Water Supply

Water Supply Vol 22 No 3, 3254 doi: 10.2166/ws.2021.401

Uncertain time series forecasting method for the water demand prediction in Beijing

Haiyan Li[a], Xiaosheng Wang[b,*] and Haiying Guo[b]

[a] School of Economics, Nankai University, Tianjin 300071, China
[b] School of Mathematics and Physics, Hebei University of Engineering, Handan 056038, China
*Corresponding author. E-mail: xswang@hebeu.edu.cn

ABSTRACT

Water demand prediction is crucial for effective planning and management of water supply systems to handle the problem of water scarcity. Taking into account the uncertainties and imprecisions within the framework of water demand forecasting, the uncertain time series prediction method is introduced for water demand prediction. Uncertain time series is a sequence of imprecisely observed values that are characterized by uncertain variables and the corresponding uncertain autoregressive model (UAR) is employed to describe it for predicting future values. The main contributions of this paper are shown as follows. Firstly, by defining the auto-similarity of uncertain time series, the identification algorithm of UAR model order is proposed. Secondly, a new parameter estimation method based on the uncertain programming is developed. Thirdly, the imprecisely observed values are assumed as the linear uncertain variables and a ratio-based method is presented for constructing the uncertain time series. Finally, the proposed methodologies are applied to model and forecast Beijing's water demand under different confidence levels and compared with the traditional time series, i.e. ARIMA method. The experimental results are evaluated on the basis of performance criteria, which shows that the proposed method outperforms over the ARIMA method for water demand prediction.

Key words: uncertainty theory, uncertain time series, water demand prediction

HIGHLIGHTS

- Considering the uncertainty of water demand, the uncertain time series method for demand estimation of water resources is presented.
- The auto-similarity of uncertain time series is defined, and the identification algorithm of uncertain autoregressive model order is proposed.
- An uncertain programming approach to estimate the parameters of model is proposed.
- The construction of liner uncertain time series is investigated due to the interval-valued data frequently encountered in real life.

1. INTRODUCTION

Water is an indispensable natural resource on our planet, which plays an important role in man's life and activity. Apart from drinking and personal hygiene, water is still a necessary resource for agricultural and industrial production, economic and ecological development (Deng *et al.* 2015; Liu *et al.* 2015). However, due to climate change, socio-economic development and population growth, water consumption (especially for freshwater) is growing rapidly, and water supply is facing many challenges (Choksi *et al.* 2015), especially the problem of water scarcity (Frederick 1997; Pahl-Wostl 2007; Arnell & Lloyd-Hughes 2014; Wang *et al.* 2015). This has led to the need for effective planning, managing and operating of finite water resources (Oduro-Kwarteng *et al.* 2009; Wang *et al.* 2018). Therefore, water demand forecasting is a fundamental phase for optimal allocation of water resources and aims to provide the simulated view of future demand, which can assist decision makers in devising appropriate management schemes to relieve the conflict between growing demand and limited supply of water resources. To this end, many researchers have proposed different methods to model and forecast water demand.

Time series analysis is one of the commonly used methods for water demand prediction. It was proposed by Box and Jenkins who considered the dependence among the data. The model, namely, ARIMA, is regarded as a classical forecasting technique, describing a predicted value as a linear function of previous data and random errors and including a cyclical or seasonal component. For example, Maidment *et al.* (1985) applied the time series model of daily municipal water use as a function of rainfall and air temperature for short-term forecasting of daily water use in Austin, Texas. Smith (1988) developed an autoregressive process with randomly varying mean to forecast the daily municipal water use, which captured the seasonality and day-week effects in the model through the unit demand function. Aly & Wanakule (2004) utilized a deterministic smoothing algorithm that considered level, trend and seasonality components of time series to estimate monthly water use. Zhai *et al.* (2012) employed the time series forecasting method for predicting the future needs of water in Beijing by analyzing the driving mechanism of changes of water consumption and water consumed structure.

In a time series modelling application, the determination of the model order is the fundamental step towards describing any dynamic process and has been of considerable interest for a long time. Primarily, the determination method of the order of time series model is based on the properties of sample autocorrelation coefficient and partial autocorrelation coefficient. Following that, several order selection approaches based on information theoretic criteria, such as Akaike's information criterion (AIC) (Akaike 1974), Akaike's final prediction error (FPE) (Akaike 1970), minimum description length (MDL) (Rissanen 1978; Liang *et al.* 1993) and so on, have been developed. Another common method, namely, the linear algebraic method based upon the determinant and rank testing algorithms, was proposed in (Fuchs 1987) and Sadabadi *et al.* (2007). Apart from two methods above, many other methods like bayesian information criterion (BIC) (Schwarz 1978), edge detection-based approach (Al-Smadi & Al-Zaben 2005), optimal instrumental variable (IV) algorithm (Sadabadi *et al.* 2009) and so on were investigated to estimate the order of the time series model. Another estimation problem that has also been considerably investigated is the aspect of coefficients determination of time series model. Commonly used methods for estimation of unknown coefficients are least-squares (LS) estimator and maximum likelihood estimator (MLE) methods.

Based on the mentioned methods of model order identification and parameters estimation, the time series models can be formulated to forecast water demand. It is worth noting that the aforementioned models provided a single valued forecast of water demand, disregarding the uncertainty inherent in some situations where the influential factors that affect water demand are uncertain, which leads to the uncertainty of water demand. This would limit the usefulness of these deterministic models. One classical way to handle the uncertainty is to use a probabilistic model (Almutaz *et al.* 2012; Haque *et al.* 2014) based on the Monte Carlo Stimulations (MCS) to obtain the distribution of water demand and provide an estimate of the overall uncertainty in the predictions connected to uncertainty of influential factors.

Unfortunately, the distribution function obtained in most practical problems is not close enough to the actual frequency, especially in the case of emergencies and lack of history data. In addition, the water demand data possess uncertain characteristics caused by inaccuracies in measurements that need to be given by experts. This motivates us to apply a new mathematical tool to deal with a range of uncertainties inherent in certain water demand data. Recently, uncertainty theory was proposed by Liu (2009) in 2007, which is an effective way to solve previous problem for imprecisely observed values. Based on the uncertainty theory, many researchers have done a lot of work including the determination of uncertain distribution (Wang *et al.* 2012a; Wang & Peng 2014), hypothesis test (Guo *et al.* 2017; Ye & Liu 2021), and uncertain regression analysis (Wang *et al.*, 2012b; Lio & Liu 2018; Yao & Liu 2018; Ye & Liu 2020). Furthermore, the concept of uncertain time series was firstly proposed by Yang & Liu (2019) based on uncertain theory in 2019. Like the traditional time series analysis, there may be more than one approach to model time series. However, in their study, to describe uncertain time series, the UAR model was employed to predict the future values based on previously imprecisely observed values that are characterized in terms of uncertain variables. Based on the imprecisely observed values, Yang & Liu (2019) presented the least-squares method to estimate the coefficients of the UAR model for predicting the carbon emission.

However, there are still many important issues that have not been touched. Firstly, the identification of UAR model order is one of these, because it is the first step in estimating the model parameters. In the work of Yang & Liu (2019), the 2-order UAR model was directly employed to forecast the future values. This method is too subjective and lacks a certain theoretical foundation, which may reduce the prediction accuracy of the model. So, in this paper, by defining the auto-similarity of uncertain time series, an algorithm for determining the optimal order of autoregressive model is designed. Secondly, its novel parameter estimation approach is developed based on uncertain programming. Within the proposed method, the original problem including uncertain measure is transformed to the equivalent crisp mathematical programming. Thirdly, in our daily life, most information is uncertain in nature. For example, water demand naturally takes different values with minimum water demand and maximum water demand, which are inherently imprecisely observed values at times t, $t = 1, 2, \cdots, n$, respectively, so the linear uncertain variables are selected for this purpose. Hence, it is a critical issue for us to determine the lower and upper bounds for the actual data belonging to a range. That is, how to construct an uncertain time series based on observed historical point data. Referring to the work of Huarng (2006), we introduced a novel ratio-based approach to determine the effective uncertain time series. Furthermore, the proposed uncertain time series forecasting approach is used to predict the urban water demand. As the second-largest city of China, Beijing's rapid development has attracted many immigrants in recent years and water consumption is growing rapidly, which led to the increasingly sharp conflict between demand and supply of water resources. This situation has become an important constraint on the sustainable development of Beijing. Therefore, the water demand prediction of Beijing is a fundamental stage for water resources planning and utilization, which contributes to harmonious development between the socio-economy and resources environment in Beijing. To further verify the accuracy of the proposed methodologies, traditional time series method is selected as a competitor. The results are judged on the basis of presented criteria, i.e. the prediction reliability and accuracy compared to ARIMA method.

The organization of this paper is as follows. Section 2 briefly presents some fundamental concepts properties and theorems in uncertainty theory. Section 3 introduces the forecasting procedure of uncertain time series analysis. Section 4 provides an experimental analysis to validate the effectiveness of the proposed method and access its performance by comparing with the conventional time series (ARIMA) method. Finally, some conclusions are drawn.

2. PRELIMINARIES

In this section, we will present some fundamental definitions and theorems on uncertainty theory.

Definition 1. (Liu 2007) *Let Γ be a nonempty set, and $\mathcal{L}$ be a σ-algebra over Γ. Each element $\Lambda \in \mathcal{L}$ is called an event. A number $\mathcal{M}\{\Lambda\}$ indicates the belief degree that Λ will occur. Then $\mathcal{M}$ is called an uncertain measure if it satisfies the following axioms:*

Axiom 1: *(Normality Axiom) $\mathcal{M}\{\Gamma\} = 1$ for the nonempty set Γ.*

Axiom 2: *(Duality Axiom) $\mathcal{M}\{\Lambda\} + \mathcal{M}\{\Lambda^c\} = 1$ for any event Λ.*

Axiom 3: *(Subadditivity Axiom) For every countable sequence of events $\Lambda_i,\ i = 1, 2, \cdots$, we have*

$$\mathcal{M}\left\{\bigcup_{i=1}^{\infty} \Lambda_i\right\} \leq \sum_{i=1}^{\infty} \mathcal{M}\{\Lambda_i\} \tag{1}$$

In this case, the triplet $(\Gamma, \mathcal{L}, \mathcal{M})$ is called an uncertainty space.

Then the product uncertain measure on the product σ-algebra $\mathcal{L}$ was defined by Liu (2009), producing the fourth axiom of uncertain measure.

Axiom 4: *(Product Axiom) Let $(\Gamma_k, \mathcal{L}_k, \mathcal{M}_k)$ be uncertainty spaces for $k = 1, 2, \cdots$. The product uncertain measure $\mathcal{M}$ is an uncertain measure satisfying*

$$\mathcal{M}\left\{\prod_{k=1}^{\infty} \Lambda_k\right\} = \bigwedge_{k=1}^{\infty} \mathcal{M}_k\{\Lambda_k\} \tag{2}$$

where Λ_k are arbitrarily chosen events from $\mathcal{L}$ for $k = 1, 2, \cdots$, respectively.

The concept of uncertain variable ξ was introduced by Liu as a measurable function from an uncertainty space $(\Gamma, \mathcal{L}, \mathcal{M})$ to the set of real numbers.

Definition 2. (Liu 2007) *An uncertain variable is a measure function ξ from an uncertain space $(\Gamma, \mathcal{L}, \mathcal{M})$ to the set of real number. That is, for any Borel set B, the set*

$$\{\xi \in B\} = \{\gamma \in \mathcal{L} | \xi(\gamma) \in B\} \tag{3}$$

is an event.

Definition 3. (Liu 2007) *Let $\xi_1, \xi_2, \cdots, \xi_n$ be uncertain variables, and let f be a real-valued measurable function. Then $\xi = f(\xi_1, \xi_2, \cdots, \xi_n)$ is an uncertain variable defined by*

$$\xi(\gamma) = f(\xi_1(\gamma), \xi_2(\gamma), \cdots, \xi_n(\gamma)), \quad \forall \gamma \in \Gamma. \tag{4}$$

Definition 4. (Liu 2007) *The uncertainty distribution Φ of an uncertain variable ξ is*

$$\Phi(x) = \mathcal{M}\{\xi \leq x\}, \quad \forall x \in R. \tag{5}$$

Definition 5. (Liu 2007) *An uncertain variable is called linear if it has a linear uncertainty distribution*

$$\Phi(x) = \begin{cases} 0, & \text{if} \quad x \le a \\ (x-a)/(b-a), & \text{if} \quad a \le x \le b \\ 1, & \text{if} \quad x \ge b \end{cases} \tag{6}$$

denoted by $L(a, b)$ *where* a *and* b *are real numbers with* $a < b$.

Definition 6. (Liu 2010a) *An uncertainty distribution* $\Phi(x)$ *is said to be regular if it is continuous and strictly increasing function with respect to* x *at which* $0 < \Phi(x) < 1$, *and*

$$\lim_{x\to-\infty} \Phi(x) = 0, \ \lim_{x\to\infty} \Phi(x) = 1. \tag{7}$$

Definition 7. (Liu 2010a) *Let* ξ *be an uncertain variable with regular uncertainty distribution* $\Phi(x)$. *Then the inverse function* $\Phi^{-1}(\alpha)$ *is called the inverse uncertainty distribution of* ξ.

Example 1. The inverse uncertainty distribution of linear uncertain variable $L(a, b)$ is

$$\Phi^{-1}(\alpha) = (1-\alpha)a + \alpha b. \tag{8}$$

Definition 8. (Liu 2007) *Let* ξ *be an uncertain variable. Then the expected value of* ξ *is defined by*

$$E[\xi] = \int_0^{+\infty} \mathcal{M}\{\xi \ge r\}dr - \int_{-\infty}^{0} \mathcal{M}\{\xi \le r\}dr \tag{9}$$

provided that at least one of the two integral is finite.

Theorem 1. (Liu 2010a) *Let* ξ *be an uncertain variable with regular uncertainty distribution* Φ. *Then*

$$E[\xi] = \int_0^{+\infty} \Phi^{-1}(\alpha)d\alpha. \tag{10}$$

Theorem 2. (Liu 2010a) *Let* ξ *and* η *be independent uncertain variables with finite expected values. Then for any real number* a *and* b, *we have*

$$E[a\xi + b\eta] = aE[\xi] + bE[\eta.] \tag{11}$$

Theorem 3. (Liu 2010b) *Let* $\xi_1, \xi_2, \cdots, \xi_n$ *be independent uncertain variables with regular uncertainty distributions* $\Phi_1, \Phi_2, \cdots, \Phi_n$, *respectively. If* $f(\xi_1, \xi_2, \cdots, \xi_n)$ *is strictly increasing with respect to* $\xi_1, \xi_2, \cdots, \xi_m$ *and strictly decreasing with respect to* $\xi_{m+1}, \xi_{m+2}, \cdots, \xi_n$, *then*

$$\xi = f(\xi_1, \xi_2, \cdots, \xi_n) \tag{12}$$

has an inverse uncertainty distribution

$$\Psi^{-1}(\alpha) = f(\Phi_1^{-1}(\alpha), \cdots, \Phi_m^{-1}(\alpha), \Phi_{m+1}^{-1}(1-\alpha), \cdots, \Phi_n^{-1}(1-\alpha)). \tag{13}$$

Theorem 4. (Liu 2010b) *Let* $\xi_1, \xi_2, \cdots, \xi_n$ *be independent uncertain variables with regular uncertainty distributions* $\Phi_1, \Phi_2, \cdots, \Phi_n$, *respectively. If constraint function* $f(x, \xi_1, \xi_2, \cdots, \xi_n)$ *is strictly increasing with respect to* $\xi_1, \xi_2, \cdots, \xi_m$ *and strictly decreasing with respect to* $\xi_{m+1}, \xi_{m+2}, \cdots, \xi_n$, *then*

$$\mathcal{M}\{f(x, \xi_1, \xi_2, \cdots, \xi_n) \le 0\} \ge \alpha \tag{14}$$

holds if and only if

$$f(x, \Phi_1^{-1}(\alpha), \cdots, \Phi_m^{-1}(\alpha), \Phi_{m+1}^{-1}(1-\alpha), \cdots, \Phi_n^{-1}(1-\alpha)) \leq 0. \tag{15}$$

3. UNCERTAIN TIME SERIES FORECASTING METHOD

Uncertain time series was proposed by Yang & Liu (2019) in 2019 so as to predict the future values based on previously imprecisely observed values that are described by uncertain variables. The basic definition of uncertain time series is as follows.

Definition 10. (Yang & Liu 2019) *An uncertain time series is a sequence of imprecisely observed values that are characterized in terms of uncertain variables. Mathematically, an uncertain time series is represented by*

$$X = \{X_1, X_2, \cdots, X_n\} \tag{16}$$

where X_t are imprecisely observed values (i.e. uncertain variables) at times t, $t = 1, 2 \cdots, n$, respectively.

After giving the uncertain time series, it is necessary to formulate function relations between the observations at time t and those at previous times to describe uncertain time series. Generally, the relationship between uncertain variables can be expressed by the following function

$$X_t = f(X_{t-1}, X_{t-2}, \cdots, X_{t-p}) + \varepsilon_t, \quad t = p+1, p+2, \cdots n \tag{17}$$

According to the research (Yang & Liu 2019), the method for modelling uncertain time series is the autoregressive model,

$$X_t = \varphi_0 + \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \cdots + \varphi_p X_{t-p} + \varepsilon_t, \quad t = p+1, p+2, \cdots, n \tag{18}$$

where $\varphi_0, \varphi_1, \cdots, \varphi_p$ are unknown parameters, ε_t is an uncertain variable, and p is called the order of the autoregressive model. In order to recognize a good UAR model as a forecast tool for the given data, the following three problems need to be solved which include: the determination of order of UAR model, parameter estimation and the construction of uncertain time series. In the following subsections, all the detailed procedure of using our methodology to make predictions is presented, which can be divided into four stages, clearly differentiated in Figure 1.

3.1. The determination of the order of UAR model

During the modeling process of uncertain time series, a fundamental phase is the identification of order of UAR model. Generally, we make predictions about the future to make strategies, which should not only get information from the data before but also get information from the near past, although they may not have the same effect strength. Therefore, it is crucial to find an appropriate order to determine the lagged variables existing in the model and establish the truly effective model. If the model order is not recognized efficiently, the accuracy of the predictions produced by the model will be compromised. Just like traditional time series analysis, correlation is a very important concept used in analyzing data in the time series. We often identify the model based on the trailing or truncating properties of the autocorrelation coefficient and the partial correlation coefficient. However, for the imprecisely observed values represented by uncertain variables, the traditional statistical method above is problematic. In this subsection, we introduce an order determination method based on the notion of similarity, which is generated from the distance measure defined by Li & Liu (2015). It can be represented by

$$D_p(\xi, \eta) = (E|\xi - \eta|^p)^{\frac{1}{p+1}}, p > 0 \tag{19}$$

where ξ and η are uncertain variables. It is easy to understand that the distance between uncertain variables essentially reflects their difference. The greater the distance is, the smaller the similarity is, and vice versa. Here, we just set out one, i.e. $p = 1$, as the application for the following definitions.

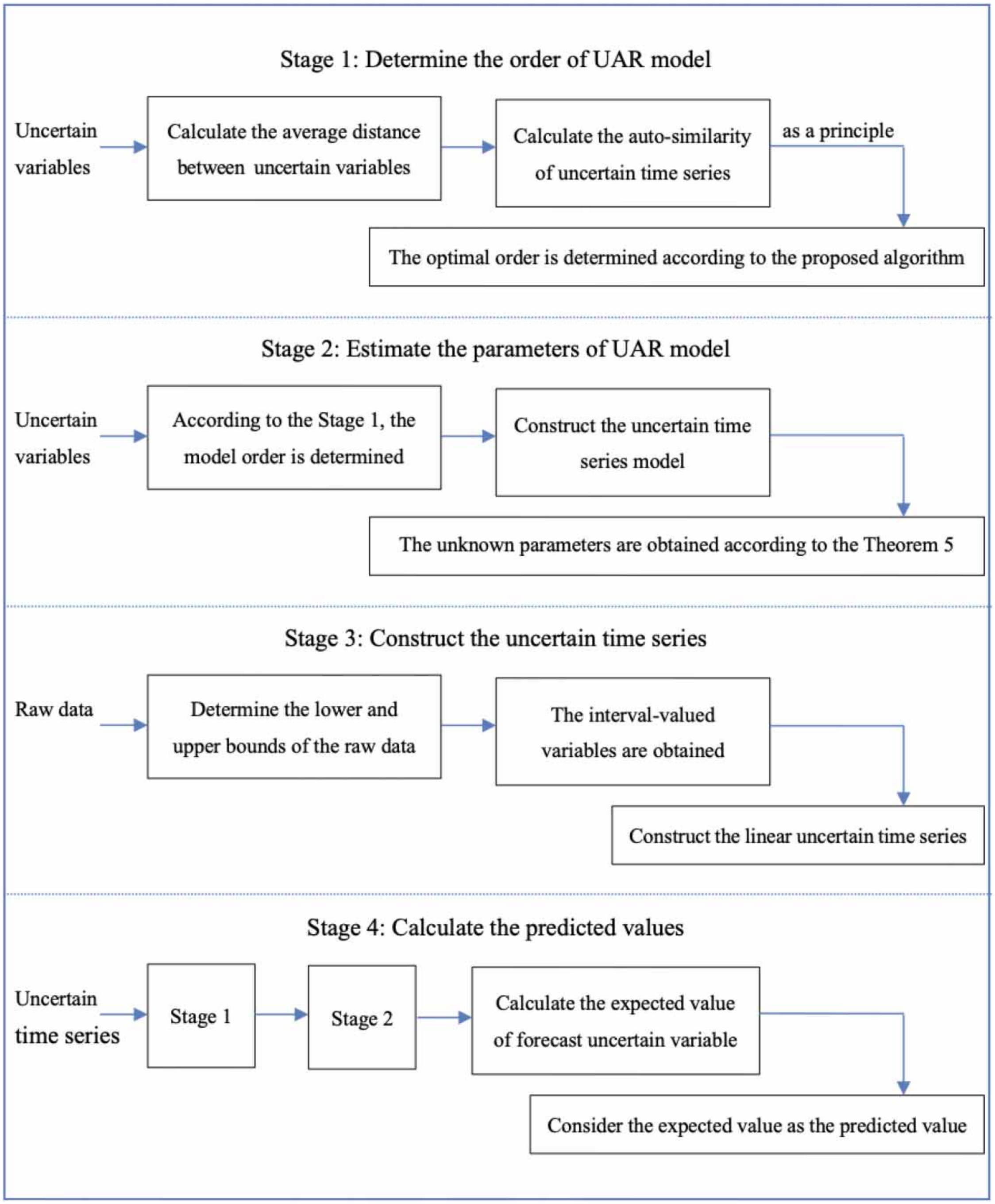

Figure 1 | The detailed four stages of the proposed uncertain time series method.

3.1.1. Basic definitions

Based on the above distance measure, the average distance between two uncertain variables in the uncertain time series is defined as below.

Definition 9. *Average distance*

Let $\{X_t\}(t = 1, 2, \cdots, n)$ *be an uncertain time series and* $X_1, X_2, \cdots, X_n$ *be imprecisely observed values characterized in terms of independent uncertain variables with regular uncertain distributions* $\Phi_1, \Phi_2, \cdots, \Phi_n$ *respectively, m denotes the* *experimental order. Then the average distance between uncertain variables* X_t *and* X_{t-k} $(k = 1, 2, \cdots m)$ *is defined as*

$$\begin{aligned} AD_k &= \frac{1}{n-k}\sum_{t=k+1}^{n} D_1(X_t, X_{t-k}) \\ &= \frac{1}{n-k}\sum_{t=k+1}^{n} [E|X_t - X_{t-k}|]^{\frac{1}{2}}, k = 1, 2, \cdots, m. \end{aligned} \quad (20)$$

To further provide ease of use, the study applies the following definition to make some adjustments.

Definition 10. *Auto-similarity of uncertain time series*

Let $AD_k(k = 1, 2, \cdots, m)$ *be the average distance between uncertain variables* X_t *and* X_{t-k} $(k = 1, 2, \cdots, m)$ *of uncertain time series*$\{X_t\}(t = 1, 2, \cdots, n)$, m *denotes the experimental order. Then the auto-similarity of uncertain time series is defined as*

$$AS_k = 1 - \frac{AD_k - \min_{1\le k\le m}\{AD_k\}}{\max_{1\le k\le m}\{AD_k\} - \min_{1\le k\le m}\{AD_k\}}, k = 1, 2, \cdots, m. \quad (21)$$

It is clear that $AS_k \in [0, 1]$, $k = 1, 2, \cdots, m$. Definition 10 shows that the greater AS_k is, the higher the similarity between the uncertain variables X_t and X_{t-k} is.

3.1.2. Model order selection algorithm

According to the above definitions, an algorithm for determining the appropriate order of UAR model is presented as follows.

Step 1. We set $k = 1$ as an alternative order.

Step 2. Determine the confidence level α.

Step 3. If $|AS_k - AS_{k+1}| \le \alpha$, then select the $(k+1)^{th}$ order to add the set of alternative order numbers. That is, enter into Step 4.

If $|AS_k - AS_{k+1}| > \alpha$ and $AS_k > AS_{k+1}$, then choose the k^{th} order as the optimal order.

If $|AS_k - AS_{k+1}| > \alpha$ and $AS_k < AS_{k+1}$, then into Step 4.

Step 4. Set $k = k + 1$ and return to Step 2.

Step 5. Optimal order obtained from Step 1 to Step 4 is regarded as the order of UAR model.

Step 6. If we cannot find the effective order until the predetermined experimental order m is reached, then let $k = 1$.

We want to note again that the proposed algorithm is also presented in Figure 2.

3.2. A new parameter estimation method based on uncertain programming

Once the order of the model is determined, we need to estimate the parameters of the UAR model to make predictions. Based on the imprecisely observed values, Yang & Liu (2019) investigated the least squares approach to estimate coefficients of the UAR model. Different from the previous research, in this subsection, we will propose a new method of parameter estimation based on uncertain programming, which can give more flexibility to the UAR model to make predictions. In general, we hope that the given parameters should make the differences between the predicted values $\hat{X}_t$ and observed values X_t as small as possible. In uncertain time series model, let $X_1, X_2, \cdots, X_n$ be imprecisely observed values characterized in terms of independent uncertain variables with regular uncertain distributions $\Phi_1, \Phi_2, \cdots, \Phi_n$, respectively. Then, the estimation of unknown parameters $\varphi_0, \varphi_1, \cdots, \varphi_p$ in the UAR model

$$\hat{X}_t = \varphi_0 + \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \cdots + \varphi_p X_{t-p} + \varepsilon_t, \quad t = p+1, p+2, \cdots, n \quad (22)$$

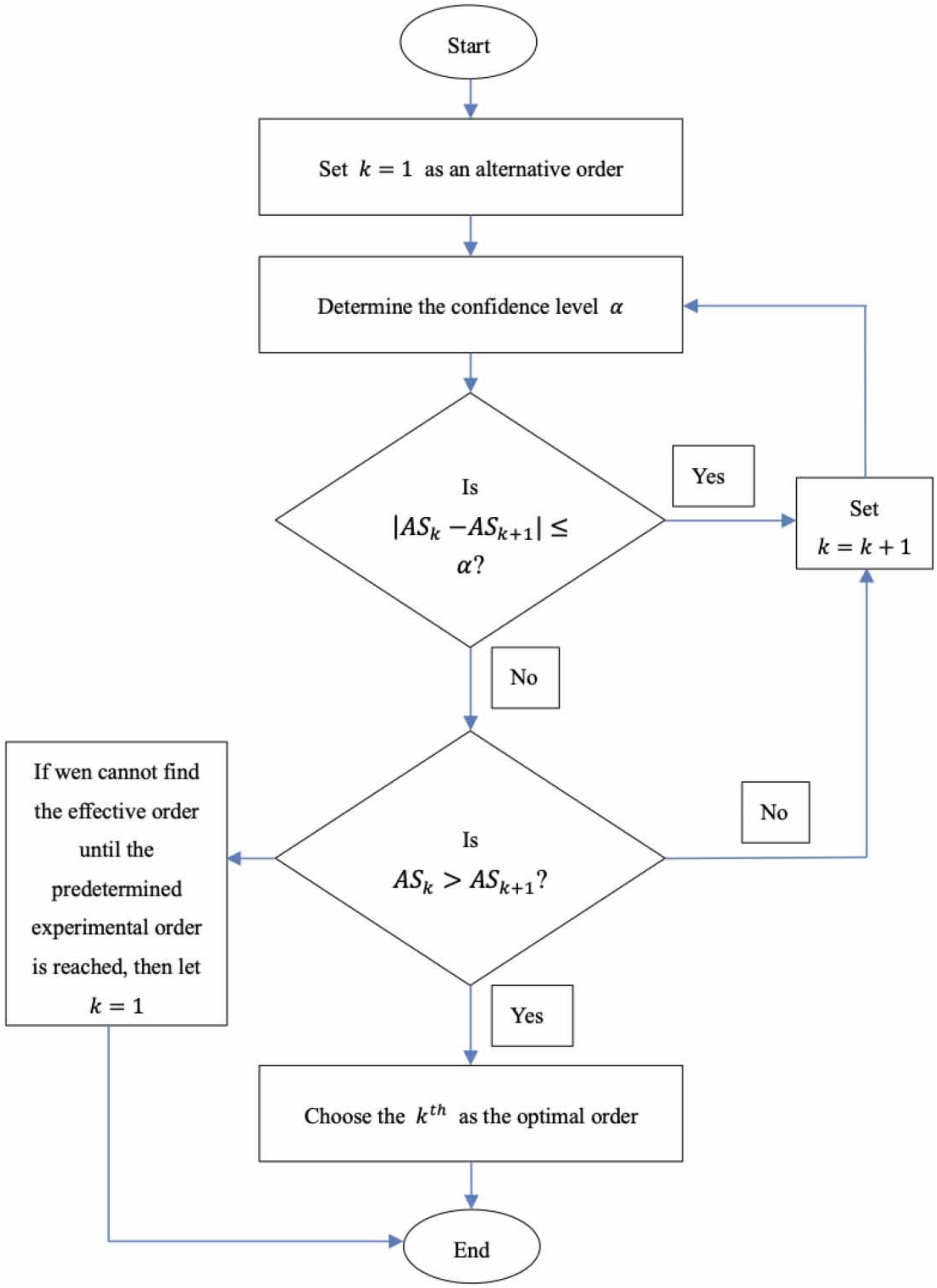

Figure 2 | Flowchart of the algorithm for determining the model order.

can solve the following programming model

$$\begin{cases} \min \sum_{t=p+1}^{n} a_t \\ subject\ to \\ \mathcal{M}\{(X_t - \hat{X}_t)^2 \leq a_t\} \geq \beta_t \\ t = p+1, p+2, \cdots, n \end{cases} \tag{23}$$

where a_t is the target variable and $\beta_t \in (0, 1)$ is a given level by the domain experts according to their experience knowledge.

In order to obtain the optimal solution, we need to transform it into an equivalent deterministic model. The following theorem will address this problem.

Theorem 5. *The model (23) is equivalent to the crisp mathematical programming*

$$\begin{cases} \min \sum_{t=p+1}^{n} a_t \\ \textit{subject to} \\ \left(\Phi_t^{-1}(\beta_t) - \varphi_0 - \sum_{i=1}^{p} \varphi_i \Upsilon_{t-i}^{-1}(\beta_t)\right)^2 - a_t \le 0 \\ t = p+1, p+2, \cdots, n \end{cases} \tag{24}$$

where

$$\Upsilon_{t-i}^{-1}(\beta_t, \varphi_i) = \begin{cases} \Phi_{t-i}^{-1}(1-\beta_t), & \text{if} \quad \varphi_i \ge 0 \\ \Phi_{t-i}^{-1}(\beta_t), & \text{if} \quad \varphi_i < 0 \end{cases} \tag{25}$$

for $i = 1, 2, \cdots, p$.

Proof: It follows from Theorem 4 immediately.

3.3. The construction of uncertain time series

Uncertain time series was proposed so as to deal with such forecasting problems where the historical data are not crisp numbers but are imprecisely observed values. Besides, in practical cases, most traditional point data possess uncertainty characteristics due to the measurement errors. For instance, the water demand variables that naturally take a finite set of numerical values varying between a lower and upper bound are regarded as interval-valued variables, which stand for the inaccuracies in measurements. For doing so, each interval-valued variable is assumed as the linear uncertain variable. However, for the same uncertain time series model, the difference of constructed intervals by adopting different ways can result in different forecasting performance. So, how to use an efficient way to choose effective length of interval is especially critical to improve uncertain time series forecasting performance. A key point in determining the proper length of interval is that they should not too large or small. When an effective length of interval is too wide, the prediction results will be meaningless in the uncertain time series. If the length is too small, the uncertain time series will become very close to the traditional time series and the result is not intended. On the other hand, many traditional time series have the momentum to vibrate in a certain period of time. Therefore, in the process of constructing uncertain time series, we should consider the trend information of data of time series itself, which makes the determined interval series more reasonable and really reflects the variation tendency of data of time series. By following the two requirements, in this subsection, we propose a new ratio-based approach to determine the length of interval to obtain the high forecasting accuracy. The step of the algorithm of the method presented can be given as follows:

Step 1. Take the first order of differences between any two consecutive observations $y_t - y_{t-1}$ for any y_t and y_{t-1}, $t = 2, 3, \cdots, n$.

Step 2. Calculate relative differences $r_t = (y_t - y_{t-1})/y_{t-1}$ for all t, $t = 2, 3, \cdots, n$.

Step 3. Determine the lower and upper bounds of the initial value.

Let

$$y_1^U - y_1^L = \frac{1}{n}\sum_{t=1}^{n}\left| y_t - \frac{1}{n}\sum_{t=1}^{n} y_t \right| \tag{26}$$

and

$$y_1^U + y_1^L = 2y_1. \tag{27}$$

Step 4. Determine the lower and upper bounds on the interval series.

Let

$$y_t^L = y_{t-1}^L \times (1 + r_t) \tag{28}$$

and

$$y_t^U = y_{t-1}^U \times (1 + r_t) \tag{29}$$
$$t = 2, 3, \cdots, n,$$

be the lower and upper bounds for observation y_t at time t, $t = 2, 3, \cdots, n$ respectively.

Step 5. Construct the uncertain time series. Following steps above, the interval-valued variables, $X_t = \{[y_t^L, y_t^U]: y_t^L, y_t^U \in R, y_t^L < y_t^U\}$, $t = 1, 2, \cdots, n$, are obtained. We can consider each interval variable X_t at time t as the linear uncertain variable with linear uncertain distribution $L(y_t^L, y_t^U)$. Then, an uncertain time series can be represented as

$$X = \{X_1, X_2, \cdots, X_n\}. \tag{30}$$

4. CASE STUDY

This section presents the application of the proposed methods for water demand forecast in Beijing. In Section 4.1, the location and dataset used in model development are given. Section 4.2 provides the implementations of the proposed uncertain time series model. For the purpose of comparison, the ARIMA model is selected to contrast the forecasting performance, and the classical measure methods are adopted to evaluate the forecast accuracy of the models in Section 4.3.

4.1. Location and dataset

In this work, the study area is located in Beijing. As the capital of China, Beijing is China's political, cultural, and international communication center, located at the interlaced terrace of North China Plain and Mongolian Plateau. In Beijing, drinking water is mainly supplied by the Yongdinghe and Chaobaihe rivers. As a result of China's rapid development and dense population, Beijing's water demand consumption is increasing rapidly and Beijing is experiencing a shortage of water resources. According to the Beijing Water Authority (BWA), the annual water resources per capita is less than 300 m^3, which is only 12.5% of the national average and far below the internationally recognized minimum standard of 1000 m^3 per year. This situation of increasing water demands and limited water resource supplies has also become the vital restrictive factor affecting the socio-economic development and environmental health of Beijing for a long time into the future. Therefore, it is particularly important to apply the proposed method to forecast the water demand in Beijing, which contributes to the water resources planning and management in the near future.

All the research dataset of uncertain time series were obtained from Beijing Water Resources Bulletin during the time period between 1988 and 2016. A total of 29 data points were collected and are shown in Table 1. In order to illustrate the effectiveness of the proposed uncertain time series method, the data from 1988 to 2013 are used as an estimation sample to determine the coefficients of the estimation model, while the rest of data are reserved as the hold-out sample, used to test the model and access the performance of prediction.

4.2. Methodologies implementations

4.2.1. Construction of uncertain time series

As mentioned previously, water demand data are not crisp numbers but imprecisely observed values, so the data pre-processing is essential. For the purpose of implementation, we utilize the linear uncertain variables to describe the water demand observations and construct the uncertain time series by following the algorithm in the previous subsection:

(1) Take the first order of differences between any two consecutive observations, which are listed in the fourth column of Table 1.
(2) Calculate the relative differences between any two consecutive observations. All the relative differences are listed in the fifth column of Table 1.
(3) Determine the lower and upper bounds of the interval-valued series. Firstly, the initial interval is calculated as follows.

Table 1 | Imprecisely observed data where $L(a, b)$ represents linear uncertain variable

Observation t	Time	Total water demand $y_t(10^7 m^3)$	First order difference	Relative difference	y_t^L	y_t^U	X_t
1	1988	424			406	442	$L(406, 442)$
2	1989	446	22	0.05189	427	465	$L(427, 465)$
3	1990	411	−35	−0.0785	394	428	$L(394, 428)$
4	1991	423	12	0.0292	405	441	$L(405, 441)$
5	1992	464	41	0.0969	444	484	$L(444, 484)$
6	1993	452	−12	−0.0259	433	471	$L(433, 471)$
7	1994	459	7	0.0155	440	478	$L(440, 478)$
8	1995	449	−10	−0.0218	430	468	$L(430, 468)$
9	1996	400	−49	−0.1091	383	417	$L(383, 417)$
10	1997	403	3	0.0075	386	420	$L(386, 420)$
11	1998	404	1	0.0025	387	421	$L(387, 421)$
12	1999	417	13	0.0322	399	435	$L(399, 435)$
13	2000	400	−17	−0.0408	383	417	$L(383, 417)$
14	2001	389	−11	−0.0275	372	406	$L(372, 406)$
15	2002	346	−43	−0.1105	331	361	$L(331, 361)$
16	2003	358	12	0.0347	343	373	$L(343, 373)$
17	2004	346	−12	−0.0335	331	361	$L(331, 361)$
18	2005	345	−1	−0.0029	330	360	$L(330, 360)$
19	2006	343	−2	−0.0058	328	358	$L(328, 358)$
20	2007	348	5	0.0146	333	363	$L(333, 363)$
21	2008	351	3	0.0086	336	366	$L(336, 366)$
22	2009	355	4	0.0114	340	370	$L(340, 370)$
23	2010	352	−3	−0.0085	337	367	$L(337, 367)$
24	2011	360	8	0.0227	345	375	$L(345, 375)$
25	2012	359	−1	−0.0028	344	374	$L(344, 374)$
26	2013	364	5	0.01393	349	379	$L(349, 379)$

By

$$\begin{aligned} y_1^U - y_1^L &= \frac{1}{n}\sum_{t=1}^{n}\left|y_t - \frac{1}{n}\sum_{t=1}^{n} y_t\right| \\ &= \frac{1}{26}\sum_{t=1}^{26}\left|y_t - \frac{1}{26}\sum_{t=1}^{26} y_t\right| \\ &= \frac{1}{26}\sum_{t=1}^{26}|y_t - 391| \\ &= 36 \end{aligned} \tag{31}$$

and

$$\begin{aligned} y_1^U + y_1^L &= 2 \times y_1 \\ &= 848. \end{aligned} \tag{32}$$

we have

$$y_1^L = 406, y_1^U = 442. \tag{33}$$

Then, it is obvious that $X_1 = [406, 442]$. We consider the initial interval-valued variable X_1 as the linear uncertain variable with linear uncertain distribution $L(y_1^L, y_1^U)$. Secondly, by following Step 3 in Section 3.3, the corresponding interval series can be constructed as shown in Table 1. Thus, we can obtain an uncertain time series that are characterized in terms of linear uncertain variables, i.e.

$$X = \{X_1, X_2, \cdots, X_{26}\}. \tag{34}$$

4.2.2. The determination of the model order

In the model order selection phase, different experimental orders are examined based on the definition of auto-similarity of uncertain time series and the best one among them is selected. Generally, we set the maximum lagging order $m = 5$. The details of the calculation can be found in the following paragraphs.

Firstly, we assume that linear uncertain variables $X_1, X_2, \cdots, X_{26}$ are independent. According to Definition 9, the average distance of the first order lag of all uncertain variables is calculated as follows

$$\begin{aligned} AD_1 &= \frac{1}{25}\sum_{t=2}^{26} D_1(X_t, X_{t-1}) \\ &= \frac{1}{25}\sum_{t=2}^{26} [E|X_t - X_{t-1}|]^{\frac{1}{2}} \\ &= \frac{1}{25}\sum_{t=2}^{26}\left[\int_0^1 |\Phi_t^{-1}(\alpha) - \Phi_{t-1}^{-1}(1-\alpha)|d\alpha\right]^{\frac{1}{2}} \\ &= 3.2075. \end{aligned} \tag{35}$$

Similarly, we have,

$$AD_2 = \frac{1}{24}\sum_{t=3}^{26} D_1(X_t, X_{t-2}) = 3.5781, \tag{36}$$

$$AD_3 = \frac{1}{23}\sum_{t=4}^{26} D_1(X_t, X_{t-3}) = 4.2078, \tag{37}$$

$$AD_4 = \frac{1}{22}\sum_{t=5}^{26} D_1(X_t, X_{t-4}) = 4.7491 \tag{38}$$

and

$$AD_5 = \frac{1}{21}\sum_{t=6}^{26} D_1(X_t, X_{t-5}) = 5.1953. \tag{39}$$

Then, $\min_{1\leq k\leq 5}\{AD_k\} = AD_1$, $\max_{1\leq k\leq 5}\{AD_k\} = AD_5$. It follows from Definition 10 that

$$AS_1 = 1, AS_2 = 0.8136, AS_3 = 0.4968, AS_4 = 0.2245, AS_5 = 0. \tag{40}$$

By using the algorithm in Section 3.1, we can find the appropriate order.

Let $\alpha = 0.2$. Because $|AS_1 - AS_2| = 0.1864 < 0.2 = \alpha$ and $|AS_2 - AS_3| = 0.3186 > 0.2 = \alpha$. At the same time $AS_2 = 0.1836 > 0.4968 = AS_3$, thus the order of UAR model is $k = 2$.

4.2.3. The parameter estimation for the UAR model

According to Section 4.2.2, we obtain the 2-order UAR model

$$\hat{X}_t = \varphi_0 + \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \varepsilon_t, \quad t = 3, 4, \cdots, 26. \tag{41}$$

By using Theorem 5, we obtain the following mathematical programming model to estimate the unknown coefficients $\varphi_i (i = 0, 1, 2)$.

$$\begin{cases} \min \sum_{t=3}^{26} a_t \\ \textit{subject to} \\ ((35 + 38 \times \varphi_1 + 36 \times \varphi_2) \times \beta_3 + 394 - \varphi_0 - 465 \times \varphi_1 - 442 \times \varphi_2)^2 \leq a_3 \\ ((36 + 35 \times \varphi_1 + 38 \times \varphi_2) \times \beta_4 + 405 - \varphi_0 - 428 \times \varphi_1 - 465 \times \varphi_2)^2 \leq a_4 \\ \vdots \\ ((31 + 30 \times \varphi_1 + 31 \times \varphi_2) \times \beta_{26} + 349 - \varphi_0 - 374 \times \varphi_1 - 375 \times \varphi_2)^2 \leq a_{26} \\ \varphi_0 > 0, \varphi_1 > 0, \varphi_2 > 0. \end{cases} \tag{42}$$

Let $\beta_t = 0.75$, $(t = 3, 4, \cdots, 26)$. Then we obtain the optimal solution

$$(\varphi_0, \varphi_1, \varphi_2) = (44.90, 0.81, 0.10) \tag{43}$$

and the corresponding autoregressive model

$$\hat{X}_t = 44.90 + 0.81X_{t-1} + 0.10X_{t-2}, \quad t = 3, 4, \cdots, 26. \tag{44}$$

According to different confidence levels (i.e. 0.75, 0.85, 0.90), different regression models can be established to compare with the existing traditional time series model.

4.3. Comparison with the existing method

In this subsection, we apply the proposed UAR model to forecast the water demand of Beijing from 2014 to 2016. Following the above forecasting model (Equation (44)), the forecasting result is an uncertain variable. However, in most cases, the results we require are often crisp values, so the research uses the expected value of the uncertain variable as the predicted value. The prediction results under different confidence levels (i.e. 0.75, 0.85, 0.90) are presented in Table 2. In order to further verify the effectiveness of the proposed methodologies, the traditional ARIMA time series method is selected as a competitor to contrast the forecasting performance. According to the historical data of water demand from 1988 to 2013, the result of ARIMA model is represented as follows

$$\hat{X}_t = X_{t-1} + \varepsilon_t. \tag{45}$$

The performances of the models are evaluated based on the classical measure methods, i.e. average relative error (ARE) and total absolute error (AE). Obviously, lower ARE and AE values lead to better performance. The definitions of all

Table 2 | Predicted results of Beijing's total water demand

Time	Actual values	ARIMA model	UAR model		
			$\beta_t = 0.75$	$\beta_t = 0.85$	$\beta_t = 0.90$
2014	375	361.59	375.64	383.74	385.74
2015	382	359.17	385.57	400.56	404.32
2016	388	356.75	398.53	416.54	422.04

Table 3 | Comparison of above models

Time	Absolute error				Relative error			
		UAR model				UAR model		
	ARIMA model	$\beta_t = 0.75$	$\beta_t = 0.85$	$\beta_t = 0.90$	ARIMA model	$\beta_t = 0.75$	$\beta_t = 0.85$	$\beta_t = 0.90$
2014	13.41	0.64	8.74	10.74	0.0358	0.0017	0.0233	0.0286
2015	22.83	3.57	18.56	22.32	0.0598	0.0093	0.0486	0.0584
2016	31.25	10.53	28.54	34.04	0.0805	0.0271	0.0736	0.0877
Total predicted error	67.49	14.74	55.84	67.09				
Maximum relative error					0.0805	0.0271	0.0736	0.0877
Average relative error					0.0587	0.0127	0.0485	0.0583

these performance criteria are represented by

$$ARE = \frac{1}{s}\sum_{j=1}^{s}\left|\frac{\hat{y}_j - y_j}{y_j}\right| \tag{46}$$

and

$$AE = \sum_{j=1}^{s}|\hat{y}_j - y_j| \tag{47}$$

where s is the total number of data needed to predict, $\hat{y}_j$ and y_j denote the predicted value and the actual observed value, respectively.

Therefore, the prediction performance are shown in Table 3 and Figure 3. From the experimental results obtained, it can be concluded that the proposed uncertain time series forecasting method has better forecasting performance than the ARIMA method under the considered levels.

As far as the comparison between the proposed method under the 0.75 and 0.85 confidence levels and the ARIMA method is concerned, the former outperforms the latter in all cases. The total prediction error is reduced by 78.15 and 17.26% respectively, and the average relative error by 78.36 and 17.38% respectively. In addition, for the prediction error of each observation, the proposed method under the 0.75 and 0.85 levels is smaller than the ARIMA method. Especially for the associated 0.75 confidence level, the improvements of forecasting performance are more obvious.

When considering the comparison between the proposed method under the 0.90 confidence level and the ARIMA method in all cases, we can see that the maximum prediction error of the proposed method is a little higher than the ARIMA method. Overall, the former almost wins. The total prediction error is reduced by 0.59% and the average relative error by 0.68%. This reduction is crucial in the planning and management of water supply systems.

Apart from the statistical criteria discussed above, we implement the forecasting trend to evaluate the performance of the above methods. According to the current trend, Beijing's total water demand was increasing from 2014 to 2016. However, the prediction trend of the ARIMA method was declining. This is not in accordance with the reality of Beijing's total water

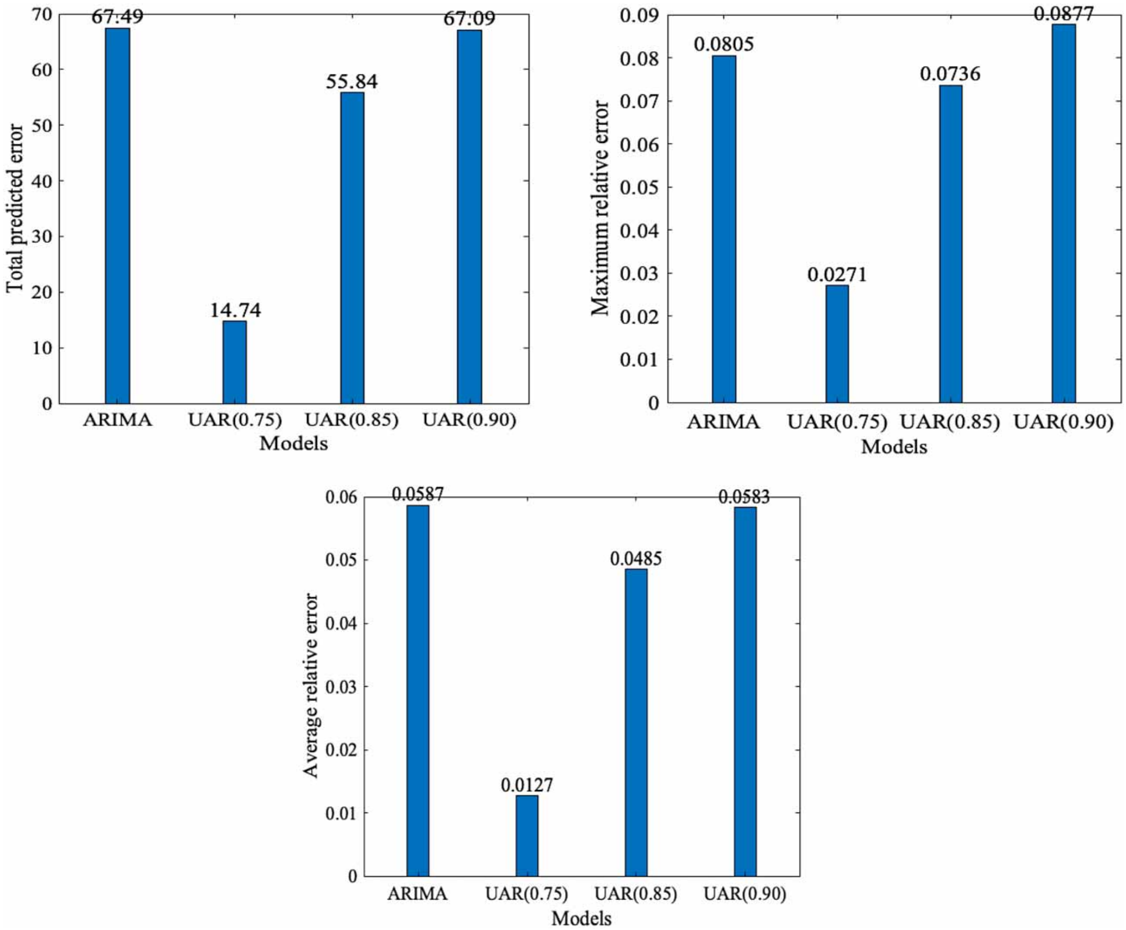

Figure 3 | Comparisons of prediction accuracy for water demand prediction.

demand during 2014 and 2016. So, these predicted results may not provide support to the water resource management in the near future. It is worthy to note that the predicted results of the proposed method could reflect the realistic water demand trend in the short term and help the decision makers to devise reasonable management schemes.

5. CONCLUSION

In this study we presented a modified time series method for demand estimation of water resources in Beijing. Considering the uncertainty of water demand in real life, we attempted to combine the uncertainty theory with a time series model, called uncertain time series, to handle the above problems. In the presented method, we employed the UAR model to describe uncertain time series for predicting future values. First, the auto-similarity of uncertain time series, as a principle of justifiable recognition, is defined and the identification algorithm of determining the optimal model order is proposed, which enables the estimation of the correct parameters of the model. Second, we propose an uncertain programming approach for estimating the parameters of the model. Then, the imprecisely observed values are assumed as the linear uncertain variables and a ratio-based method is presented for constructing the uncertain time series. Finally, we tested the performance of the proposed model and the traditional time series model (ARIMA) based on the statistical criteria. The results demonstrated that the proposed model provided much better accuracy over the traditional model mentioned above for water demand predictions. The possible reason is that the traditional model cannot effectively handle the imprecisely observed values, this allows the possibility of the loss of effective information, which leads to the reduction of prediction accuracy.

Although the proposed UAR model has greatly improved the traditional water demand time series method, there are still some limitations which need to be improved, such as the determination of the model order and the construction of the uncertain time series. In the future, we will further study the algorithm optimization of model order, and provide better solutions for the UAR model applications which improve the accuracy of forecast. On the other hand, we only investigate the construction of linear uncertain time series due to the interval-valued data that are encountered frequently in multiple situations. Further study may attempt to construct the normal uncertain time series to expand the application field of the model.

ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China (No.61873084) and the Foundation of Hebei Education Department (No. ZD2017016).

DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

REFERENCES

Akaike, H. 1970 Statistical predictor identification. *Annals of Institute of Statistical Mathematics* **22**, 203–217.

Akaike, H. 1974 A new look at the statistical model identification. *IEEE Transactions on Automatic Control* **AC-19** (6), 716–723.

Almutaz, I., Ajbar, A., Khalid, Y. & Ali, E. 2012 A probabilistic forecast of water demand for a tourist and desalination dependent city: case of Mecca, Saudi Arabia. *Desalination* **294**, 53–59.

Al-Smadi, A. & Al-Zaben, A. 2005 ARMA model order determination using edge detection: a new perspective. *Circuits, Systems and Signal Processing* **24** (6), 723–732.

Aly, A. & Wanakule, N. 2004 Short-term forecasting for urban water consumption. *Water Resources Planning and Management* **130**, 405–410.

Arnell, N. & Lloyd-Hughes, B. 2014 The global-scale impacts of climate change on water resources and flooding under new climate and socio-economic scenarios. *Climatic Change* **122**, 127–140.

Choksi, K. N., Sheth, M. A. & Mehta, D. 2015 To assess the performance of Sewage Treatment Plant: a case study of Surat city. *International Journal of Engineering and Technology (IRJET)* **2** (08), 1071–1075.

Deng, W., Wang, G. & Zhang, X. 2015 A novel hybrid water quality time series prediction method based on cloud model and fuzzy forecasting. *Chemometrics and Intelligent Laboratory Systems* **149**, 39–49.

Frederick, K. D. 1997 Adapting to climate impacts on the supply and demand for water. *Climatic Change* **37** (1), 141–156.

Fuchs, J. J. 1987 On estimating the order of an ARMA process. *Automatica* **23** (6), 779–782.

Guo, H., Wang, X. & Gao, Z. 2017 Uncertain hypothesis testing for two experts empirical data. *Mathematical and Computer Modelling* **55**, 1478–1482.

Haque, M. M., Rahman, A., Hagare, D. & Kibria, G. 2014 Probabilistic water demand forecasting using projected climatic data for Blue Mountains water supply system in Australia. *Water Resources Management* **294**, 53–59.

Huarng, K. 2006 Ratio-based lengths of intervals to improve fuzzy time series forecasting. IEEE transactions on systems. *Man, and Cybernetics-Part B: Cybernetics* **36**, 328–340.

Li, X. & Liu, Y. 2015 Distance and similarity measures between uncertain variables. *Journal of Intelligent and Fuzzy Systems* **28** (5), 2073–2081.

Liang, G., Wilkes, D. M. & Cadzow, J. A. 1993 ARMA model order estimation based on the eigenvalues of the covariance matrix. *IEEE Transactions on Signal Processing* **41** (10), 3003–3009.

Lio, W. & Liu, B. 2018 Residual and confidence interval for uncertain regression model with imprecise observations. *Journal of Intelligent and Fuzzy Systems* **35** (2), 2573–2583.

Liu, B. 2007 *Uncertain Theory*, 2nd edn. Springer-Verlag, Berlin.

Liu, B. 2009 Some research problems in uncertain theory. *Journal of Uncertain Systems.* **3** (1), 3–10.

Liu, B. 2010a *Uncertain Theory: A Branch of Mathematics for Modeling Human Uncertainty*. Springer-Verlag, Berlin.

Liu, B. 2010b Uncertain risk analysis and uncertain reliability analysis. *Journal of Uncertain Systems* **4** (3), 163–170.

Liu, X., Hen, Y., Guo, Y., Li, S. & Guo, B. 2015 Modeling demand/supply of water resources in the arid region of northwestern China during the late 1980s to 2010. *Geographical Sciences* **25** (5), 573–591.

Maidment, D. R., Miaou, S. P. & Crawford, M. M. 1985 Transfer function models of daily urban water use. *Water Resources Research* **21** (4), 425–432.

Oduro-Kwarteng, S., Nyarko, K., Odai, S. & Aboagye-Sarfo, P. 2009 Water conservation potential in educational institutions in developing countries: case study of a university campus in Ghana. *Journal of Urban Water* **6** (6), 449–455.

Pahl-Wostl, C. 2007 Transitions towards adaptive management of water facing climate and global change. *Water Resources Management* **21** (1), 49–62.

Rissanen, J. 1978 Modeling by shortest data description. *Automatica* **14**, 465–471.

Sadabadi, M. S., Shafiee, M. & Karrari, M. 2007 Determination of the two-dimensional ARMA model order using rank test based approach. *Proceedings of 16th IEEE International Conference on Control Applications*, 1156–1160

Sadabadi, M. S., Shafiee, M. & Karrari, M. 2009 Two-dimensional ARMA model order determination. *ISA Transactions* **48**, 247–253.

Schwarz, G. 1978 Estimating the dimension of a model. *Annals of Statistics* **6**, 461–464.
Smith, J. A. 1988 A model of daily municipal water use for short-term forecasting. *Water Resources Research* **24** (2), 153–164.
Wang, X. & Peng, Z. 2014 Method of moments for estimating uncertainty distribution. *Journal of Uncertainty Analysis and Applications* **2** (1), 1–10.
Wang, X., Gao, Z. & Guo, H. 2012a Delphi method for estimating uncertainty distributions. *Information: An International Interdisciplinary Journal* **15** (2), 449–460.
Wang, X., Gao, Z. & Guo, H. 2012b Uncertain linear regression model and its application. *Journal of Intelligent Manufacturing* **28** (3), 559–564.
Wang, X., Zhang, J., Shahid, S., He, R., Xia, X. & Mou, X. 2015 Potential impact of climate change on future water demand in Yulin city. *Northwest China. Mitigation and Adaptation Strategies for Global Change* **20** (1), 1–19.
Wang, X., Zhang, J., Shahid, S., Bi, S., Elmahdi, A., Liao, C. & Li, Y. 2018 Forecasting industrial water demand in Huaihe River Basin due to environmental changes. *Mitigation and Adaptation Strategies for Global Change* **23**, 469–483.
Yang, X. & Liu, B. 2019 Uncertain time series analysis with imprecise observations. *Fuzzy Optimization and Decision Making.* **18** (3), 263–278.
Yao, K. & Liu, B. 2018 Uncertain regression analysis: an approach for imprecise observation. *Soft Computing* **22** (17), 5573–5578.
Ye, T. & Liu, B. 2021 Uncertain hypothesis test with application to uncertain regression analysis. *Fuzzy Optimization and Decision Making*. https://doi.org/10.1007/s10700-021-09365-w.
Ye, T. & Liu, Y. 2020 Multivariate uncertain regression model with imprecise observations. *Journal of Ambient Intelligence and Humanized Computing*. https://doi.org/10.1007/s12652-020-01763-z.
Zhai, Y., Wang, J., Teng, Y. & Zuo, R. 2012 Water demand forecasting of Beijing using the time series forecasting method. *Geographical Sciences* **22** (5), 919–932.

First received 25 June 2021; accepted in revised form 9 November 2021. Available online 22 November 2021