have a probabilistic meaning, given the criterion used for their definition, and can be considered as an estimate of the 'validation uncertainty' of the model considered.

The comparison between the bands produced by the grey formulation of the ANN and EPR models has shown that those produced by the grey ANN model are usually narrower (the percentage of observed values included being equal). Furthermore, the grey ANN model maintains a good performance in the validation period while the grey EPR model deteriorates significantly.

The bands produced by the grey formulation of the two models have also been compared to those produced by the Box–Cox method. The bands produced by the grey ANN model generally have a smaller width, the percentage of coverage being equal. The opposite result is obtained in the case of the grey EPR.

The different results are explained on the basis of the structure of the two data-driven models: in particular, the predominant role of the power terms in the EPR model heightens the effect of the grey exponents when high input values are considered, thus producing broad and narrow bands in correspondence to the peaks and low flows, respectively.

Overall, these findings highlight that the approach based on grey numbers can be a valuable tool for representing the uncertainty associated with the selected model; its performance is therefore in some way dependent on its characteristics and structure. Although this approach can be applied to any hydrologic forecasting model from a theoretical point of view, its use always requires an analysis of the response of the model selected to the 'greyification' process.

## ACKNOWLEDGEMENTS

This study was conducted within the framework of the PRRITT 2008-LAB.ENVIREN project. The authors wish to thank Dr Tommaso Moramarco (IRPI-CNR-Perugia (I)) for providing the data. The authors are grateful to Prof. Ezio Todini and to the second anonymous reviewer for their helpful and constructive comments that helped us to improve the quality and clarity of the paper.

## REFERENCES

Alvisi, S. & Franchini, M. 2010 Pipe roughness calibration in water distribution systems using grey numbers. *J. Hydroinform.* **12** (4), 424–445.

Alvisi, S. & Franchini, M. 2011 Fuzzy neural networks for water level and discharge forecasting with uncertainty. *Environ. Model. Softw.* **26** (4), 523–537.

ASCE Task Committee on the application of ANN in Hydrology 2000 Artificial neural networks in hydrology. II: Hydrological applications. *J. Hydrol. Eng.* **5** (2), 124–137.

Berardi, L., Laucelli, D. & Giustolisi, O. 2010 Accounting for uncertainty of variables in data-driven modeling by EPR. In *Proc. of 9th Int. Conf. on Hydroinformatics*, Tianjin, China.

Box, G. E. P. & Cox, D. 1974 An analysis of transformations. *J. R. Stat. Soc.* **26**, 211–252.

Campolo, M., Andreussi, P. & Soldati, A. 1999 River flood forecasting with neural network model. *Water Resour. Res.* **35** (4), 1191–1197.

Demuth, H. B. & Beale, M. 2000 *Neural Network Toolbox for use with Matlab*. The Math Works, Inc., Natick.

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.

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.

El-Baroudy, I., Elshorbagy, A., Carey, S. K., Giustolisi, O. & Savic, D. 2010 Comparison of three data-driven techniques in modelling the evapotranspiration process. *J. Hydroinform.* **12** (4), 365–379.

Feyen, L., Vrugt, J. A., ÓNuallain, B., van der Knijff, J. & De Roo, A. 2007 Parameter optimisation and uncertainty assessment for large-scale Streamflow simulation with the LISFLOOD model. *J. Hydrol.* **332**, 276–289.

Giustolisi, O. & Savic, D. A. 2006 A symbolic data-driven technique based on evolutionary polynomial regression. *J. Hydroinform.* **8** (3), 207–222.

Giustolisi, O. & Savic, D. A. 2009 Advances in data-driven analyses and modelling using EPR-MOGA. *J. Hydroinform.* **11** (3), 225–236.

Giustolisi, O., Doglioni, A., Savic, D. A. & di Pierro, F. 2008 An Evolutionary multi-objective strategy for the effective management of groundwater resources. *Water Resour. Res.* **44**, W01403.

Hagan, M. T. & Menhaj, M. 1994 Training feedforward networks with the Marquardt algorithm. *IEEE Trans. Neural Netw.* **5** (6), 989–993.

Hagan, M. T., Demuth, H. B. & Beale, M. 1996 *Neural Network Design*. PWS Publishing Company, Boston.

Hornik, K., Stinchombe, M. & White, H. 1989 Multilayer feedforward networks are universal approximators. *Neural Netw.* **2**, 359–366.

Jacquin, A. P. & Shamseldin, A. Y. 2007 Development of a possibilistic method for the evaluation of predictive uncertainty in rainfall-runoff modelling. *Water Resour. Res.* **43**, W04425

Krzysztofowicz, R. 1999 Bayesian theory of probabilistic forecasting via deterministic hydrologic model. *Water Resour. Res.* **35** (9), 2739–2750.

Kuczera, G. & Parent, E. 1998 Monte Carlo assessment of parameter uncertainty in conceptual catchment models: the Metropolis algorithm. *J. Hydrol.* **211**, 69–85.

Liu, S. & Lin, Y. 2006 *Grey Information: Theory and Practical Applications*. Springer-Verlag, London.

Maier, H. R. & Dandy, G. C. 2000 Neural Networks for prediction and forecasting of water resources variables: a review of modelling issue and applications. *Environ. Model. Softw.* **15**, 101–124.

Montanari, A. & Grossi, G. 2008 Estimating the uncertainty of hydrological forecasts: a statistical approach. *Water Resour. Res.* **44**, W00B08.

Powell, M. J. D. 1983 Variable metric methods for constrained optimization. In: *Mathematical Programming: The State of the Art* A. (Bechem, M. Grostschel & B. Korte, eds.). Springer Verlag, Germany, 288–311.

Rosenblatt, F. 1958 The Perceptron: a probabilistic model for information storage and organization in the brain. *Psychol. Rev.* **65**, 386–408.

Schitlowski, K. 1985 NLQPL: a Fortran-subroutine solving constrained nonlinear programming problems. *Operations Res.* **5**, 485–500.

Shrestha, D. L. & Solomatine, D. P. 2008 Data-driven approaches for estimating uncertainty in rainfall runoff modelling. *J. River Basin Manage* **6** (2), 109–122.

Solomatine, D. P. & Shrestha, D. L. 2009 A novel method to estimate model uncertainty using machine learning techniques. *Water Resour. Res.* **45**, W00B11.

Thirumalaiah, K. & Deo, M. C. 2000 Hydrological forecasting using neural networks. *J. Hydrol. Eng.* **5** (2), 180–189.

Todini, E. 1988 Rainfall-runoff modelling: past, present and future. *J. Hydrol.* **100**, 341–352.

Todini, E. 2004 Role and treatment of uncertainty in real-time flood forecasting. *Hydrol. Process.* **18**, 2743–2746.

Todini, E. 2007 Hydrological catchment modelling: past, present and future. *Hydrol. Earth Syst. Sci.* **11** (1), 468–482.

Todini, E. 2008 A model conditional processor to assess predictive uncertainty in flood forecasting. *Intl. J. River Basin Manage.* **6** (2), 123–137.

Todini, E. 2010 Predictive uncertainty in flood forecasting and emergency management. In *Proceedings of 17th Congress of IAHR-APD 2010*, Auckland, New Zealand. Available from: (http://ebookbrowse.com/arlai-iahr-apd2010-pdf-d91676578).

Tung, Y. K. 1996 Uncertainty and reliability analysis. In: *Water Resources Handbook* (L. W. Mays, ed.). McGraw-Hill Book Company, New York, pp. 71–765.

Vrugt, J. A., Gupta, H. V., Bouten, W. & Sorooshian, S. 2003 A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters. *Water Resour. Res.* **39** (8), 1201–1218.

Xiong, L., Wan, M., Wei, X. & O'Oonnor, K. 2009 Indices for assessing the prediction bounds of hydrological models and application by generalised likelihood uncertainty estimation. *Hydrol. Sci. J.* **54** (5), 852–871.

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*, Loughborough, UK, pp. 239–246.

Zhang, X., Liang, F., Srinivasan, R. & Van Liew, M. 2009 Estimating uncertainty of streamflow simulation using Bayesian neural networks. *Water Resour. Res.* **45**, W02403.

First received 21 January 2011; accepted in revised form 25 November 2011. Available online 3 May 2012