Chaos, Solitons and Fractals 165 (2022) 112769

Contents lists available at ScienceDirect

Chaos, Solitons and Fractals

journal homepage: www.elsevier.com/locate/chaos

Parameter estimation of uncertain differential equation by implementing an optimized artificial neural network

Idin Noorani, Farshid Mehrdoust *

Department of Applied Mathematics, Faculty of Mathematical Sciences, University of Guilan, P.O. Box: 41938-1914, Rasht, Iran

ARTICLE INFO

Keywords:
Artificial neural network
Liu process
Nelder–Mead optimization method
Parameter estimation
Uncertainty theory

ABSTRACT

This study suggests a novel method for estimation of uncertain stock model parameters driven by Liu process. The proposed method decomposes the parameter estimation problem into two sub-problems: the first sub-problem implements an optimized artificial neural network based on the observed data, and the next sub-problem estimates the uncertain model parameters according to the optimized artificial neural network. We apply Nelder–Mead algorithm to optimize the artificial neural network and parameter estimation problem. The main supremacy of the presented method is that the estimation problem is independent of time intervals among observations and can be used to model future data. Providing a comparative method shows that the proposed approach can be effective for non-linear problems in which the artificial neural network structures perform well.

1. Introduction

Uncertainty and probability theories are two systems in the mathematical sciences that logically examine indefiniteness. As pointed out by Liu [1], probability theory is appropriate for frequency issues, while uncertainty theory is convenient to dealing with belief degrees. Based on a determined quantity whose distribution function is definite, it must be investigated whether the efficiency of the uncertainty theory is appropriate or the probability theory. In Liu [2], a persuasive example according to the uncertainty theory is presented to demonstrate that the probability theory must be used, if the considered distribution function can properly fit the frequency of the observed data; otherwise, the uncertainty theory is suitable. Regrettably, the desired distribution function cannot fit the frequency of observations well on most empirical studies. Thus, uncertainty theory is a significant mathematical instrument for time series analysis. The main subjects of uncertainty theory are: uncertain finance [3–5], uncertain energy [6,7], uncertain risk analysis [8] and uncertain logic [9].

Development and utilization of high-constancy models in the field of financial market modeling requires their advanced control and design. A critical phase in these models development is solving the parameter estimation problem. Problem-solving methods for estimation of parameter can be arranged according to the type of system model. Types of models are presented by equations that contain only deterministic expressions or those that stochastic term or uncertain expressions. As being one of the most important steps in modeling processes, estimation problems associated with the parameterized uncertain differential equations (UDEs) have been researched over recent years. Yao and Liu [10] presented a calibration method of moments. Sheng et al. [11] proposed a least squares estimation. Yang et al. [12] proposed a minimum cover estimation. Liu and Liu [13] used the maximum likelihood estimation. Liu [14] used the generalized moment estimation. Liu and Liu [15] considered the method of moments based on residuals. In these methods, UDEs have been used to model real-life problems such as pharmacokinetics [16], chemical reaction [17] and epidemic spread [18].

The parameter estimation of UDE according to the method of moments is based on difference scheme. However, when the time intervals between observations are not short enough, this method is not applicable. From this view point, Liu and Liu [15] recently introduced the method of moments based on residuals to estimate the parameters of UDE. After that, Mehrdoust et al. [7] applied this method to estimate the parameters of uncertain energy model. In this study, we intend to eliminate the factor of time interval between observations that appeared in the parameter estimation problem based on the method of moments. For this purpose, we suggest a method in which a dataset with future observations is first produced by the artificial neural network (ANN) structure and then using the new generated dataset, we estimate the UDE parameters. In this approach, instead of variations of observations relative to their step size, we find an alternative expression based on the ANN structure without the step size term. Other advantages of using ANNs include the fact that they

* Corresponding author.
E-mail address: far.mehrdoust@gmail.com (F. Mehrdoust).

https://doi.org/10.1016/j.chaos.2022.112769
Received 1 August 2022; Received in revised form 30 September 2022; Accepted 30 September 2022
Available online 19 October 2022

are used effectively to approximate non-linear functions and can be trained for multi-dimensional variables [19]. Mathematically, ANN is a functional approximation which has a behavioral pattern similar to the human brain. Theoretically, the ANN network can be trained based on experience and given the appropriate number of hidden layers and sufficient computational sources, and can approximate any complex function.

Another important advantage of ANN is that using this structure, it is possible to estimate the parameters of the desired model from unobserved data (forecast data) in order to model the future data. The use of ANN to estimate stochastic differential equation (SDE) parameters [20], ordinary differential equation parameters [21], partial differential equation parameters [22] and ecological system parameters [23] was investigated by the researchers over recent years. However, using ANN structure to estimate the UDE parameters as a significant approach for modeling natural phenomena has not been studied so far. Therefore, in this paper, we intend to fill this gap by applying the optimized ANN to estimate the parameters of several well-known UDEs. To do this, we apply Nelder–Mead (NM) optimization method and evaluate its performance based on numerical results. For this purpose, we use UDE parameters and their output values from the solutions of the equations to train neural networks, and use the trained networks to estimate the UDE parameters for given output data.

The rest of the paper is organized as follows. In the next section some required basic definitions of the uncertain process and the parameter estimation of UDE driven by Liu process based on the method of moments is proposed. Section 3 suggests the presented methodology for estimation of parameter based on the optimized ANN. Empirical studies and investigation of the parameter estimation problem on highly noisy environments are given in Sections 4 and 5, respectively. Finally, conclusion of this study is reported in Section 6.

2. Preliminary

In this section, following the method of moments presented by Yao and Liu [10], we generally provide the parameter estimation problem of the UDE. For this purpose, we first present some required basic definitions of the uncertain process. Uncertainty theory was founded by Liu [24] in 2007 and refined by Liu [25] in 2009.

Definition 2.1. Let $\mathcal{L}$ be a σ-algebra on a nonempty set Γ. A set function $\mathcal{M} : \mathcal{L} \longrightarrow [0,1]$ is called an uncertain measure if it satisfies the following axioms

- (Normality Axiom) $\mathcal{M}\{\Gamma\} = 1$ for the universal set Γ.
- (Duality Axiom) $\mathcal{M}\{\Lambda\} + \mathcal{M}\{\Lambda^C\} = 1$ for any event Λ.
- (Subadditivity Axiom) For every countable sequence of events $\Lambda_1, \Lambda_2, \ldots$, we have
  $$\mathcal{M}\{\bigcup_{i=1}^{\infty} \Lambda_i\} \le \sum_{i=1}^{\infty} \mathcal{M}\{\Lambda_i\}.$$
  The triple $(\Gamma, \mathcal{L}, \mathcal{M})$ is called an uncertainty space.
- (Product Axiom) Let $(\Gamma_i, \mathcal{L}_i, \mathcal{M}_i)$ be uncertainty spaces for $i = 1, 2, \ldots, n$. Then, the product uncertain measure $\mathcal{M}$ is an uncertain measure of product on σ-algebra $\mathcal{L}_1 \times \mathcal{L}_2 \times \cdots \times \mathcal{L}_n$ satisfying
  $$\mathcal{M}\{\prod_{i=1}^{n} \Lambda_i\} = \min_{1\le i\le n} \mathcal{M}\{\Lambda_i\},$$
  where Λ_i are arbitrarily chosen events from $\mathcal{L}_i$ for $i = 1, 2, \ldots$ respectively.

An uncertain variable ξ is a measurable function from an uncertainty space $(\Gamma, \mathcal{L}, \mathcal{M})$ to the set $\mathbb{R}$ of real numbers. Its uncertainty distribution Φ is defined as $\Phi(x) = \mathcal{M}\{\xi \le x\}$ for any real number x. If the inverse function Φ^{-1} exists and is unique for each $\alpha \in (0,1)$, then it is called the inverse uncertainty distribution of ξ.

Definition 2.2. An uncertain variable ξ is called normal if it has a normal uncertainty distribution

$$\Phi(x) = \left(1 + \exp\{\frac{\pi(\mu - x)}{\sqrt{3}\sigma}\}\right)^{-1}, \quad x \in \mathbb{R},$$

denoted by $\mathcal{N}(\mu, \sigma)$ where μ and σ are real numbers with $\sigma > 0$.

Definition 2.3. Let ξ be an uncertain variable, and k be a positive integer. Then the kth moment of ξ is defined by

$$E\left[\xi^k\right] = \int_0^{+\infty} \mathcal{M}\left\{\xi^k \ge r\right\} \mathrm{d}r - \int_{-\infty}^{0} \mathcal{M}\left\{\xi^k \le r\right\} \mathrm{d}r,$$

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

Liu [26] stated that if $\Phi^{-1}(\alpha)$ is an uncertainty distribution of ξ, then

$$E\left[\xi^k\right] = \int_0^1 \left(\Phi^{-1}(\alpha)\right)^k \mathrm{d}\alpha.$$

Definition 2.4. An uncertain process C_t is called a Liu process if

- $C_0 = 0$ and almost all sample paths are Lipschitz continuous,
- C_t has stationary and independent increments,
- the increment $C_{s+t} - C_s$ has a normal uncertainty distribution
  $$\Phi_t(x) = \left(1 + \exp\left(-\frac{\pi x}{\sqrt{3}t}\right)\right)^{-1}, \quad x \in \mathbb{R}.$$

Definition 2.5. Suppose that C_t is a Liu process, and f and g are two measurable real functions. Then

$$dX_t = f(t, X_t)dt + g(t, X_t)dC_t, \tag{2.1}$$

is called an UDE.

Definition 2.6 (*[27]*)**.** The α-path $(0 < \alpha < 1)$ of the UDE (2.1) with an initial value X_0 is a deterministic function X_t^α with respect to t that solves the corresponding ordinary differential equation

$$dX_t^\alpha = f(t, X_t^\alpha)dt + \mid g(t, X_t^\alpha) \mid \Phi^{-1}(\alpha)dt, \quad X_0^\alpha = X_0,$$

where $\Phi^{-1}(\alpha)$ is the inverse uncertainty distribution of standard normal uncertain variable, i.e.,

$$\Phi^{-1}(\alpha) = \frac{\sqrt{3}}{\pi} \ln \frac{\alpha}{1-\alpha}, \quad 0 < \alpha < 1.$$

Consider an UDE

$$dX_t = f(t, X_t, \theta)dt + g(t, X_t, \theta)dC_t, \tag{2.2}$$

where $\theta := (\theta_1, \ldots, \theta_p) \in \mathbb{R}^p$ are unknown parameters that should be estimated. On the other hand, the difference form of Eq. (2.2) is

$$X_{t_{i+1}} = X_{t_i} + f\left(t_i, X_{t_i}; \theta\right)\left(t_{i+1} - t_i\right) + g\left(t_i, X_{t_i}; \theta\right)\left(C_{t_{i+1}} - C_{t_i}\right)$$

which can be rewritten as

$$\frac{X_{t_{i+1}} - X_{t_i} - f\left(t_i, X_{t_i}; \theta\right)\left(t_{i+1} - t_i\right)}{g\left(t_i, X_{t_i}; \theta\right)\left(t_{i+1} - t_i\right)} = \frac{C_{t_{i+1}}^{(h)} - C_{t_i}^{(h)}}{t_{i+1} - t_i}$$

According to Definition 2.4, the right term

$$\frac{C_{t_{i+1}} - C_{t_i}}{t_{i+1} - t_i} \sim \mathcal{N}(0, 1).$$

Therefore, we have

$$\frac{X_{t_{i+1}} - X_{t_i} - f\left(t_i, X_{t_i}; \theta\right)\left(t_{i+1} - t_i\right)}{g\left(t_i, X_{t_i}; \theta\right)\left(t_{i+1} - t_i\right)} \sim \mathcal{N}(0, 1) \tag{2.3}$$

Assume that there are N observations $x := (x_{t_1}, x_{t_2}, \ldots, x_{t_N})$ of the solution X_t at time set $t := (t_1, t_2, \ldots, t_N)$ with $t_1 < t_2 < \cdots < t_N$, respectively. Substituting X_{t_i} and $X_{t_{i+1}}$ with the observations x_{t_i} and $x_{t_{i+1}}$ in the Eq. (2.3), we write

$$\begin{aligned} H_i(x;\theta) :=& \frac{x_{t_{i+1}} - x_{t_i} - f\left(t_i, x_{t_i};\theta\right)\left(t_{i+1}-t_i\right)}{g\left(t_i, x_{t_i};\theta\right)\left(t_{i+1}-t_i\right)} \\ =& \Big(\frac{x_{t_{i+1}} - x_{t_i}}{t_{i+1}-t_i}\Big)\Big(\frac{1}{g\left(t_i, x_{t_i};\theta\right)}\Big) \\ &- \frac{f\left(t_i, x_{t_i};\theta\right)}{g\left(t_i, x_{t_i};\theta\right)}, \qquad i = 1,2,\dots,N-1, \end{aligned} \tag{2.4}$$

which are real functions of the parameters $\theta^* := (\theta_1^*, \theta_2^*, \dots, \theta_p^*)$. We notice that the kth moment of sample is

$$\frac{1}{N-1}\sum_{i=1}^{N-1}\left(H_i\left(x;\theta^*\right)\right)^k, \quad k = 1,2,\dots.$$

On the other hand, the kth moment of population is expressed as follows

$$\Big(\frac{\sqrt{3}}{\pi}\Big)^k \int_0^1 \Big(\ln\frac{\alpha}{1-\alpha}\Big)^k \,\mathrm{d}\alpha, \quad k = 1,2,\dots,$$

Hence, we set

$$\frac{1}{N-1}\sum_{i=1}^{N-1}\left(H_i(x;\theta^*)\right)^k = \Big(\frac{\sqrt{3}}{\pi}\Big)^k \int_0^1 \Big(\ln\frac{\alpha}{1-\alpha}\Big)^k \,\mathrm{d}\alpha, \quad k = 1,2,\dots,p. \tag{2.5}$$

Equivalently, the values $\theta_1^*, \theta_2^*, \dots, \theta_p^*$ can be estimated by solving the generalized moment estimation problem as follows:

Problem I.

$$\min_{\theta^*}\sum_{k=1}^{p}\left(\frac{1}{N-1}\sum_{i=1}^{N-1}\left(H_i\left(x;\theta^*\right)\right)^k - I_k\right)^2,$$

where

$$I_k = \Big(\frac{\sqrt{3}}{\pi}\Big)^k \int_0^1 \Big(\ln\frac{\alpha}{1-\alpha}\Big)^k \,\mathrm{d}\alpha.$$

In the next section, based on the ANN structure, an alternative expression for $(x_{t_{i+1}} - x_{t_i})/(t_{i+1}-t_i)$ in Eq. (2.4) without the step size term $t_{i+1}-t_i$ is presented.

3. Optimized artificial neural network

ANN structures can approximate nonlinear functions efficiently even for big dataset. According to the ANN characteristics, it has been used for the solution of various mathematical programming problems and tested on several financial applications. As shown in Fig. 1, ANN structures are analogous to biological neural models that are connected by a network of nodes called neurons. In such structures, the network is made up of one input layer, one output layer, and one or more hidden layers. However, some theoretical studies (such as [28]) have confirmed that applying one hidden layer is enough to estimate any intricate non-linear function. Moreover, based on the experimental document presented by Günay [29], it is useless to use more than one layer because using more hidden layers exposes the network to local minimum and further layers make no difference for more efficient and more accurate forecasting. Here, the input layer is formed based on times $t := \{t_0, \dots, t_{N-1}\}$ and the output layer is constructed by time series values $\mathbf{x} := \{\mathbf{x}_0, \dots, \mathbf{x}_{N-1}\}$. The purpose of ANN models is show the correlation between the input and output layers by the biases and weights of the network neurons.

The obtained ANN model is presented as follows:

$$x_t^{\mathrm{ANN}} = w_2^T v(t) + b_2 \tag{3.1}$$

with

$$v(t) = \tanh(a_t), \tag{3.2}$$

and

$$a_t = w_1 t + b_1, \tag{3.3}$$

where $w_1 \in \mathbb{R}^{\ell\times 1}, w_2 \in \mathbb{R}^{\ell\times r}$ are the constant weight matrices and $b_1 \in \mathbb{R}^{\ell\times 1}, b_2 \in \mathbb{R}^{r\times 1}$ are constant bias vectors. Further, r and ℓ represent the neurons number in the output and hidden layers, respectively. Assuming that a_t is given by a linear combination of the input vector t, the output vector of the hidden layer is expressed as $v(t)$ and it is provided by non-linear transformation $\tanh(a_t)$. In this case, based on the linear combination $v(t)$, the network output x_t^{ANN} is established. It should be noted that, the values w_1, w_2, b_1 and b_2 are given by minimizing the mean square error between the network output $x_{t_i}^{\mathrm{ANN}}$ and actual values x_{t_i}:

Problem II.

$$\min_{w_1,w_2,b_1,b_2}\sum_{i=1}^{N-1}\left(x_{t_i}^{\mathrm{ANN}} - x_{t_i}\right)^2$$

subject to: Eqs. (3.1)–(3.3).

Differentiating (3.1) with respect to t gives

$$\frac{dx_t^{\mathrm{ANN}}}{dt} = \sum_{\ell} w_{1_{\ell,1}} w_{2_{\ell,r}}(1 - v_\ell^2(t)). \tag{3.4}$$

To provide an estimate of the UDE parameters, we apply an approximate solution based on the proposed differential term in Eq. (3.4). For this purpose, according to the observed data, we apply the ANN structure to capture the approximate solution. For the rth output, this method can be institutionalized by solving the non-linear problem as follows:

Problem III.

$$\min_{\theta^*}\sum_{k=1}^{p}\left(\frac{1}{Q-1}\sum_{q=1}^{Q-1}\left(\widetilde{H}_q\left(x^{\mathrm{ANN}};\theta^*\right)\right)^k - I_k\right)^2, \tag{3.5}$$

where

$$\begin{aligned} \widetilde{H}_q(x^{\mathrm{ANN}};\theta^*) =& \Big(\sum_{\ell} w_{1_{\ell,1}} w_{2_{\ell,r}}(1 - v_\ell^2(t_q))\Big)\Big(\frac{1}{g\left(t_q, x_{t_q}^{\mathrm{ANN}};\theta^*\right)}\Big) \\ &- \frac{f\left(t_q, x_{t_q}^{\mathrm{ANN}};\theta^*\right)}{g\left(t_q, x_{t_q}^{\mathrm{ANN}};\theta^*\right)}, \qquad q = 0,1,\dots,Q-1. \end{aligned}$$

In expression above, based on the ANN structure, the term $(x_{t_{i+1}} - x_{t_i})/(t_{i+1}-t_i)$ in Eq. (2.4) replaced by Eq. (3.4).

As stated in Problem III, the time spots are described by t_q and are presently not the identical as the basic time spots t_i. In other words, the ANN structure can be used to estimate the parameters of model, when the several data are not available. Due to the fact that an approximation from the actual dataset is derived by ANN structure as trained dataset, this issue is feasible. Therefore, an extended dataset can be used to solve the parameter estimation problem.

A prevalent method for finding the minimums of a multivariate objective function based on deterministic search is the NM algorithm. This method was proposed by John Nelder and Roger Mead in 1965 (see [30]). In fact, the NM method appertains on comparing the values of the function at the vertex of a generic simplex. In this algorithm, a novel simplex is generated by replacing the vertex with the highest value by another point. This simplex adjusts itself to the local optimal value through shrink operators, contraction, expansion and reflection, and then achieves the optimal value.

Assuming that $L(\Theta)$ represents the objective function in Problems I–III, we apply the NM method to minimize the function $L(\Theta)$, where $\Theta = (\theta_1, \theta_2, \dots, \theta_p)$ or $\Theta = (w_1, w_2, b_1, b_2)$. The NM algorithm steps can be summarized as follows:

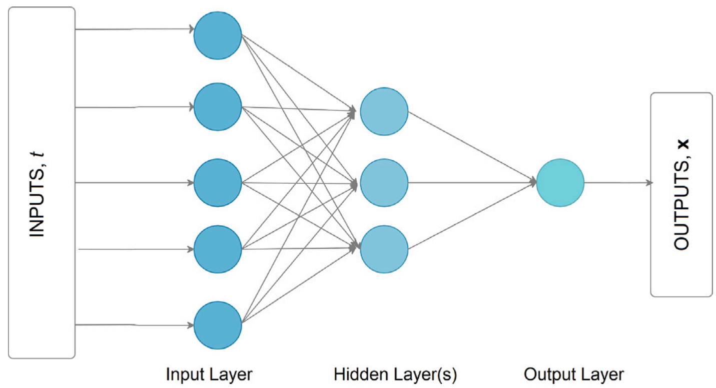

Fig. 1. ANN structure with one input layer, one hidden layer and one output layer.

Step 1. Recognize the coefficients of the NM algorithm as $\alpha > 0$ (reflection coefficient), $\gamma > 0$ (expansion coefficient), $0 < \rho \leq 0.5$ (contraction coefficient) and β (shrink coefficient). These are commonly taken as $\alpha = 1, \gamma = 2, \rho = 0.5$ and $\beta = 0.5$ in the literature.

Step 2. Recognize initial values $\Theta_1, \dots, \Theta_4$.

Step 3. The vertices are arranged in ascending order, that is, $L(\Theta_1) \leq L(\Theta_2) \leq L(\Theta_3) \leq L(\Theta_4)$.

Step 4. Calculate Θ_0 which is the centroid of Θ_1, Θ_2 and Θ_3.

Step 5. Compute the reflected point $\Theta_r = \Theta_0 + \alpha(\Theta_0 - \Theta_4)$. If $L(\Theta_1) \leq L(\Theta_r) < L(\Theta_3)$, generate a novel simplex by switching Θ_4 with Θ_r, and then go to Step 3.

Step 6. If $L(\Theta_r) < L(\Theta_1)$, compute the expanded point $\Theta_e = \Theta_0 + \gamma(\Theta_r - \Theta_0)$. If $L(\Theta_e) < L(\Theta_r)$, generate a new simplex by switching Θ_4 with Θ_e, and go to Step 3. Otherwise, produce a novel simplex by switching Θ_4 with Θ_r, and go to Step 3.

Step 7. If $L(\Theta_r) \geq L(\Theta_1)$, compute the contracted point $\Theta_c = \Theta_0 + \rho(\Theta_4 - \Theta_0)$). If $L(\Theta_c) < L(\Theta_4)$, generate a novel simplex by switching Θ_4 with Θ_c, and go to Step 3. Otherwise, go on to Step 8.

Step 8. For any $i \in \{2, 3, 4\}$, switch the point with $\Theta_i = \Theta_1 + \beta(\Theta_i - \Theta_1)$, up to the convergence criteria is established. Then go to Step 3.

Step 9. When the algorithm ceases, the point with the under-most value is an optimal minimum.

We notice that Problem I minimizes the mean square error between the population moments and uncertain sample moments, whilst the Problem III minimizes the mean square error between the population moments and uncertain sample moments obtained by ANN structure based on the observed data with predicted value. As demonstrated, satisfactory estimates can be achieved from this approach.

The parameter estimation problem is described based on two other problems. The first problem is expressed the ANN structure to achieve the approximation of the observed dataset and to predict them. The model presented by Eqs. (3.1)–(3.3) is proceeded from the observed dataset and the second problem is formulated by the optimized ANN and the derivatives of the ANN structure in Eq. (3.5). Solving the Problem III carry out an estimation of the UDE parameters, the scheme of the presented approach is illustrated in Fig. 2, in which step B can be applied to better improve the parameters estimation. Theoretically, the ANN network can be trained based on experience and given the appropriate number of hidden layers and sufficient computational sources, and can approximate any complex function. Using this feature of the ANN structure, we estimated the parameters of the UDE based on the method of moments. The main benefits of the proposed method are:

(i) The approximate dataset obtained by the ANN structure through minimizing the mean square error between the observed dataset and the ANN structure can be wholly efficient.

(ii) This method and the corresponding enhanced dataset, rather than all the dataset can be used for parameter estimation. For instance as shown in Fig. 2, the original dataset has 15 data points (given by the circle) in step A and 20 data points (5 data given by the square as forecast data) in step C. This process is achieved by the optimized ANN in stage B. In fact, by performing step D, in addition to simplifying the problem (without the time step part) and estimating the parameters of the uncertainty model, future data can be modeled based on the UDE driven by Liu process.

(iii) Differentiability features of the proposed ANN structure make it possible for using the formularization of the first sub-problem.

(iv) ANN structures are wholly efficient and can approximate the non-linear functions.

On the other hand, global optimization algorithms based on population methods are effective and can avoid local minima and finding the best regions of the design space. Although, they can favorably determine a basin for the local optimum or global optimum, but they may not be able to find the best solution in the basin. The computation time cost in global optimization methods is inevitable even for a small number of parameters. However, it may be worth in cases where the value of verifying the performance is high, or the cost of being wrong about the reliability or safety is high (see Boyd et al. [31]).

In the next section, in addition to dealing with the parameter estimation problem, numerically, we compare the particle swarm optimization (PSO) method and the NM algorithm for the parameter estimation of the UDE in Problem III. The PSO algorithm is a global search method, which is modeled based on the social behavior of bird flock (see Kennedy and Eberhart [32]). It is necessary to mention that the implementation of this method in the MATLAB environment can be easily performed by using the command *particleswarm*.

4. Empirical examples

In this section, some examples based on the real financial dataset are proposed to illustrate the efficiency of the presented method.

Consider the following UDE

$$dX_t = f(t, X_t; \theta^*)dt + g(t, X_t; \theta^*)dC_t,$$

where f and g are two measurable real functions, C_t is the uncertain process and $\theta^* := (\theta_1^*, \dots, \theta_p^*)$ is the estimated parameters. As Yang et al. [12] stated, the expected value of the solution of the UDE at time t_{n+1} can be considered as the forecast value at this time. In other words, according to Definition 2.3 and assuming that $\hat{X}(t_{n+1})$ is the forecast value and $\Phi_{t_{n+1}}^{-1}(\alpha) = X^{\alpha}(t_{n+1})$ is the distribution function, we have

$$\hat{X}(t_{n+1}) = \int_0^1 \Phi_{t_{n+1}}^{-1}(\alpha) d\alpha. \tag{4.1}$$

To show the efficiency of the uncertainty model with the estimated parameters according to the artificial neural network, we predict the

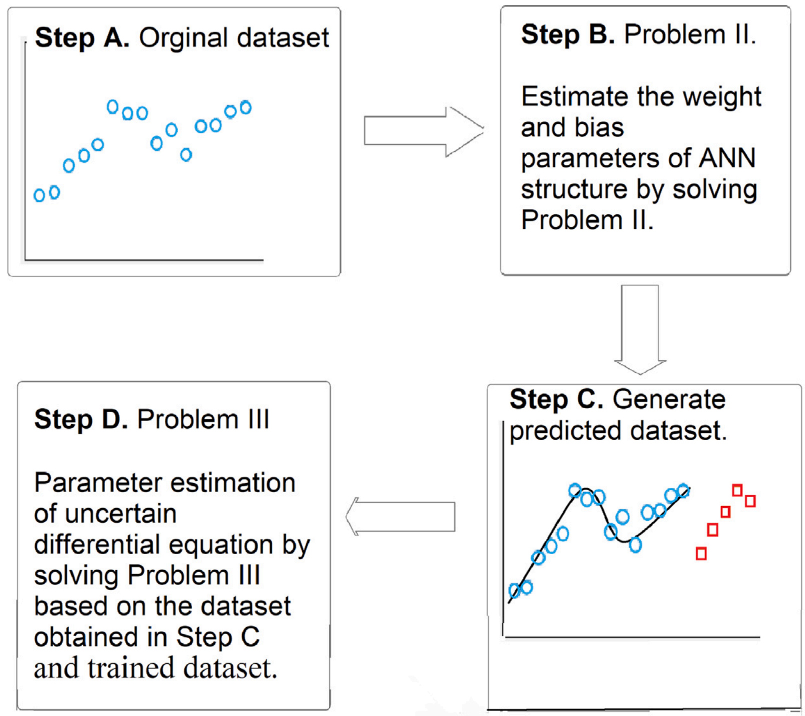

Fig. 2. Presented method for estimation parameter of UDE.

stock data based on Eq. (4.1). We also consider the one-step ahead approach under the stochastic model as a comparative method.

Consider the following SDE

$$dY_t = f(t, Y_t; \theta^*)dt + g(t, Y_t; \theta^*)dW_t,$$

where f and g are two measurable real function, W_t is a Wiener process and $\theta^* := (\theta_1^*, \ldots, \theta_p^*)$ is the estimated parameters. Let us $\mathcal{F}_{t_n}$, $n = 0, 1, \ldots$, be the filtration generated by the market information. In this case, the one-step ahead forecast at time t_n can be calculated based on the conditional expectation $\mathbb{E}[Y_{t_{n+1}} \mid \mathcal{F}_{t_n}]$.

It is often appropriate to implement the optimization algorithm several times and record the optimal point found at each run to ensure that relatively good solutions are found. Based on this fact, we consider one of these two stopping criteria for NM and PSO algorithms at each implementation: (1) The change in the best utility function value (i.e. the change to the minimal value found in previous iterations) is less than the function tolerance $TolFun$ (=10e-10); (2) The maximum number of iteration of particle swarm and the maximum number of iterations in the NM algorithm are two hundred times the total number of the unknown parameters at each implementation. Furthermore, two algorithms are stopped, when the number of implementations is more than 100 or the value of the objective function is less than 10e−05. It should be noted that at each implementation, we generate p standard normal random variables as a new initial point in the NM algorithm, where p is the total number of the unknown parameters.

Example 1. Consider the geometric Liu process X_t

$$dX_t = \mu X_t dt + \sigma X_t dC_t, \tag{4.2}$$

where parameters μ and $\sigma > 0$ must be estimated. According to Table 1, we have 20 observed data as training dataset in Intel Corporation from Mar 01, 2022–Mar 28, 2022. Notice that these 20 dataset are applied to estimate the weight and bias parameter of the ANN model by solving Problem II. We consider the next 3 data (Mar 29, 2022–Mar 31, 2022) as test dataset. In this case, we will have 23 data totally for parameter estimation of the UDE (4.2). Fig. 3 compares real price with trained and tested dataset. In this figure, the mean square error (MSE), mean absolute error (MAE) and mean relative error (MRE) criteria for the trained data and test data are reported. As it turns out, the optimized ANN has acceptable accuracy. According to the Problem III, the estimated values μ^* and σ^* are obtained by solving the non-linear problem based on the NM algorithm:

$$\min_{\mu^*,\sigma^*} \sum_{k=1}^{2} \left(\frac{1}{22} \sum_{q=0}^{22} \left(\widetilde{H}_q(x^{\mathrm{ANN}}; \mu_1^*, \mu_2^*, \sigma^*) \right)^k - \left(\frac{\sqrt{3}}{\pi} \right)^k \int_0^1 \left(\ln \frac{\alpha}{1-\alpha} \right)^k \mathrm{d}\alpha \right)^2, \tag{4.3}$$

where

$$\widetilde{H}_q(x^{\mathrm{ANN}}; \mu^*, \sigma^*) = \left(\frac{1}{\sigma^* x_{t_q}^{\mathrm{ANN}}} \right) \left(\sum_{\ell=1}^{3} w_{1_{\ell,1}} w_{2_{\ell,1}} (1 - v_\ell^2(t_q)) - \mu^* \right).$$

Estimation of parameters by solving the problem above is obtained as follows

$$\mu^* = 0.0893, \quad \sigma^* = 0.3129.$$

Moreover, the estimated results of the weight and bias parameters of the proposed ANN by solving the Problem II are obtained as follows

$$w_1 = \begin{pmatrix} -6.0722 \\ -9.3619 \\ -10.8718 \end{pmatrix}, \quad w_2 = \begin{pmatrix} -19.6447 \\ 6.1356 \\ 11.0025 \end{pmatrix}, \quad b_1 = \begin{pmatrix} 3.5747 \\ 3.7500 \\ 6.6830 \end{pmatrix}, \quad b_2 = 50.3052.$$

Then, we have

$$dX_t = 0.0893 X_t dt + 0.3129 X_t dC_t, \tag{4.4}$$

Due to Fig. 4, all the actual dataset are situated in the region between the 0.85-path and the 0.15-path of the UDE (4.4). Thus, the estimated parameters are admissible.

Putting the increment dW_t of Wiener process instead of the increment dC_t of Liu process in expression (4.2), we have the following SDE

$$dX_t = \hat{\mu} X_t dt + \hat{\sigma} X_t dW_t, \tag{4.5}$$

Now, by applying the maximum likelihood estimation method and 23 stock price data (the dataset reported in Table 1 and 3 data obtained

Table 1
Observed data in Intel Corporation from Mar 01, 2022–Mar 28, 2022.

Trading day	Day (t_i)	Stock price (x_{t_i})	Trading day	Day (t_i)	Stock price (x_{t_i})
3/1/2022	0	46.82	2/15/2022	0.4762	44.81
3/2/2022	0.0476	48.87	3/16/2022	0.5238	46.63
3/3/2022	0.0952	47.93	3/17/2022	0.5714	47.14
3/4/2022	0.1429	48.07	3/18/2022	0.6190	47.45
3/7/2022	0.1905	47.68	3/21/2022	0.6667	47.39
3/8/2022	0.2381	47.50	3/22/2022	0.7143	48.39
3/9/2022	0.2857	47.63	3/23/2022	0.7619	48.27
3/10/2022	0.3333	46.66	3/24/2022	0.8095	51.62
3/11/2022	0.3810	45.83	3/25/2022	0.8571	51.63
3/14/2022	0.4286	44.40	3/28/2022	0.9048	51.51

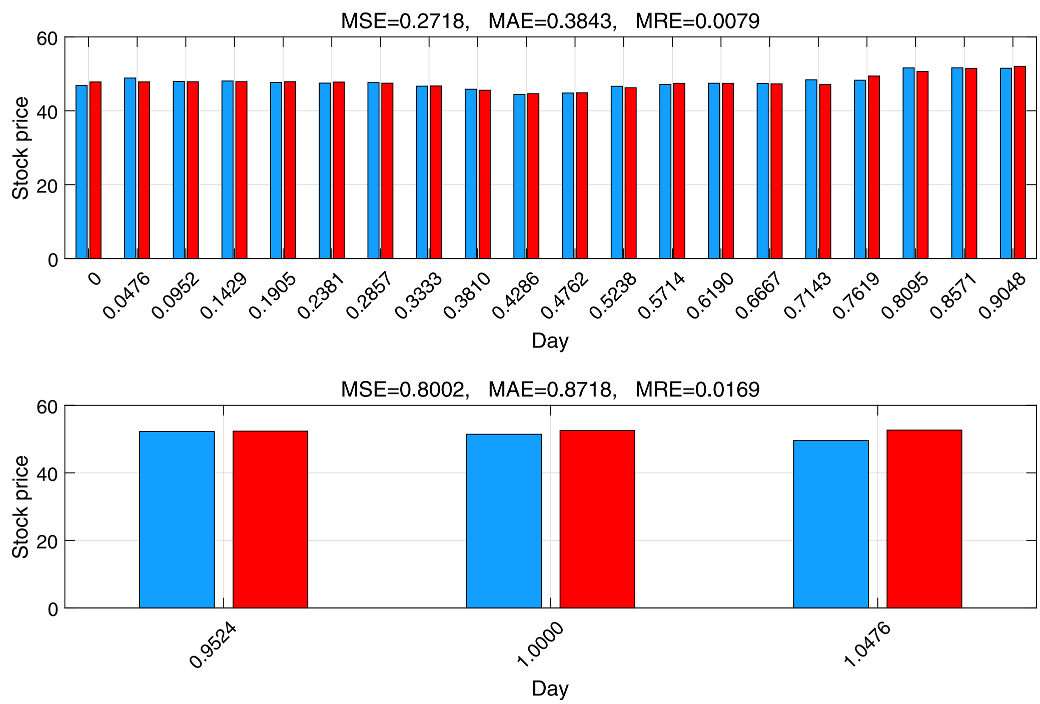

Fig. 3. Comparing observed data (blue color) with trained data (red color) in figure above and test data (red color) in figure below. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

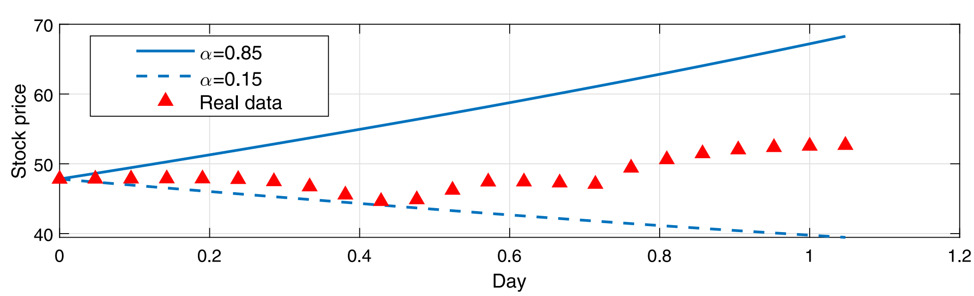

Fig. 4. Actual dataset and α-paths of X_t in Example 1.

by the ANN structure), the estimated parameters $\hat{\mu}^*$ and $\hat{\sigma}^*$ in Eq. (4.5) are obtained as follows

$$\hat{\mu}^* = 0.9856, \quad \hat{\sigma}^* = 0.0192.$$

In this case, the forecast one-step ahead at time t_n, can be expressed as follows

$$\mathbb{E}[X_{t_{n+1}} \mid \mathcal{F}_{t_n}] = X_{t_n} + 0.7813 X_{t_n}(t_{n+1} - t_n).$$

Using expression (4.1) and the one-step ahead method leads to the results obtained in Table 2. A close look at this table shows that the values predicted by the uncertain model, whose its parameters are estimated based on the ANN, are closer to the future price of Intel Corporation.

In Fig. 5, the logarithmic values of the objective function in the problem (4.3) based on the NM and PSO optimization algorithms with various number of iterations are graphed. As it turns out, the NM algorithm has less computational cost and in practice it can be applied to solve the parameters estimation problem of the geometric Liu process. The reason of this matter is that the NM algorithm reaches the optimal point in fewer implementations. It should be noted that from the two stop conditions of algorithms, that is, the number of implementations more than 100 or the value of the objective function

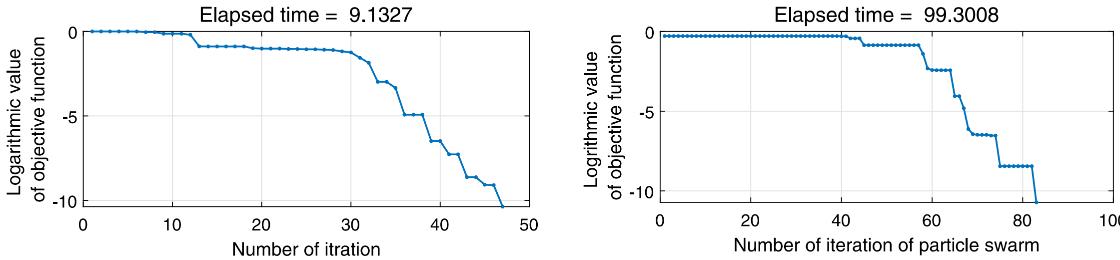

Fig. 5. The logarithmic values of the objective function in problem (4.3), which have been obtained by NM (left figure) and PSO (right figure) algorithms with different number of iterations.

Table 2
Comparing the future price of Intel Corporation with the predicted value obtained by UDE and SDE. [·] denotes the absolute error between the actual value and forecast value.

Trading day	4/1/2022	4/4/2022	4/5/2022
Actual value	48.11	49.20	48.13
Forecast value by the UDE	49.76 [1.65]	48.31 [0.88]	49.40 [1.27]
Forecast value by the SDE	51.88 [3.77]	50.36 [1.16]	51.50 [3.37]

less than 10e−05, according to the second condition, both NM and PSO algorithms have given the results in Fig. 5.

Example 2. Consider the uncertain Ornstein–Uhlenbeck process X_t

$$dX_t = (\mu_1 - \mu_2 X_t)dt + \sigma dC_t, \tag{4.6}$$

with three parameters μ_1, μ_2 and $\sigma > 0$ to be estimated. According to Table 3, we have 20 observed data as training dataset in Xiaomi Corporation from Mar 01, 2022–Mar 28, 2022. Notice that these 20 dataset are applied to estimate the weight and bias parameter of the ANN model by solving Problem II. We consider the next 3 data (Mar 29, 2022–Mar 31, 2022) as test dataset. In this case, we will have 23 data totally for parameter estimation of the UDE (4.6). Fig. 6 compares real price with trained and tested dataset. In this figure, the MSE, MAE and MRE criteria for the trained data and test data are reported. As it turns out, the optimized ANN has acceptable accuracy.

According to the Problem III, the estimated values μ_1^*, μ_2^* and σ^* are obtained by solving the non-linear problem based on the NM algorithm:

$$\min_{\mu_1^*,\mu_2^*,\sigma^*} \sum_{k=1}^{3} \left(\frac{1}{22} \sum_{q=0}^{22} \left(\widetilde{H}_q(x^{\text{ANN}}; \mu_1^*, \mu_2^*, \sigma^*) \right)^k - \left(\frac{\sqrt{3}}{\pi} \right)^k \int_0^1 \left(\ln \frac{\alpha}{1-\alpha} \right)^k \mathrm{d}\alpha \right)^2, \tag{4.7}$$

where

$$\widetilde{H}_q(x^{\text{ANN}}; \mu_1^*, \mu_2^*, \sigma^*) = \left(\frac{1}{\sigma^*} \right) \left(\sum_{\ell=1}^{3} w_{1_{\ell,1}} w_{2_{\ell,1}} (1 - v_\ell^2(t_q)) - \mu_1^* + \mu_2^* x_{t_q}^{\text{ANN}} \right).$$

Estimation of parameters by solving the problem above is obtained as follows

$$\mu_1^* = 0.9122, \quad \mu_2^* = 0.5447, \quad \sigma^* = 15.7886.$$

Moreover, the estimated results of the weight and bias parameters of the proposed ANN by solving the Problem II are obtained as follows

$$w_1 = \begin{pmatrix} 2.6799 \\ 20.8555 \\ -17.6313 \end{pmatrix}, \quad w_2 = \begin{pmatrix} -68.7096 \\ -16.8643 \\ -17.5411 \end{pmatrix}, \quad b_1 = \begin{pmatrix} 2.0416 \\ -9.1477 \\ 7.8337 \end{pmatrix}, \quad b_2 = 82.1899.$$

Then, we have

$$dX_t = (0.9122 - 0.5447X_t)dt + 15.7886dC_t, \tag{4.8}$$

Due to Fig. 7, all the actual dataset are situated in the region between the 0.75-path and the 0.25-path of the UDE (4.8). Thus, the estimated parameters are admissible.

Putting the increment dW_t of Wiener process instead of the increment dC_t of Liu process in expression (4.6), we have the following SDE

$$dX_t = (\hat{\mu}_1 - \hat{\mu}_2 X_t)dt + \hat{\sigma} dW_t, \tag{4.9}$$

Now, by applying the maximum likelihood estimation method and 23 stock price data (the dataset reported in Table 3 and 3 data obtained by the ANN structure), the estimated parameters $\hat{\mu}_1^*, \hat{\mu}_2^*$ and $\hat{\sigma}^*$ in Eq. (4.9) are obtained as follows

$$\hat{\mu}_1^* = 0.5668, \quad \hat{\mu}_2^* = 0.8541, \quad \hat{\sigma}^* = 0.4397.$$

In this case, the forecast one-step ahead at time t_n, can be expressed as follows

$$\mathbb{E}[X_{t_{n+1}} \mid \mathcal{F}_{t_n}] = X_{t_n} + (0.5668 - 0.8541X_{t_n})(t_{n+1} - t_n).$$

Using expression (4.1) and the one-step ahead method leads to the results obtained in Table 4. A close look at this table shows that the values predicted by the uncertain model, whose its parameters are estimated based on the ANN, are closer to the future price of Xiaomi Corporation.

In Fig. 8, the logarithmic values of the objective function in the problem (4.7) based on the NM and PSO optimization algorithms with various number of iterations are graphed. As it turns out, the NM algorithm has less computational cost and in practice it can be applied to solve the parameters estimation problem of the uncertain Ornstein–Uhlenbeck process. It should be noted that from the two stop conditions of algorithms, that is, the number of implementations more than 100 or the value of the objective function less than 10e−05, according to the second condition, both NM and PSO algorithms have given the results in Fig. 8.

Example 3. Consider the uncertain mean-reverting process X_t

$$dX_t = (\mu_1 - \mu_2 X_t)dt + \sigma X_t dC_t, \tag{4.10}$$

with three parameters μ_1, μ_2 and $\sigma > 0$ to be estimated. According to Table 5, we have 20 observed data as training dataset in ADM company from Mar 01, 2022–Mar 28, 2022. Notice that these 20 dataset are applied to estimate the weight and bias parameter of the ANN model by solving Problem II. We consider the next 3 data (Mar 29, 2022–Mar 31, 2022) as test dataset. In this case, we will have 23 data totally for parameter estimation of the UDE (4.10). Fig. 9 compares real price with trained and tested dataset. In this figure, the MSE, MAE and MRE criteria for the trained data and test data are reported. As it turns out, the optimized ANN has acceptable accuracy.

According to the Problem III, the estimated values μ_1^*, μ_2^* and σ^* are obtained by solving the non-linear problem based on the NM algorithm:

$$\min_{\mu_1^*,\mu_2^*,\sigma^*} \sum_{k=1}^{3} \left(\frac{1}{22} \sum_{q=0}^{22} \left(\widetilde{H}_q(x^{\text{ANN}}; \mu_1^*, \mu_2^*, \sigma^*) \right)^k - \left(\frac{\sqrt{3}}{\pi} \right)^k \int_0^1 \left(\ln \frac{\alpha}{1-\alpha} \right)^k \mathrm{d}\alpha \right)^2, \tag{4.11}$$

Table 3
Observed data in Xiaomi Corporation from Mar 01, 2022–Mar 28, 2022.

Trading day	Day (t_i)	Stock price (x_{t_i})	Trading day	Day (t_i)	Stock price (x_{t_i})
3/1/2022	0	14.68	2/15/2022	0.4762	11.50
3/2/2022	0.0476	14.82	3/16/2022	0.5238	13.4
3/3/2022	0.0952	14.72	3/17/2022	0.5714	13.94
3/4/2022	0.1429	13.96	3/18/2022	0.6190	13.52
3/7/2022	0.1905	13.68	3/21/2022	0.6667	13.38
3/8/2022	0.2381	12.94	3/22/2022	0.7143	14.20
3/9/2022	0.2857	13.30	3/23/2022	0.7619	14.78
3/10/2022	0.3333	13.58	3/24/2022	0.8095	14.50
3/11/2022	0.3810	13.66	3/25/2022	0.8571	14.42
3/14/2022	0.4286	12.48	3/28/2022	0.9048	14.48

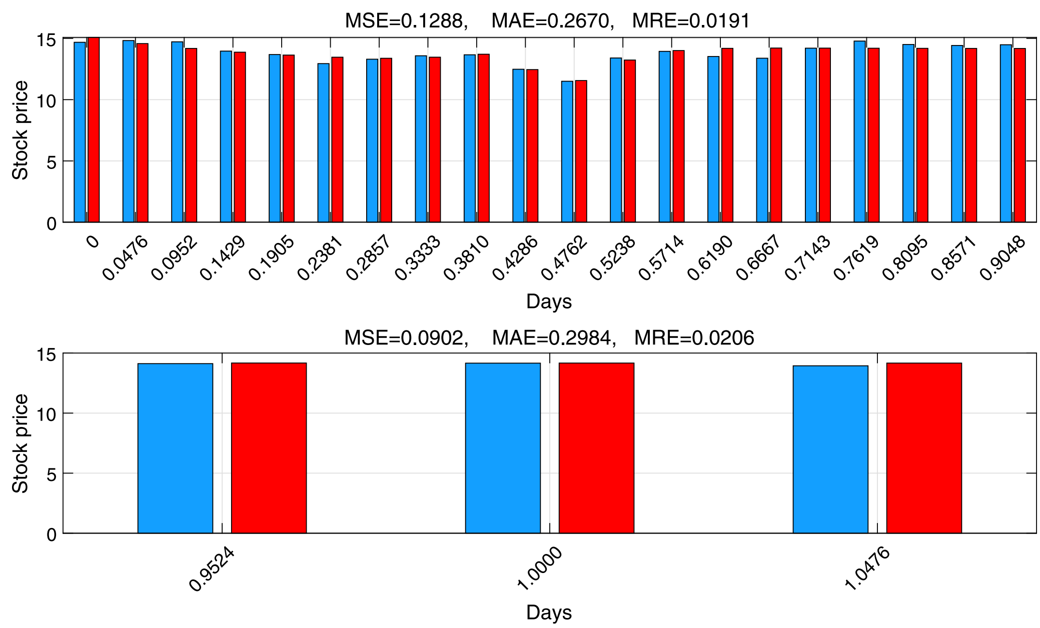

Fig. 6. Comparing observed data (blue color) with trained data (red color) in figure above and test data (red color) in figure below. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

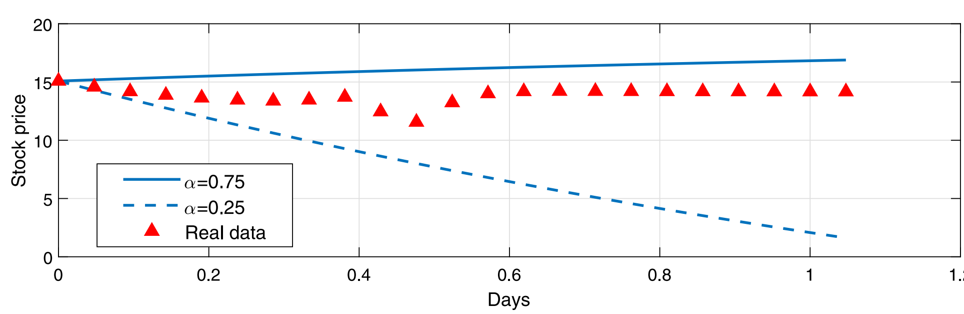

Fig. 7. Actual dataset and α-paths of X_t in Example 2.

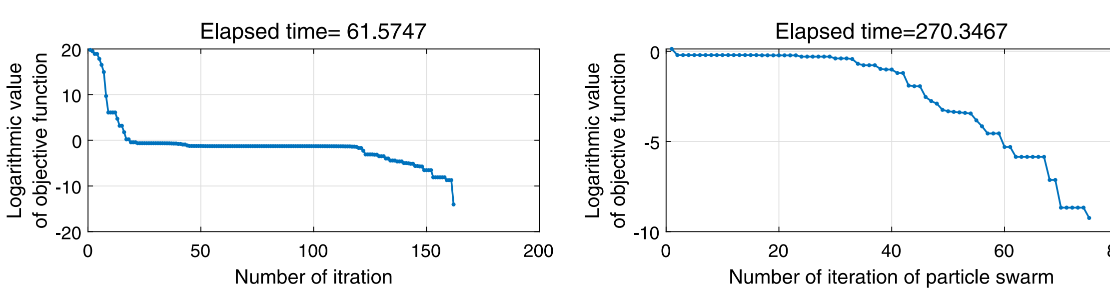

Fig. 8. The logarithmic values of the objective function in problem (4.7), which have been obtained by NM (left figure) and PSO (right figure) algorithms with different number of iterations.

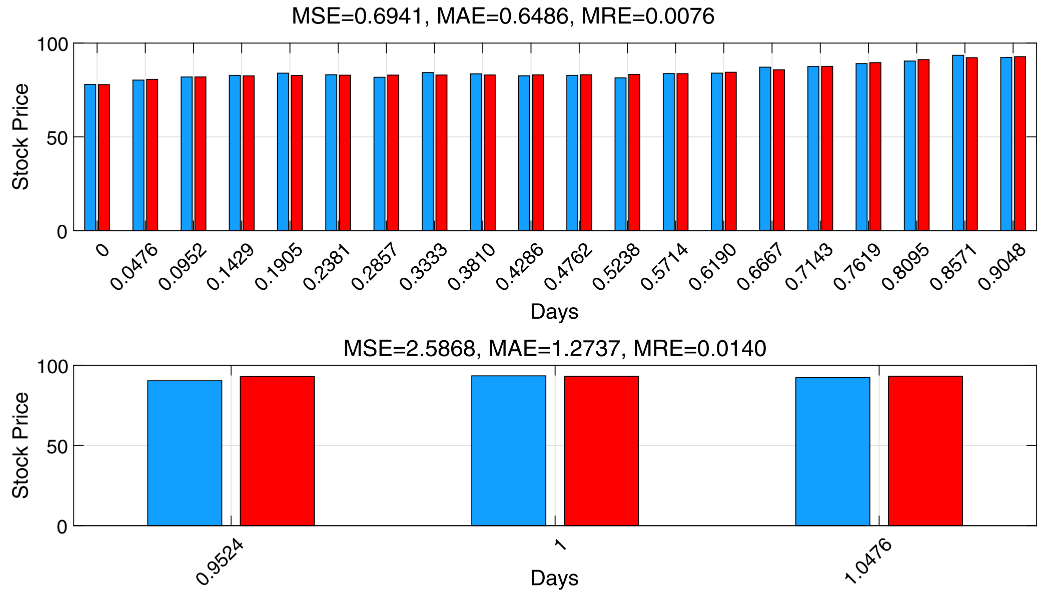

Fig. 9. Comparing observed data (blue color) with trained data (red color) in figure above and test data (red color) in figure below. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 4
Comparing the future price of Xiaomi Corporation with the predicted value obtained by UDE and SDE. [·] denotes the absolute error between the actual value and forecast value.

Trading day	4/1/2022	4/4/2022	4/6/2022
Actual value	13.78	14.04	13.66
Forecast value by the UDE	13.63 [0.14]	13.47 [0.56]	13.72 [0.06]
Forecast value by the SDE	13.40 [0.37]	13.24 [0.79]	13.49 [0.16]

where

$$\widetilde{H}_q(x^{\text{ANN}}; \mu_1^*, \mu_2^*, \sigma^*) = \left(\frac{1}{\sigma^* x_{t_q}^{\text{ANN}}}\right)\left(\sum_{\ell=1}^{3} w_{1_{\ell,1}} w_{2_{\ell,1}} (1 - v_\ell^2(t_q)) - \mu_1^* + \mu_2^* x_{t_q}^{\text{ANN}}\right).$$

Estimation of parameters by solving the problem above is obtained as follows

$$\mu_1^* = 29.4338, \quad \mu_2^* = 0.0231, \quad \sigma^* = 0.2948.$$

Moreover, the estimated results of the weight and bias parameters of the proposed ANN by solving the Problem II are obtained as follows

$$w_1 = \begin{pmatrix} 8.7888 \\ 8.2935 \\ 8.6604 \end{pmatrix}, \quad w_2 = \begin{pmatrix} 15.4253 \\ 5.1576 \\ 47.7545 \end{pmatrix}, \quad b_1 = \begin{pmatrix} -23.4348 \\ -6.0266 \\ 1.4383 \end{pmatrix}, \quad b_2 = 55.8032.$$

Then, we have

$$dX_t = (29.4338 - 0.0231 X_t)dt + 0.2948 X_t dC_t, \tag{4.12}$$

Due to Fig. 10, all the actual dataset are situated in the region between the 0.85-path and the 0.15-path of the UDE (4.12). Thus, the estimated parameters are admissible.

Putting the increment dW_t of Wiener process instead of the increment dC_t of Liu process in expression (4.10), we have the following SDE

$$dX_t = (\hat{\mu}_1 - \hat{\mu}_2 X_t)dt + \hat{\sigma} X_t dW_t, \tag{4.13}$$

Now, by applying the maximum likelihood estimation method and 23 stock price data (the dataset reported in Table 5 and 3 data obtained by the ANN structure), the estimated parameters $\hat{\mu}_1^*, \hat{\mu}_2^*$ and $\hat{\sigma}^*$ in Eq. (4.13) are obtained as follows

$$\hat{\mu}_1^* = 0.5238, \quad \hat{\mu}_2^* = 0.7842, \quad \hat{\sigma}^* = 0.6231.$$

In this case, the forecast one-step ahead at time t_n, can be expressed as follows

$$\mathbb{E}[X_{t_{n+1}} \mid \mathcal{F}_{t_n}] = X_{t_n} + (0.5238 - 0.7842 X_{t_n})(t_{n+1} - t_n).$$

Using expression (4.1) and the one-step ahead method leads to the results obtained in Table 6. A close look at this table shows that the values predicted by the uncertain model, whose its parameters are estimated based on the ANN, are closer to the future price of ADM company.

In Fig. 11, the logarithmic values of the objective function in the problem (4.11) based on the NM and PSO optimization algorithms with various number of iterations are graphed. As it turns out, the NM algorithm has less computational cost and in practice it can be applied to solve the parameters estimation problem of the uncertain mean-reverting process. It should be noted that from the two stop conditions of algorithms, that is, the number of implementations more than 100 or the value of the objective function less than 10e−05, the NM algorithm according to the second condition and the PSO algorithm according to the first condition have given the results in Fig. 11.

Example 4. Consider the uncertain exponential Ornstein–Uhlenbeck process

$$dX_t = \mu_1(1 - \mu_2 \ln X_t)X_t dt + \sigma X_t dC_t, \tag{4.14}$$

with three parameters μ_1, μ_2 and $\sigma > 0$ to be estimated. According to Table 7, we have 20 observed data as training dataset in Cisco Systems Corporation from Mar 01, 2022–Mar 28, 2022. Notice that these 20 dataset are applied to estimate the weight and bias parameter of the ANN model by solving Problem II. We consider the next 3 data (Mar 29, 2022–Mar 31, 2022) as test dataset. In this case, we will have 23 data totally for parameter estimation of the UDE (4.14). Fig. 12 compares real price with trained and tested dataset. In this figure, the MSE, MAE and MRE criteria for the trained data and test data are reported. As it turns out, the optimized ANN has acceptable accuracy.

According to the Problem III, the estimated values μ_1^*, μ_2^* and σ^* are obtained by solving the non-linear problem based on the NM algorithm:

$$\min_{\mu_1^*, \mu_2^*, \sigma^*} \sum_{k=1}^{3} \left(\frac{1}{22} \sum_{q=0}^{22} \left(\widetilde{H}_q(x^{\text{ANN}}; \mu_1^*, \mu_2^*, \sigma^*) \right)^k - \left(\frac{\sqrt{3}}{\pi}\right)^k \int_0^1 \left(\ln \frac{\alpha}{1-\alpha} \right)^k \mathrm{d}\alpha \right)^2, \tag{4.15}$$

Table 5
Observed data in ADM company from Mar 01, 2022–Mar 28, 2022.

Trading day	Day (t_i)	Stock price (x_{t_i})	Trading day	Day (t_i)	Stock price (x_{t_i})
3/1/2022	0	77.99	2/15/2022	0.4762	82.81
3/2/2022	0.0476	80.28	3/16/2022	0.5238	81.45
3/3/2022	0.0952	81.93	3/17/2022	0.5714	83.78
3/4/2022	0.1429	82.80	3/18/2022	0.6190	83.98
3/7/2022	0.1905	83.97	3/21/2022	0.6667	87.16
3/8/2022	0.2381	83.13	3/22/2022	0.7143	87.56
3/9/2022	0.2857	81.78	3/23/2022	0.7619	89.09
3/10/2022	0.3333	84.31	3/24/2022	0.8095	90.43
3/11/2022	0.3810	83.60	3/25/2022	0.8571	93.49
3/14/2022	0.4286	82.57	3/28/2022	0.9048	92.35

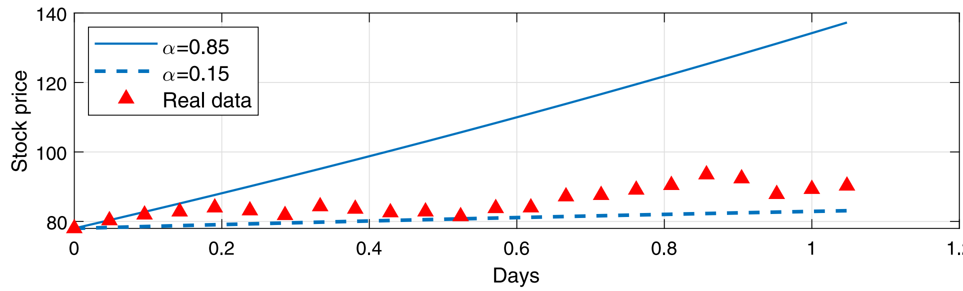

Fig. 10. Actual dataset and α-paths of X_t in Example 3.

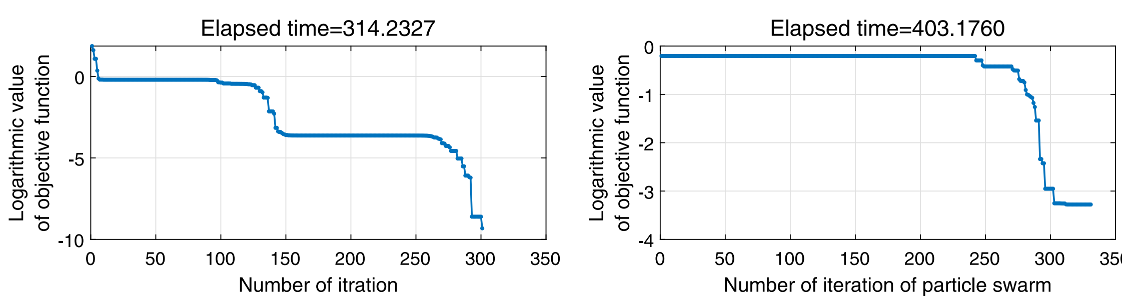

Fig. 11. The logarithmic values of the objective function in problem (4.11), which have been obtained by NM (left figure) and PSO (right figure) algorithms with different number of iterations.

Table 6
Comparing the future price of ADM company with the predicted value obtained by UDE and SDE. [·] denotes the absolute error between the actual value and forecast value.

Trading day	4/1/2022	4/4/2022	4/5/2022
Actual value	91.17	90.43	90.90
Forecast value by the UDE	91.57 [0.40]	92.48 [2.05]	91.74 [0.84]
Forecast value by the SDE	86.91 [4.25]	87.79 [2.63]	87.07 [3.82]

where

$$\widetilde{H}_q(x^{\mathrm{ANN}};\mu_1^*,\mu_2^*,\sigma^*)=\left(\frac{1}{\sigma^* x_{t_q}^{\mathrm{ANN}}}\right)\left(\sum_{\ell=1}^{3} w_{1_{\ell,1}}w_{2_{\ell,1}}(1-v_\ell^2(t_q)) - \mu_1^*(1-\mu_2^*\ln x_{t_q}^{\mathrm{ANN}})x_{t_q}^{\mathrm{ANN}}\right).$$

Estimation of parameters by solving the problem above is obtained as follows

$$\mu_1^*=0.0669,\quad \mu_2^*=0.0764,\quad \sigma^*=1.2352.$$

Moreover, the estimated results of the weight and bias parameters of the proposed ANN by solving the Problem II are obtained as follows

$$w_1=\begin{pmatrix}153.7080\\0.6064\\16.7552\end{pmatrix},\quad w_2=\begin{pmatrix}1.1621\\-22.3490\\84.5401\end{pmatrix},\quad b_1=\begin{pmatrix}-72.9179\\0.8248\\2.1291\end{pmatrix},\quad b_2=-11.2559.$$

Then, we have

$$dX_t = 1.3713(1-0.1293\ln X_t)X_t dt + 1.0117X_t dC_t \tag{4.16}$$

Due to Fig. 13, all the actual dataset are situated in the region between the 0.65-path and the 0.35-path of the UDE (4.16). Thus, the estimated parameters are admissible.

Putting the increment dW_t of Wiener process instead of the increment dC_t of Liu process in expression (4.14), we have the following SDE

$$dX_t = \hat{\mu}_1(1-\hat{\mu}_2\ln X_t)X_t dt + \hat{\sigma}X_t dW_t, \tag{4.17}$$

Now, by applying the maximum likelihood estimation method and 23 stock price data (the dataset reported in Table 7 and 3 data obtained by the ANN structure), the estimated parameters $\hat{\mu}_1^*, \hat{\mu}_2^*$ and $\hat{\sigma}^*$ in Eq. (4.17) are obtained as follows

$$\hat{\mu}_1^*=0.8419,\quad \hat{\mu}_2^*=0.5189,\quad \hat{\sigma}^*=0.6032.$$

In this case, the forecast one-step ahead at time t_n, can be expressed as follows

$$\mathbb{E}[X_{t_{n+1}}\mid \mathcal{F}_{t_n}] = X_{t_n} + 0.8419(1-0.5189\ln X_{t_n})(t_{n+1}-t_n)X_{t_n}.$$

Using expression (4.1) and the one-step ahead method leads to the results obtained in Table 8. A close look at this table shows that the values predicted by the uncertain model, whose its parameters are estimated based on the ANN, are closer to the future price of Cisco Systems Corporation.

Table 7
Observed data in Cisco Systems Corporation from Mar 01, 2022–Mar 28, 2022.

Trading day	Day (t_i)	Stock price (x_{t_i})	Trading day	Day (t_i)	Stock price (x_{t_i})
3/1/2022	0	54.62	2/15/2022	0.4762	55.60
3/2/2022	0.0476	56.15	3/16/2022	0.5238	55.81
3/3/2022	0.0952	56.07	3/17/2022	0.5714	56.17
3/4/2022	0.1429	56.51	3/18/2022	0.6190	55.97
3/7/2022	0.1905	55.59	3/21/2022	0.6667	56.00
3/8/2022	0.2381	54.40	3/22/2022	0.7143	56.17
3/9/2022	0.2857	55.92	3/23/2022	0.7619	54.48
3/10/2022	0.3333	54.71	3/24/2022	0.8095	54.50
3/11/2022	0.3810	54.69	3/25/2022	0.8571	54.97
3/14/2022	0.4286	54.30	3/28/2022	0.9048	55.14

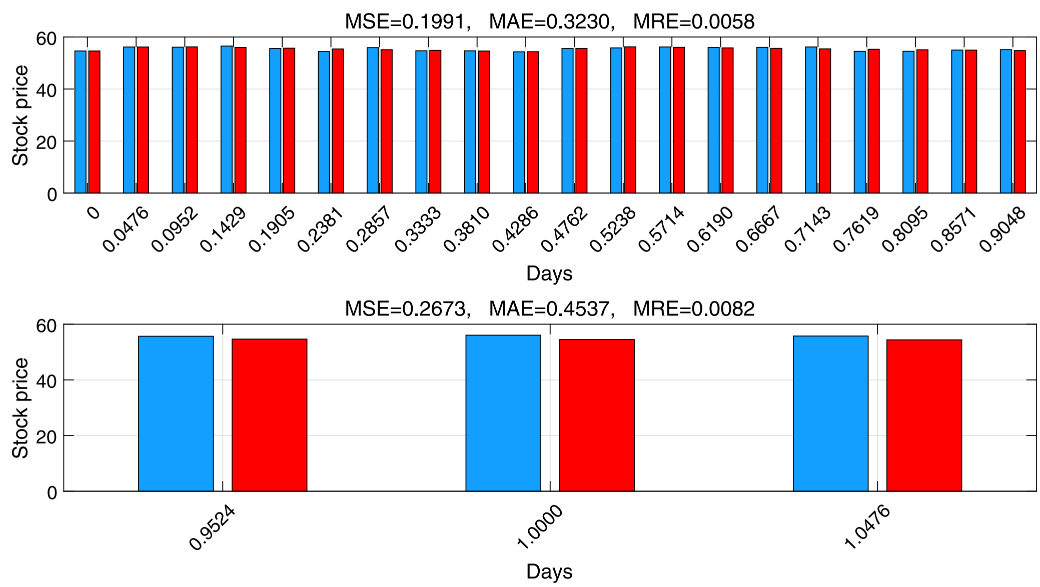

Fig. 12. Comparing observed data (blue color) with trained data (red color) in figure above and test data (red color) in figure below. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

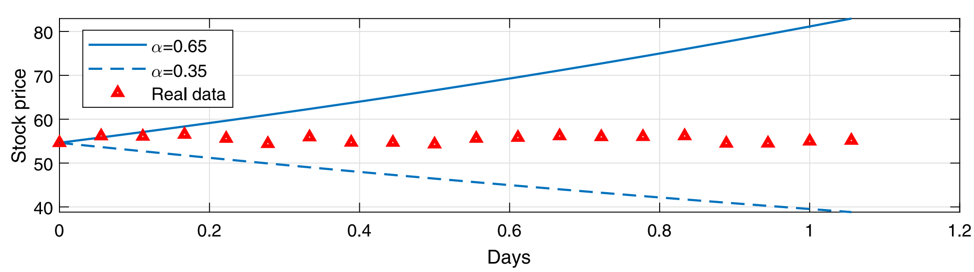

Fig. 13. Actual dataset and α-paths of X_t in Example 4.

Table 8
Comparing the future price of Cisco Systems Corporation with the predicted value obtained by UDE and SDE. [·] denotes the absolute error between the actual value and forecast value.

Trading day	4/1/2022	4/4/2022	4/5/2022
Actual value	55.66	55.87	54.92
Forecast value by the UDE	56.02 [0.36]	55.92 [0.05]	56.13 [1.21]
Forecast value by the SDE	53.33 [2.32]	53.23 [2.63]	53.43 [1.48]

In Fig. 14, the logarithmic values of the objective function in the problem (4.15) based on the NM and PSO optimization algorithms with various number of iterations are graphed. As it turns out, the NM algorithm has less computational cost and in practice it can be applied to solve the parameters estimation problem of the uncertain exponential Ornstein–Uhlenbeck process. It should be noted that from the two stop conditions of algorithms, that is, maximum number of implementations more than 100 or the value of the objective function less than 10e−05, according to the second condition, both NM and PSO algorithms have given the results in Fig. 14.

5. Parameter estimation on highly noisy environments

Some time series related to the behavior of a phenomenon may contain data with high noise or outliers. These factors are observed in the behavior of a time series (such as stock prices) when sudden events such as changes in government policies, system failures, climate change, etc. occur. In this section, assuming that the mentioned events have affected on the behavior of a phenomenon, we evaluate the approach presented to parameters estimation of the uncertain differential equation.

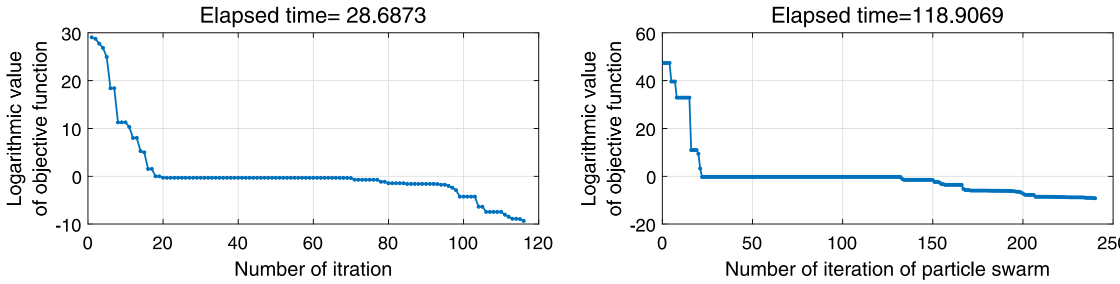

Fig. 14. The logarithmic values of the objective function in problem (4.15), which have been obtained by NM (left figure) and PSO (right figure) algorithms with different number of iterations.

Example 5. Assume that the observed data of a particular stock market is described by the following dynamics

$$dX_t = 0.0893X_t dt + 0.3129\frac{\sqrt{3}}{\pi}\ln\frac{0.85}{0.15}X_t dt + \frac{2}{\epsilon}dW_t, \quad X_0 = 0.5, \tag{5.1}$$

where ϵ is a constant number and dW_t is the Wiener process increment. Assume $\delta := t_i - t_{i-1}$ is the step size and define $\epsilon = \sqrt{\delta}$. In this case, according to the model (5.1), the observed data $x_1, x_2, \ldots, x_N$ can be generated as follows

$$x_i = x_{i-1} + 0.0893\delta x_{i-1} + 0.3129\frac{\sqrt{3}}{\pi}\ln\frac{0.85}{0.15}\delta x_{i-1} + 2Z_i,$$
$$i = 1, 2, \ldots, N, \quad x_0 = 0.5, \tag{5.2}$$

where $Z_1, Z_2, \ldots, Z_N$ are standard normal random variables. We note that the generated dataset in Eq. (5.2) are governed by combining 0.95-path of UDE in (4.4) and Gaussian noise. In addition, to increasing noise in the dataset, we multiplied the Brownian motion increments by 2. Due to the fact that in this case study, our dataset is based on stock prices, thus the dataset used to solve the parameter estimation problem must be positive. On the other hand, the dataset generated from model (5.1) can be negative due to Gaussian noise, so we consider the absolute value of these dataset to estimate the model parameters. According to Table 9, we have 12 observed data $y_0 := |x_0|, \ldots, y_{11} := |x_{11}|$ as training dataset, which $x_0, x_1, \ldots, x_{11}$ are generated by discretized form (5.2). Notice that these 12 dataset $y_0 := y_0, \ldots, y_{11}$ are applied to estimate the weight and bias parameter of the ANN structure by solving Problem II. Fig. 15 compares the observed dataset with trained values. In this figure, the MSE, MAE and MRE criteria for the trained dataset are reported. As it turns out, the optimized ANN has acceptable accuracy and can be used for highly noisy dataset.

We now estimate the parameters involved in the geometric Liu process based on the highly noisy dataset reported in Table 9. Consider the geometric Liu process X_t

$$dX_t = \mu X_t dt + \sigma X_t dC_t, \tag{5.3}$$

where parameters μ and $\sigma > 0$ must be estimated.

According to the Problem III, the estimated values μ^* and σ^* are obtained by solving the non-linear problem based on the NM algorithm:

$$\min_{\mu^*,\sigma^*}\sum_{k=1}^{2}\left(\frac{1}{11}\sum_{i=0}^{11}\left(\widetilde{H}_i(x^{\text{ANN}};\mu_1^*,\mu_2^*,\sigma^*)\right)^k - \left(\frac{\sqrt{3}}{\pi}\right)^k\int_0^1\left(\ln\frac{\alpha}{1-\alpha}\right)^k \mathrm{d}\alpha\right)^2,$$

where

$$\widetilde{H}_i(x^{\text{ANN}};\mu^*,\sigma^*) = \left(\frac{1}{\sigma^* x_{t_i}^{\text{ANN}}}\right)\left(\sum_{\ell=1}^{3} w_{1_{\ell,1}} w_{2_{\ell,1}}(1 - v_\ell^2(t_i)) - \mu^*\right).$$

Estimation of parameters by solving the problem above is obtained as follows

$$\mu^* = 1.0760, \quad \sigma^* = 3.5845.$$

Table 9
Absolute value of data generated by model (5.2).

Day (t_i)	Stock price (y_{t_i})	Day (t_i)	Stock price (y_{t_i})
0	0.5000	0.5455	0.9914
0.0909	1.0207	0.6364	0.9783
0.1818	1.5212	0.7273	1.6779
0.2727	2.5442	0.8182	1.9532
0.3636	2.8526	0.9091	1.6692
0.4545	2.3519	1.0000	1.7721

Moreover, the estimated results of the weight and bias parameters of the proposed ANN by solving the Problem II are obtained as follows

$$w_1 = \begin{pmatrix} 7.0488 \\ -0.3446 \\ 6.0958 \end{pmatrix}, \quad w_2 = \begin{pmatrix} -19.4667 \\ 10.3944 \\ 20.0063 \end{pmatrix}, \quad b_1 = \begin{pmatrix} -3.1680 \\ 5.1787 \\ -2.6923 \end{pmatrix}, \quad b_2 = -9.0333.$$

Then, we have

$$dX_t = 1.0760X_t dt + 3.5845X_t dC_t,$$

Example 6. In this example, we first generate the outliers by the compound Poisson process and next add them to the geometric Liu process (4.4). Assume that the observed data of a particular stock market is described by the following dynamics

$$dX_t = 0.0893X_t dt + 0.3129\frac{\sqrt{3}}{\pi}\ln\frac{0.85}{0.15}X_t dt + \sum_{j=1}^{Y_t} Z_j, \quad X_0 = 1, \tag{5.4}$$

where dC_t is the Liu process increment, Y_t is the Poisson process with intensity rate $\lambda := 1$, and $Z_1, Z_2, \ldots$ are the log-normal random variables with parameters $a := 1$ and $b := 0.6$. It is necessary to mention that putting $\lambda = 1$ means that the average number of outliers in the dataset is equal to 1. Let $\delta := t_i - t_{i-1}$ be the step size. In this case, according to the model (5.4), the observed data $x_1, x_2, \ldots, x_N$ can be generated as follows

$$x_i = x_{i-1} + 0.0893\delta x_{i-1} + 0.3129\frac{\sqrt{3}}{\pi}\ln\frac{0.85}{0.15}\delta x_{i-1} + Q_{t_i} - Q_{t_{i-1}}$$
$$i = 1, 2, \ldots, N, \quad x_0 = 1. \tag{5.5}$$

where $Q_t = \sum_{j=1}^{Y_t} Z_j$. We note that the generated dataset in Eq. (5.5) are governed by combining 0.95-path of UDE in (4.4) and compound Poisson process. According to Table 9, we have 12 observed data $x_0, x_1 \ldots, x_{11}$ as training dataset, which are generated by discretized form (5.5). Notice that these 12 dataset are applied to estimate the weight and bias parameter of the ANN structure by solving Problem II. Fig. 16 compares the observed dataset with trained values. In this figure, the MSE, MAE and MRE criteria for the trained dataset are reported. As it turns out, the optimized ANN has acceptable accuracy and can be used for dataset polluted with outliers.

We now estimate the parameters involved in the geometric Liu process based on the dataset reported in Table 10. Consider the geometric

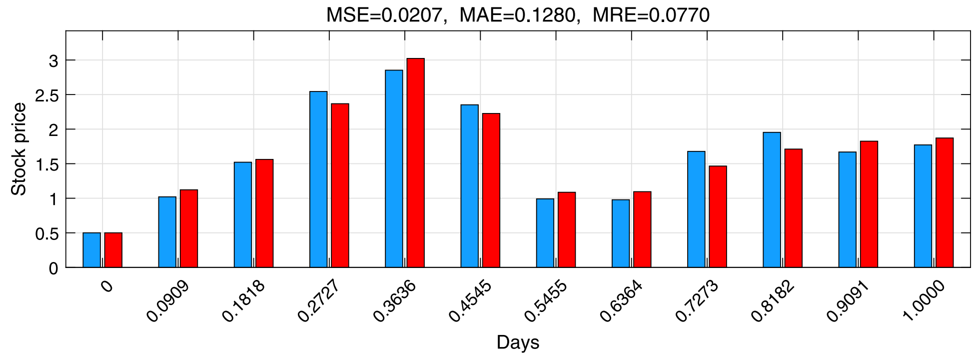

Fig. 15. Comparing observed data (blue color) with trained data (red color). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

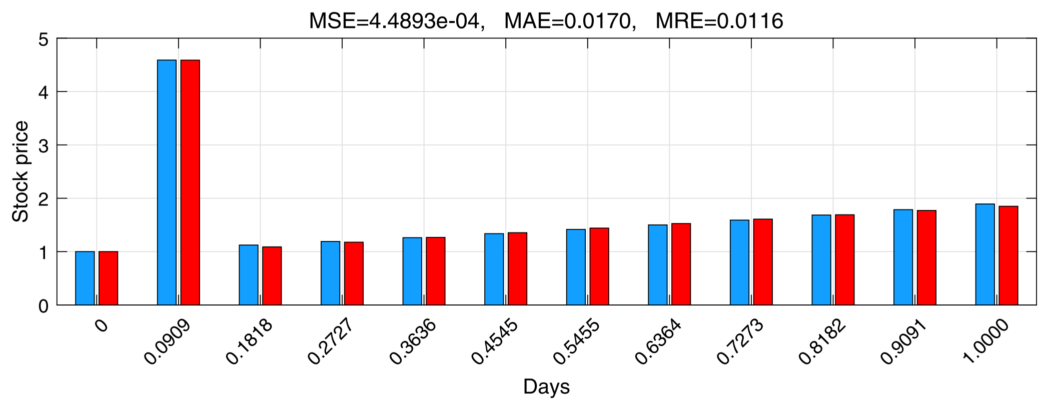

Fig. 16. Comparing observed data (blue color) with trained data (red color). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Liu process X_t

$$dX_t = \mu X_t dt + \sigma X_t dC_t, \tag{5.6}$$

where parameters μ and $\sigma > 0$ must be estimated.

According to the Problem III, the estimated values μ^* and σ^* are obtained by solving the non-linear problem based on the NM algorithm:

$$\min_{\mu^*,\sigma^*} \sum_{k=1}^{2} \left(\frac{1}{11} \sum_{i=0}^{11} \left(\widetilde{H}_i(x^{\text{ANN}}; \mu_1^*, \mu_2^*, \sigma^*) \right)^k - \left(\frac{\sqrt{3}}{\pi} \right)^k \int_0^1 \left(\ln \frac{\alpha}{1-\alpha} \right)^k \mathrm{d}\alpha \right)^2,$$

where

$$\widetilde{H}_i(x^{\text{ANN}}; \mu^*, \sigma^*) = \left(\frac{1}{\sigma^* x_{t_i}^{\text{ANN}}} \right) \left(\sum_{\ell=1}^{3} w_{1_{\ell,1}} w_{2_{\ell,1}} (1 - v_\ell^2(t_i)) - \mu^* \right).$$

Estimation of parameters by solving the problem above is obtained as follows

$$\mu^* = 0.5872, \quad \sigma^* = 0.7722.$$

Moreover, the estimated results of the weight and bias parameters of the proposed ANN by solving the Problem II are obtained as follows

$$w_1 = \begin{pmatrix} -0.9062 \\ -57.4711 \\ -0.1081 \end{pmatrix}, \quad w_2 = \begin{pmatrix} -25.9041 \\ 1.8497 \\ -95.5188 \end{pmatrix}, \quad b_1 = \begin{pmatrix} -41.0076 \\ 6.9999 \\ -1.8077 \end{pmatrix},$$
$$b_2 = -113.6709. \tag{5.7}$$

Then, we have

$$dX_t = 0.5872 X_t dt + 0.7722 X_t dC_t,$$

Table 10
Data generated by model (5.5).

Day (t_i)	Stock price (x_{t_i})	Day (t_i)	Stock price (x_{t_i})
0	1.0000	0.5455	1.4163
0.0909	4.5894	0.6364	1.5009
0.1818	1.1230	0.7273	1.5905
0.2727	1.1901	0.8182	1.6855
0.3636	1.2612	0.9091	1.7862
0.4545	1.3365	1.0000	1.8929

6. Conclusion remarks

In this study, the method of moments based on the artificial neural network was worked out to estimate the uncertain differential equation parameters. This method is based on the structure of the artificial neural networks to approximate non-linear functions with multi input and output. Theoretically, the proposed problem in this paper indicates the local optimization problem, although from applied viewpoint one only requires an appropriate fit of the dataset. As evident in the literature, this issue can be successfully performed based on the artificial neural networks. The main advantage of the proposed method in this paper is that by dividing the main problem into two other sub-problems through artificial neural network, the problem of parameter estimation is independent of the step size. We investigated the proposed method on various uncertain differential equations with various financial dataset and obtained satisfactory results. For the proposed parameter estimation problem, we also compared the local optimization method based on the Nelder–Mead algorithm and the global search method based on the particle swarm optimization algorithm. We found that the Nelder–Mead algorithm has less computational time and acceptable accuracy. Ultimately, the introduced approach for the parameter estimation of the uncertain differential equation on the dataset polluted by high noise and outliers was evaluated. Numerical results established that the suggested method based on the artificial neural network structure gives appropriate accuracy. Future study will centralize on parameter estimation of financial derivatives and calibrating them through in-sample and out-of-sample dataset.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that support the findings of this study are available from www.finance.yahoo.com.

Ethical approval

This article does not contain any studies with human participants performed by any of the authors.

References

[1] Liu B. Why is there a need for uncertainty theory. J Uncertain Syst 2012;6(1):3–10.
[2] Liu B. Uncertain urn problems and ellsberg experiment. Soft Comput 2019;23(15):6579–84.
[3] Hassanzadeh S, Mehrdoust F. Valuation of European option under uncertain volatility model. Soft Comput 2018;22(12):4153–63.
[4] Hassanzadeh S, Mehrdoust F. European option pricing under multifactor uncertain volatility model. Soft Comput 2020;24(12):8781–92.
[5] Mehrdoust F, Najafi AR. An uncertain exponential Ornstein–Uhlenbeck interest rate model with uncertain CIR volatility. Bull Iran Math Soc 2020;46(5):1405–20.
[6] Noorani I, Mehrdoust F, Lio W. Electricity spot price modeling by multi-factor uncertain process: a case study from the Nordic region. Soft Comput 2021;25(21):13105–26.
[7] Mehrdoust F, Noorani I, Wei Xu. Uncertain energy model for electricity and gas futures with application in spark-spread option price. Fuzzy Optim Decis Mak 2022. http://dx.doi.org/10.1007/s10700-022-09386-z.
[8] Liu B. Uncertain risk analysis and uncertain reliability analysis. J Uncertain Syst 2010;4(3):163–70.
[9] Liu B. Uncertain logic for modeling human language. J Uncertain Syst 2011;5(1):3–20.
[10] Yao K, Liu B. Parameter estimation in uncertain differential equations. Fuzzy Optim Decis Mak 2020;19(1):1–12.
[11] Sheng Y, Yao K, Chen X. Least squares estimation in uncertain differential equations. IEEE Trans Fuzzy Syst 2020;28(10):2651–5.
[12] Yang X, Liu Y, Park G. Parameter estimation of uncertain differential equation with application to financial market. Chaos Solitons Fractals 2020;139:110026.
[13] Liu Y, Liu B. Estimating unknown parameters in uncertain differential equation by maximum likelihood estimation. Soft Comput 2022;26(6):2773–80.
[14] Liu Z. Generalized moment estimation for uncertain differential equations. Appl Math Comput 2021;392:125724.
[15] Liu Y, Liu B. Residual analysis and parameter estimation of uncertain differential equations. Fuzzy Optim Decis Mak 2022. http://dx.doi.org/10.1007/s10700-021-09379-4.
[16] Liu Z, Yang X. A linear uncertain pharmacokinetic model driven by Liu process. Appl Math Model 2021;89(2):1881–99.
[17] Tang H, Yang X. Uncertain chemical reaction equation. Appl Math Comput 2021;411:126479.
[18] Lio W, Liu B. Initial value estimation of uncertain differential equations and zero-day of COVID-19 spread in China. Fuzzy Optim Decis Mak 2021;20(2):177–88.
[19] Dua V. A mixed-integer programming approach for optimal configuration of artificial neural networks. Chem Eng Res Des 2010;88(1):55–60.
[20] Xie Z, Kulasiri D, Samarasinghe S, Rajanayaka C. The estimation of parameters for stochastic differential equations using neural networks. Inverse Probl Sci Eng 2007;15(6):629–41.
[21] Dua V. An artificial neural network approximation based decomposition approach for parameter estimation of system of ordinary differential equations. Comput Chem Eng 2011;35(3):545–53.
[22] Jamili E, Dua V. Parameter estimation of partial differential equations using artificial neural network. Comput Chem Eng 2021;147:107221.
[23] Wu J, Fukuhara M, Takeda T. Parameter estimation of an ecological system by a neural network with residual minimization training. Ecol Model 2005;189(3–4):289–304.
[24] Liu B. Uncertainty theory. 2nd ed.. Berlin: Springer; 2009.
[25] Liu B. Some research problems in uncertainty theory. J Uncertain Syst 2009;3(1):3–10.
[26] Liu B. Uncertainty theory. 4th ed.. Berlin: Springer; 2015.
[27] Yao K, Chen X. A numerical method for solving uncertain differential equations. J Intell Fuzzy Systems 2013;25(3):825–32.
[28] Platon R, Dehkordi VR, Martel J. Hourly prediction of a building's electricity consumption using case-based reasoning, artificial neural networks and principal component analysis. Energy Build 2015;92:10–8.
[29] Günay ME. Forecasting annual gross electricity demand by artificial neural networks using predicted values of socio-economic indicators and climatic conditions: Case of Turkey. Energy Policy 2016;90:92–101.
[30] Nelder JA, Mead R. A simplex method for function minimization. Comput J 1965;7(4):308–13.
[31] Boyd S, Boyd SP, Vandenberghe L. Convex optimization. Cambridge University Press; 2004.
[32] Kennedy J, Eberhart R. Particle swarm optimization. In: Proceedings of ICNN'95-international conference on neural networks. Vol. 4, IEEE; 1995, p. 1942–8.