

Article

Parameter Estimation of General Uncertain Differential Equations via the Principle of Least Squares with Its Application in Economic Field

Xiaoya Xu and Youde Dong *

School of Economics, Shanghai University, Shanghai 200444, China; xxy150716@163.com

* Correspondence: dongyoude415@163.com

Abstract

The parameter estimation problem is one of the research hotspots in the field of uncertain differential equations. However, most studies at present focus on parameter estimation based on residuals of uncertain differential equations, which relies strictly on the solvability of residuals. In view of this disadvantage, this paper derives a symmetrical statistical invariant, which is different from residuals based on the difference scheme, and proposes the least squares estimation of general uncertain differential equations based on the statistical invariant and the principle of least squares. In order to consider parameter estimation in more general cases, this paper also studies the least squares estimation of time-varying parameters in general uncertain differential equations and designs corresponding to numerical algorithms to calculate the numerical solutions of these least squares estimations. Finally, this paper also proposes two numerical examples and an empirical study to illustrate the above methods.

Keywords: uncertainty theory; uncertain differential equations; parameter estimation; the principle of least squares; symmetrical statistical invariant



Academic Editor: Hari Mohan Srivastava

Received: 15 August 2025

Revised: 13 September 2025

Accepted: 17 September 2025

Published: 24 September 2025

Citation: Xu, X.; Dong, Y. Parameter Estimation of General Uncertain Differential Equations via the Principle of Least Squares with Its Application in Economic Field. *Symmetry* **2025**, *17*, 1594. <https://doi.org/10.3390/sym17101594>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In order to more realistically describe time-varying systems in the real world, there are two mathematical tools in total. One is the stochastic differential equation based on probability theory (proposed by Kolmogorov [1]), and the other is the uncertain differential equation based on uncertainty theory (proposed by Liu [2]). Since their proposal, stochastic differential equations have been extensively studied by numerous scholars, giving rise to numerous research branches, and successfully applied in fields such as financial research, social sciences, physical systems, and chemical evolution.

Although stochastic differential equations have produced many achievements, most studies on stochastic differential equations only focus on their theoretical results and do not discuss their applicability based on actual data. In other words, the theoretical basis of stochastic differential equations is probability theory, and the premise that probability theory can reasonably model the real world is based on the condition that the frequency of the observed data we obtain is stable enough, which is a strong enough condition. In fact, a large number of empirical studies have proven that the observed data in practice do not meet the conditions of frequency stability. At this time, the distribution function generated based on historical data cannot be close enough to the actual frequency of the research object. If it is regarded as a probability distribution function, and the stochastic differential

equation based on probability theory is used to describe the research object, unreasonable results will be obtained. For example, Yang and Ke [3] showed that stochastic differential equations are not suitable for portraying Shanghai interbank offered rates, which in turn shows that the actual data in the financial system cannot meet the frequency stability conditions. Yang and Liu [4] illustrated that if population size data are characterized by using stochastic partial differential equations, unstable frequency results will also be obtained, and Liu et al. [5] modeled online ride-hailing data in Beijing, China, by using the stochastic renewal process, and obtained results that violate common sense. These results all show that the actual data in the social system are also frequency unstable. In addition to showing, from the perspective of empirical research, that stochastic differential equations are not suitable for modeling the real world, Liu [6] also pointed out that the velocity of Brownian motion B_t is infinite (that is, the derivative of Brownian motion B_t at any moment is a normal random variable whose expected value is 0, and the variance is infinite) and theoretically deduced the paradox of the stochastic differential equation. Since the real world cannot guarantee the fulfillment of the assumption of frequency stability, stochastic differential equations cannot reasonably model time-varying systems in the real world. In order to solve the problem of how to reasonably describe the phenomenon of frequency instability, Liu [2] proposed uncertainty theory in 2007, which is an axiomatic mathematical system based on normativity, duality, subadditivity, and the product axioms. To date, uncertainty theory has been studied by many scholars and has spawned many research branches, such as the uncertain differential equation [7], uncertain optimal control [8], uncertain renewal process [9], uncertain graph [10], etc.

Among the many theoretical branches of uncertainty theory, uncertain differential equations hold a crucial position as a powerful tool for characterizing the evolution of dynamic systems over time under uncertain environments. They were proposed by Liu [11] as a class of differential equations driven by Liu processes. Since their proposal, they have attracted the attention of many scholars. In theoretical research, Chen and Liu [12] first investigated the existence and uniqueness conditions for solutions of uncertain differential equations. Based on this work, Liu [13] began research on stability for uncertain differential equations. Subsequently, Yao et al. [14] proposed several stability theorems, further advancing the field of stability analysis for uncertain differential equations. The Yao–Chen formula developed by Yao and Chen [15] is a significant contribution to the study of uncertain differential equations, which links uncertain differential equations with ordinary differential equations, demonstrating that the solution to an uncertain differential equation can be represented by the solutions of a family of ordinary differential equations. Therefore, the Yao–Chen formula provides an accurate and elegant way to obtain numerical solutions to uncertain differential equations. Based on this, Yao and Chen [15] first proposed a numerical method for solving uncertain differential equations. Based on the above pioneering work, many scholars have devoted themselves to researching and improving on the theoretical work of uncertain differential equations, producing a large number of results.

The application of uncertain differential equations in practice is the problem of parameter estimation, which is also the current research hotspot. However, the current research on parameter estimation of uncertain differential equations mainly focuses on methods based on the residuals of uncertain differential equations. Although the parameter estimation methods based on the residuals have high accuracy, such methods strictly rely on the existence of analytical solutions to uncertain differential equations or the existence of the inverse uncertainty distributions of their solutions. As we all know, only a few uncertain differential equations can obtain corresponding analytical solutions, and only regular uncertain differential equations can be solved numerically and obtain inverse uncertainty distributions. Therefore, parameter estimation methods based on the residuals of uncertain

differential equations have certain limitations. In order to study the least squares estimation of general uncertain differential equations, this paper will construct a symmetric statistical invariant based on the difference scheme of uncertain differential equations, and study its least squares estimation based on this statistical invariant and the principle of least squares. Specifically, the main contributions of this paper are as follows:

- A statistical invariant based on the difference scheme of uncertain differential equations has been constructed, and for general uncertain differential equations, the analytical expression of this statistical invariant can be obtained. This resolves the problem that the residual-based parameter estimation method for uncertain differential equations is limited by the solvability of the uncertainty distribution or the inverse uncertainty distribution of the solution.
- The least squares estimations of the constant parameters and time-varying parameters of general uncertain differential equations are proposed. This solves the problem of applying the least squares method based on uncertainty distribution to the parameter estimation of constant parameters and time-varying parameters in uncertain differential equations.
- Two numerical algorithms are designed to solve the numerical solution of the aforementioned least squares estimations.
- Two numerical examples and an empirical study are provided to illustrate the corresponding methods proposed in this paper.

2. Literature Review of Parameter Estimation of Uncertain Differential Equations

Parameter estimation of uncertain differential equations is the current research hotspot in this field and is also the research topic of this paper. Therefore, this section will conduct a literature review on parameter estimation of uncertain differential equations.

Parameter estimation for uncertain differential equations originated from the work of Yao and Liu [16], which first proposed moment estimation for uncertain differential equations based on their difference scheme and the method of moments. However, when there are too many unknown parameters, moment estimation may not be possible. To address this issue, Liu [17] further developed the generalized moment estimation of uncertain differential equations by minimizing the sum of squares of the deviations between each order of sample moments and each order of population moments. Furthermore, Sheng et al. [18] investigated a least squares estimation for uncertain differential equations by minimizing the noise terms. Yang et al. [19] explored the minimum covering estimation by ensuring that the observed data fall within a reasonable range of the α paths. Liu and Liu [20] borrowed the idea of uncertain maximum likelihood to present the maximum likelihood estimation for uncertain differential equations. These works are early achievements in this field and are all based on difference schemes for uncertain differential equations. Their disadvantage is that these difference schemes can introduce errors, and this error is particularly significant when the time step of the observation data is not sufficiently small, which is usually impossible to control.

In order to address the problems caused by the difference schemes, Liu and Liu [21] proposed the concept of residuals for uncertain differential equations and established a precise connection between uncertain differential equations and observed data. Since then, the focus of many scholars has shifted to residual-based parameter estimation methods for uncertain differential equations. Based on the residuals of uncertain differential equations, Liu and Liu [21] first studied residual-based moment estimation. Then, Liu and Liu borrowed the ideas of uncertain maximum likelihood and least squares criteria to explore residual-based maximum likelihood estimation (Liu and Liu [22]) and residual-based

least squares estimation (Liu and Liu [23]), respectively. In addition, scholars have also explored the parameter estimation of multi-factor uncertain differential equations, including residual-based moment estimation (Yao and Sheng [24]), residual-based maximum likelihood estimation (Liu et al. [25]), and residual-based least squares estimation (Wu and Liu [26]). Some scholars have also conducted research on parameter estimation of a class of uncertain partial differential equations. For example, Yang and Liu [4] studied residual-based moment estimation of a class of uncertain partial differential equations, and Yang and Liu [27] proposed residual-based least squares estimation and applied it to modeling the Chinese population. These studies are precisely the focus of parameter estimation in uncertain differential equations at the methodological level, and the advantage of these methods is that they are based on residuals that accurately connect the uncertain differential equations to the observed data, resulting in relatively more accurate results. However, their disadvantage is that they rely heavily on the solvability of the residuals. For the general uncertain differential equations that do not satisfy regularity conditions, the residuals are essentially unobtainable. In such cases, these methods will fail, and we will be forced to resort to parameter estimation methods based on difference schemes.

3. Preliminary

In this section, we will provide some concepts and theorems about uncertainty theory and uncertain differential equations to ensure that readers can understand the following content more easily.

The uncertain measure is the foundation of uncertainty theory, which is actually a set function defined on a σ -algebra, and the specific definition is as follows:

Definition 1 (Liu [2]). Let Γ denote the universal set, and let $\mathbb{L}$ denote the σ -algebra over Γ . If the measurable set function $\mathcal{M}$ defined on the σ -algebra $\mathbb{L}$ satisfies the following three conditions, then the measurable set function $\mathcal{M}$ is called an uncertain measure, and thus $(\Gamma, \mathbb{L}, \mathcal{M})$ is called an uncertainty space:

Normality Axiom: $\mathcal{M}\{\Gamma\} = 1$ for the universal set Γ .

Duality Axiom: For any event $\Lambda \in \mathbb{L}$, $\mathcal{M}\{\Lambda\} + \mathcal{M}\{\Lambda^c\} = 1$ always holds.

Subadditivity Axiom: For every countable sequence of events $\Lambda_1, \Lambda_2, \dots$,

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

always holds.

For the purpose of obtaining the uncertain measure of a composite event, the product uncertain measure $\mathcal{M}$ on the product σ -algebra $\mathbb{L}$ was defined by Liu [13] by assuming it follows the product axiom. That is, let $(\Gamma_i, \mathbb{L}_i, \mathcal{M}_i)$, $i = 1, 2, \dots$ be a sequence of uncertainty spaces. If the uncertain measure $\mathcal{M}$ defined on the σ -algebra $\mathbb{L}$ satisfies

$$\mathcal{M}\left\{\prod_{i=1}^{\infty} \Lambda_i\right\} = \bigwedge_{i=1}^{\infty} \mathcal{M}_i\{\Lambda_i\},$$

where $\Lambda_i \in \mathbb{L}$ for $i = 1, 2, \dots$, then the uncertain measure $\mathcal{M}$ is called a product uncertain measure. Based on the uncertainty space $(\Gamma, \mathbb{L}, \mathcal{M})$, an instrument used to describe quantities in uncertain environments is the uncertain variable, which was defined by Liu [2] as a

measurable function from an uncertainty space $(\Gamma, \mathbb{L}, \mathcal{M})$ to the set of real numbers and was denoted by ξ , i.e., the set

$$\{\xi \in B\} = \{\gamma \in \Gamma \mid \xi(\gamma) \in B\}$$

is always an event for any Borel set B of real numbers. Subsequently, in order to facilitate the calculation of uncertain variables in practical situations, the uncertainty distribution of an uncertain variable ξ was also defined by Liu [2] as

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

Example 1. If the uncertain variable ξ has a normal uncertainty distribution

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

then it is called a normal uncertain variable, and is represented as $\sim \mathcal{N}(e, \sigma)$, where e is the expected value and σ is the standard deviation. In addition, a normal uncertainty distribution is called standard if $e = 0$ and $\sigma = 1$.

In practice, only independent uncertain variables can be subjected to operations. In order to define what kind of uncertain variables are independent, Liu [13] suggested that if the uncertain variables $\xi_1, \xi_2, \dots, \xi_n$ satisfy

$$\mathcal{M}\left\{\bigcap_{i=1}^n (\xi_i \in B_i)\right\} = \bigwedge_{i=1}^n \mathcal{M}\{\xi_i \in B_i\}$$

for any Borel sets $B_1, B_2, \dots, B_n$ of real numbers, then the uncertain variables $\xi_1, \xi_2, \dots, \xi_n$ can be called independent. Furthermore, in order to provide a more accurate description of time-varying systems in uncertain environments, the Liu process C_t was defined by Liu [13] as follows:

Definition 2. An uncertain process C_t is said to be a Liu process if

- (i) $C_0 = 0$, and almost all sample paths are Lipschitz continuous;
- (ii) C_t has stationary and independent increments;
- (iii) every increment $C_{s+t} - C_s$ is a normal uncertain variable with expected value 0 and standard deviation t .

4. Symmetric Statistical Invariant Based on Difference Scheme

In this section, we consider the general uncertain differential equation expressed in the following form:

$$dX_t = f(t, X_t; \theta)dt + g(t, X_t; \theta)dC_t, \quad (1)$$

where $f(t, x)$ and $g(t, x)$ are two continuous real-valued functions satisfying the existence and uniqueness conditions for the solution of uncertain differential Equation (1), i.e., linear growth condition

$$|f(t, x)| + |g(t, x)| \leq L(1 + |x|), \quad \forall t \geq 0, \forall x, y \in \mathcal{R},$$

and Lipschitz condition

$$|f(t, x_1) - f(t, x_2)| + |g(t, x_1) - g(t, x_2)| \leq L|x_1 - x_2|, \quad \forall t \geq 0, \forall x_1, x_2 \in \mathcal{R}$$

for some constant L , θ is a vector of unknown parameters to be estimated, and C_t is a Liu process.

Suppose also that there exists a set of observed data

$$x_{t_1}, x_{t_2}, \dots, x_{t_n} \quad (2)$$

of the solution X_t to uncertain differential Equation (1) at time points $t_1, t_2, \dots, t_n$ with $t_1 < t_2 < \dots < t_n$, respectively. In order to link the uncertain differential Equation (1) with the observed data (2), we first discretize uncertain differential Equation (1) at the observed time points $t_1, t_2, \dots, t_n$. That is, we choose to write uncertain differential Equation (1) into the Euler difference form,

$$X_{t_i} = X_{t_{i-1}} + f(t_{i-1}, X_{t_{i-1}}; \theta)(t_i - t_{i-1}) + g(t_{i-1}, X_{t_{i-1}}; \theta)(C_{t_i} - C_{t_{i-1}}), \quad i = 2, 3, \dots, n.$$

Based on the above Euler difference form, we can obtain

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

for $i = 2, 3, \dots, n$. For each index i with $2 \leq i \leq n$, it follows from Definition 2 that the increments $C_{t_i} - C_{t_{i-1}}$ of the Liu process C_t over the time intervals $[t_{i-1}, t_i]$ are a normal uncertain variable with expected value 0 and standard deviation $t_i - t_{i-1}$. That is,

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

Thus, we have

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

Therefore, it is easy to infer that $\frac{C_{t_i} - C_{t_{i-1}}}{t_i - t_{i-1}}, i = 2, 3, \dots, n$ are always independent of each other and

$$\frac{C_{t_i} - C_{t_{i-1}}}{t_i - t_{i-1}} \sim \mathcal{N}(0, 1), \quad i = 2, 3, \dots, n$$

always holds. By using the above transformation formula of the Euler difference form of uncertain differential Equation (1), we can also infer that

$$\frac{X_{t_i} - X_{t_{i-1}} - f(t_{i-1}, X_{t_{i-1}}; \theta)(t_i - t_{i-1})}{g(t_{i-1}, X_{t_{i-1}}; \theta)(t_i - t_{i-1})}, \quad i = 2, 3, \dots$$

are always independent of each other and

$$\frac{X_{t_i} - X_{t_{i-1}} - f(t_{i-1}, X_{t_{i-1}}; \theta)(t_i - t_{i-1})}{g(t_{i-1}, X_{t_{i-1}}; \theta)(t_i - t_{i-1})} \sim \mathcal{N}(0, 1), \quad i = 2, 3, \dots, n$$

always holds. For the sake of convenience, we write

$$\zeta_i = \frac{X_{t_i} - X_{t_{i-1}} - f(t_{i-1}, X_{t_{i-1}}; \theta)(t_i - t_{i-1})}{g(t_{i-1}, X_{t_{i-1}}; \theta)(t_i - t_{i-1})}, \quad i = 2, 3, \dots, n.$$

Then, it follows from the above analysis that

$$\zeta_i \sim \mathcal{N}(0, 1), \quad i = 2, 3, \dots, n. \quad (3)$$

On the other hand, we substitute the observed data (2) into

$$\frac{X_{t_i} - X_{t_{i-1}} - f(t_{i-1}, X_{t_{i-1}}; \theta)(t_i - t_{i-1})}{g(t_{i-1}, X_{t_{i-1}}; \theta)(t_i - t_{i-1})}, i = 2, 3, \dots, n,$$

and write

$$h_i(\theta) = \frac{x_{t_i} - x_{t_{i-1}} - f(t_{i-1}, x_{t_{i-1}}; \theta)(t_i - t_{i-1})}{g(t_{i-1}, x_{t_{i-1}}; \theta)(t_i - t_{i-1})}, i = 2, 3, \dots, n. \quad (4)$$

Note that $h_i(\theta)$, $i = 2, 3, \dots, n$ are $n - 1$ real-valued functions with respect to the vector of unknown parameters θ and can be regarded as the samples of ξ_i , $i = 2, 3, \dots, n$, respectively. That is, we should have

$$h_2(\theta), h_3(\theta), \dots, h_n(\theta) \sim \mathcal{N}(0, 1). \quad (5)$$

At this point, we have constructed a series of real-valued functions $h_i(\theta)$, $i = 2, 3, \dots, n$ of the vector of unknown parameters θ based on uncertain differential Equation (1) and its observed data (2), and processed the population distribution of the real-valued functions $h_i(\theta)$, $i = 2, 3, \dots, n$ as a statistical invariant $\mathcal{N}(0, 1)$. Since the uncertainty distribution of $\mathcal{N}(0, 1)$ is symmetric about the origin, this statistical invariant is also called symmetric.

5. Least Squares Estimation of General Uncertain Differential Equations

In order to apply the least squares principle when estimating unknown parameters, Liu and Liu [23] proposed a least squares estimation based on uncertainty distribution by minimizing the sum of squares of the deviations between the empirical distribution of samples and the distribution function of the population. In this section, we will apply this method to the parameter estimation of general uncertain differential equations.

5.1. Least Squares Estimation Based on Uncertainty Distribution

Assume that there exists a population ξ in uncertain environments, whose uncertainty distribution is denoted by $\Phi_\theta(x)$, and θ is a vector of unknown parameters. Suppose also there exists a set of observed data

$$x_1, x_2, \dots, x_n \quad (6)$$

which can be regarded as the observed values of the population ξ . Based on this, Liu and Liu [23] suggested that the least squares estimation of the vector of unknown parameters θ should minimize the deviations between the empirical distribution of observed data $x_1, x_2, \dots, x_n$ and the uncertainty distribution $\Phi_\theta(x)$. That is, the least squares estimation $\hat{\theta}$ should be the solution of

$$\min_{\theta} \sum_{i=1}^n (\Phi_\theta(x_i) - F(x_i))^2. \quad (7)$$

Here, $F(x)$ is the empirical distribution of observed data $x_1, x_2, \dots, x_n$ and is defined as

$$F(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{I}(x_i \leq x), \quad (8)$$

where

$$\mathbb{I}(x_i \leq x) = \begin{cases} 1 & \text{if } x_i \leq x \\ 0 & \text{otherwise} \end{cases}$$

is the indicator function.

5.2. Least Squares Estimation of General Uncertain Differential Equations with Constant Parameters

For the general uncertain differential Equation (1) with the vector of unknown parameters θ and the given observed data (2), we can always obtain a set of independent and identically distributed samples of the symmetric statistical invariant $\mathcal{N}(0, 1)$, i.e.,

$$h_2(\theta), h_3(\theta), \dots, h_n(\theta) \sim \mathcal{N}(0, 1). \quad (9)$$

Next, we will apply the least squares estimation method based on the uncertainty distribution to estimate the unknown parameter vector θ . It is easy to infer that the empirical distribution of $h_2(\theta), h_3(\theta), \dots, h_n(\theta)$ is

$$F(x|\theta) = \frac{1}{n-1} \sum_{j=2}^n \mathbb{I}(h_j(\theta) \leq x),$$

and the uncertainty distribution of $\mathcal{N}(0, 1)$ is

$$\Phi(x) = \frac{1}{1 + \exp\left(\frac{-\pi x}{\sqrt{3}}\right)}.$$

Then, it follows from (7) that the least squares estimation of unknown parameter vector θ can be obtained by solving the following minimization problem:

$$\min_{\theta} \sum_{i=2}^n \left(\frac{1}{1 + \exp\left(\frac{-\pi h_i(\theta)}{\sqrt{3}}\right)} - \frac{1}{n-1} \sum_{j=2}^n \mathbb{I}(h_j(\theta) \leq h_i(\theta)) \right)^2. \quad (10)$$

The solution $\hat{\theta}$ of the above minimization problem is also called the least squares estimation of the general uncertain differential Equation (1).

In order to ensure that the solution to the least squares estimation can be truly achieved, we will design a numerical algorithm.

Numerical algorithms for least squares estimation

Step 0: Input the general uncertain differential equation

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

that needs to be estimated and the observed data $x_{t_1}, x_{t_2}, \dots, x_{t_n}$.

Step 1: Determine the feasible regions Θ of the unknown parameters vector θ .

Step 2: For each $\theta \in \Theta$, compute $h_j(\theta)$ by

$$h_j(\theta) = \frac{x_{t_j} - x_{t_{j-1}} - f(t_{j-1}, x_{t_{j-1}}; \theta)(t_j - t_{j-1})}{g(t_{j-1}, x_{t_{j-1}}; \theta)(t_j - t_{j-1})}$$

for $j = 2, 3, \dots, n$.

Step 3: Set $i = 2$ and $Obj(\theta) = 0$.

Step 4: Set

$$Obj(\theta) = Obj(\theta) + \left(\frac{1}{1 + \exp\left(\frac{-\pi h_i(\theta)}{\sqrt{3}}\right)} - \frac{1}{n-1} \sum_{j=2}^n \mathbb{I}(h_j(\theta) \leq h_i(\theta)) \right)^2,$$

and $i = i + 1$.

Step 5: If $i \leq n$, then go to Step 4.

Step 6: Find $\hat{\theta}$ such that $Obj(\theta)$ reaches its minimum value by using MATLAB with version number of '9.13.0.2049777 (R2022b)'.

Step 7: Output $\hat{\theta}$.

6. Least Squares Estimation of General Uncertain Differential Equations with Time-Varying Parameters

Generally speaking, the unknown parameters in uncertain differential equations tend to be time-varying rather than constant. Building on this, we will further investigate the least squares estimation problem for uncertain differential equations with time-varying parameters.

Here, we consider the following uncertain differential equations with time-varying parameters,

$$dX_t = f(t, X_t; \theta_t)dt + g(t, X_t; \theta_t)dC_t, \quad (11)$$

where all the assumptions are the same as those in (1), with the exception that θ_t is assumed to be a time-varying parameter related to time t . On the other hand, we also assume that there exists a set of observed data

$$x_{t_1}, x_{t_2}, \dots, x_{t_n} \quad (12)$$

of the solution X_t to uncertain differential Equation (11) at time points $t_1, t_2, \dots, t_n$ with $t_1 < t_2 < \dots < t_n$, respectively. By using the method proposed in Section 4, we can also obtain $n - 1$ samples of the symmetric statistical invariant $\mathcal{N}(0, 1)$, where the time-varying parameter θ_t serves as the independent variable. That is,

$$h_2(\theta_t), h_3(\theta_t), \dots, h_n(\theta_t) \sim \mathcal{N}(0, 1). \quad (13)$$

Before estimating the time-varying parameters θ_t , we first set a sliding window k , which is a given positive integer within an appropriate range. Then, for each index j with $2 \leq j \leq n - k + 1$, the sequence $h_j(\theta_t), h_{j+1}(\theta_t), \dots, h_{j+k-1}(\theta_t)$ can also be regarded as a set of samples of $\mathcal{N}(0, 1)$. In this time, since the time period for the observed data is solely within the interval $[t_{j-1}, t_{j+k-1}]$, we can assume that the variations of the time-varying parameters θ_t within this period are small. Thus, the time-varying parameters θ_t can be regarded as a constant $\theta_{t_{j-1}}$ in this period. Therefore, we can obtain the least squares estimation $\hat{\theta}_{t_{j-1}}$ of $\theta_{t_{j-1}}$ by solving the following minimization problem:

$$\min_{\theta_{t_{j-1}}} \sum_{i=j}^{j+k-1} \left(\frac{1}{1 + \exp\left(\frac{-\pi h_i(\theta_{t_{j-1}})}{\sqrt{3}}\right)} - \frac{1}{k} \sum_{l=j}^{j+k-1} \mathbb{I}(h_l(\theta_{t_{j-1}}) \leq h_i(\theta_{t_{j-1}})) \right)^2. \quad (14)$$

Iterating over j from 2 to $n - k + 1$, and repeating the above operation, we can obtain a set of least squares estimations

$$\hat{\theta}_{t_1}, \hat{\theta}_{t_2}, \dots, \hat{\theta}_{t_{n-k}} \quad (15)$$

for the time-varying parameters θ_t . Then, we select an appropriate uncertain regression model to fit the set of least squares estimations (15), and the resulting regression equation is just the least squares estimation of the time-varying parameters θ_t .

In order to ensure that the solution to the least squares estimation of the time-varying parameters θ_t can be truly achieved, we will design a numerical algorithm.

Numerical algorithms for least squares estimation with time-varying parameters

Step 0: Input the general uncertain differential equation with time-varying parameters,

$$dX_t = f(t, X_t; \theta_t)dt + g(t, X_t; \theta_t)dC_t,$$

that needs to be estimated, the observed data $x_{t_1}, x_{t_2}, \dots, x_{t_n}$, and the sliding window k .

Step 1: Determine the feasible regions Θ_t of the unknown parameters vector θ_t .

Step 2: Set $j = 2$.

Step 3: For each $\theta_{t_{j-1}} \in \Theta_{t_{j-1}}$, compute $h_i(\theta_{t_{j-1}})$ by

$$h_i(\theta_{t_{j-1}}) = \frac{x_{t_i} - x_{t_{i-1}} - f(t_{i-1}, x_{t_{i-1}}; \theta_{t_{j-1}})(t_i - t_{i-1})}{g(t_{i-1}, x_{t_{i-1}}; \theta_{t_{j-1}})(t_i - t_{i-1})}$$

for $i = j, j+1, \dots, j+k-1$.

Step 4: Set $i = j$ and $Obj(\theta_{t_{j-1}}) = 0$.

Step 5: Set

$$Obj(\theta_{t_{j-1}}) = Obj(\theta_{t_{j-1}}) + \left(\frac{1}{1 + \exp\left(\frac{-\pi h_i(\theta_{t_{j-1}})}{\sqrt{3}}\right)} - \frac{1}{k} \sum_{l=j}^{j+k-1} \mathbb{I}(h_l(\theta_{t_{j-1}}) \leq h_i(\theta_{t_{j-1}})) \right)^2,$$

and $i = i + 1$.

Step 6: If $i \leq j + k - 1$, then go to Step 5.

Step 7: Find $\hat{\theta}_{t_{j-1}}$ such that $Obj(\theta_{t_{j-1}})$ reaches its minimum value by using MATLAB, and set $j = j + 1$.

Step 8: If $j \leq n - k + 1$, then go to Step 3.

Step 9: Select the uncertain regression model to fit $\hat{\theta}_{t_1}, \hat{\theta}_{t_2}, \dots, \hat{\theta}_{t_{n-k}}$, and obtain the regression equation $\hat{\theta}_t$.

Step 10: Output $\hat{\theta}_t$.

7. Numerical Examples

In this section, two numerical examples will be provided to illustrate the least squares estimation of the constant parameters and time-varying parameters that were previously proposed.

Example 2. Consider the following linear uncertain differential equation:

$$dX_t = adt + bdC_t, \quad (16)$$

where a and b are two parameters that need to be estimated. Suppose also there is a dataset consisting of 45 sets of observational data (t_i, x_{t_i}) , $i = 1, 2, \dots, 45$, as shown in Table 1.

In order to use linear uncertain differential Equation (16) to describe the observational data in Table 1, we first need to estimate the unknown parameters a and b . By substituting the observational data (t_i, x_{t_i}) , $i = 1, 2, \dots, 45$ into (4), we can obtain

$$h_i(a, b) = \frac{x_{t_i} - x_{t_{i-1}} - a(t_i - t_{i-1})}{b(t_i - t_{i-1})} = \frac{1}{b} \left(\frac{x_{t_i} - x_{t_{i-1}}}{t_i - t_{i-1}} - a \right), \quad i = 2, 3, \dots, 45. \quad (17)$$

Table 1. Observational data in Example 2.

t_i	x_{t_i}	t_i	x_{t_i}	t_i	x_{t_i}	t_i	x_{t_i}
0.1784	−2.1082	2.3677	0.6997	5.0851	3.1839	7.5321	4.8684
0.3556	−1.9832	2.5070	0.8563	5.3262	3.2692	7.5946	4.9250
0.4238	−1.8649	2.8355	1.0740	5.4179	3.3669	7.8815	5.0983
0.6345	−1.5715	2.9297	1.2400	5.7285	3.7531	8.1775	5.4798
1.0840	−1.0478	3.2364	1.4703	6.0310	3.9168	8.3659	5.6311
1.2637	−0.8264	3.5888	1.9034	6.2205	4.0313	8.6005	5.9770
1.6510	−0.1571	3.8625	2.0181	6.6052	4.1734	8.9839	6.5978
1.9710	0.3000	3.9543	2.1535	6.7823	4.3493	9.3721	6.7845
2.0955	0.4624	4.3282	2.4031	6.8963	4.4408	9.4526	6.8835
2.1623	0.4881	4.3976	2.4375	6.9633	4.5117		
2.2616	0.5946	4.7149	2.9108	7.0561	4.5659		
2.3172	0.6549	5.0074	3.1338	7.2704	4.7023		

Then, it follows from (10) that the least squares estimation of (a, b) solves

$$\min_{a,b} \sum_{i=2}^{45} \left(\frac{1}{1 + \exp\left(\frac{-\pi}{\sqrt{3}b} \left(\frac{x_{t_i} - x_{t_{i-1}}}{t_i - t_{i-1}} - a\right)\right)} - \frac{1}{44} \sum_{j=2}^{45} \mathbb{I}\left(\frac{x_{t_j} - x_{t_{j-1}}}{t_j - t_{j-1}} \leq \frac{x_{t_i} - x_{t_{i-1}}}{t_i - t_{i-1}}\right) \right)^2.$$

By using MATLAB to solve the above minimization problem, the optimal solution can be obtained as

$$\hat{a} = 0.9406, \hat{b} = 0.5105.$$

Thus, the estimated linear uncertain differential equation is

$$dX_t = 0.9406dt + 0.5105dC_t. \quad (18)$$

By substituting the estimated $\hat{a}$ and $\hat{b}$ into (17), we can obtain 44 samples $h_2, h_3, \dots, h_{45}$ of the normal uncertainty distribution $\mathcal{N}(0, 1)$, as shown in Table 2 and Figure 1.

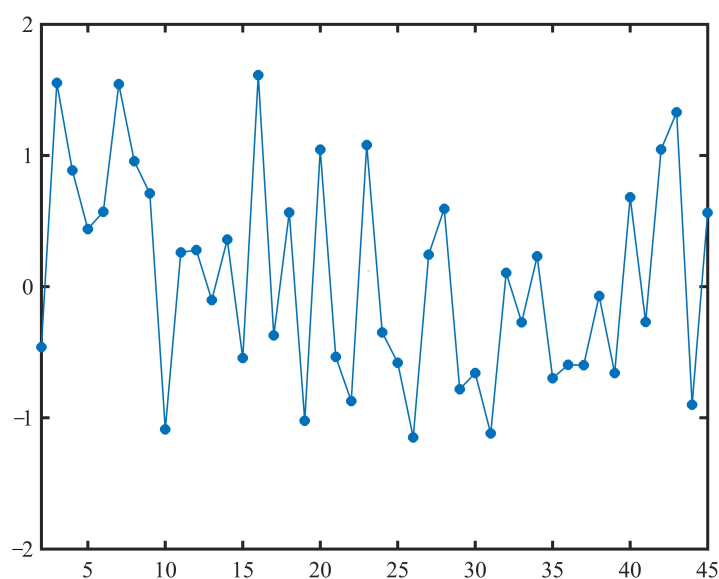
**Figure 1.** The samples of normal uncertainty distribution $\mathcal{N}(0, 1)$ obtained in Example 2.

Table 2. The samples of normal uncertainty distribution $\mathcal{N}(0, 1)$ obtained in Example 2.

−0.4610	1.5532	0.8862	0.4394	0.5711	1.5431	0.9559	0.7106	−1.0881
0.2606	0.2793	−0.1019	0.3590	−0.5446	1.6123	−0.3716	0.5653	−1.0215
1.0445	−0.5350	−0.8720	1.0799	−0.3492	−0.5799	−1.1494	0.2444	0.5937
−0.7826	−0.6585	−1.1193	0.1048	−0.2721	0.2308	−0.6980	−0.5956	−0.5991
−0.0716	−0.6589	0.6814	−0.2688	1.0459	1.3292	−0.9000	0.5638	

If the estimated parameters $\hat{a}$ and $\hat{b}$ are appropriate, then the obtained $h_2, h_3, \dots, h_{45}$ should indeed be a set of samples of normal uncertainty distribution $\mathcal{N}(0, 1)$. That is, the test of the estimated parameters $\hat{a}$ and $\hat{b}$ is equivalent to testing whether $h_2, h_3, \dots, h_{45}$ is a set of samples of $\mathcal{N}(0, 1)$, whose null hypothesis is

$$e = 0 \text{ and } \sigma = 1$$

and the alternative hypothesis is

$$e \neq 0 \text{ or } \sigma \neq 1,$$

where e is the expected value, and σ is the standard deviation of the normal population distribution of $h_2, h_3, \dots, h_{45}$, respectively. The uncertain hypothesis test presented by Ye and Liu [28] tests whether the above hypothesis test problem for a given significance level $\alpha = 0.05$ is

$$W = \{(h_2, h_3, \dots, h_{45}) : \text{there are at least 3 of indexes } i\text{'s with } 2 \leq i \leq 45 \\ \text{such that } h_i < \Phi^{-1}\left(\frac{\alpha}{2}\right) \text{ or } h_i > \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\},$$

where

$$\Phi^{-1}(\alpha) = \frac{\sqrt{3}}{\pi} \ln \frac{\alpha}{1 - \alpha}$$

is the inverse uncertainty distribution of $\mathcal{N}(0, 1)$. As shown in Table 2 and Figure 1, it is easy to see that all the points in $h_2, h_3, \dots, h_{45}$ are within the interval

$$\left[\Phi^{-1}\left(\frac{\alpha}{2}\right), \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\right] = [-2.0198, 2.0198].$$

Thus, we obtain

$$(h_2, h_3, \dots, h_{45}) \notin W.$$

That is, the obtained $h_2, h_3, \dots, h_{45}$ is indeed a set of samples of normal uncertainty distribution $\mathcal{N}(0, 1)$. Therefore, the estimated parameters $\hat{a}$ and $\hat{b}$ are appropriate.

Example 3. Consider the following exponential uncertain differential equation,

$$dX_t = \mu_t X_t dt + \sigma_t X_t dC_t, \quad (19)$$

where μ_t and σ_t are two time-varying parameters that need to be estimated. Suppose also there is a dataset consisting of 120 sets of observational data (t_i, x_{t_i}) , $i = 1, 2, \dots, 120$, as shown in Table 3. In order to use exponential uncertain differential Equation (19) to describe the observational data in Table 3, we first need to estimate the time-varying unknown parameters μ_t and σ_t .

Table 3. Observational data in Example 3.

t_i	x_{t_i}	t_i	x_{t_i}	t_i	x_{t_i}	t_i	x_{t_i}
0.0320	0.0931	1.5224	0.2614	2.9707	0.6073	4.1176	0.9295
0.1395	0.1009	1.6224	0.2860	2.9867	0.6159	4.1716	0.9777
0.2116	0.1065	1.6549	0.2934	3.0080	0.6255	4.1913	0.9738
0.2964	0.1138	1.7040	0.3024	3.0259	0.6328	4.2973	1.0292

Table 3. Cont.

t_i	x_{t_i}	t_i	x_{t_i}	t_i	x_{t_i}	t_i	x_{t_i}
0.3219	0.1161	1.7261	0.3041	3.0373	0.6349	4.3085	1.0315
0.3508	0.1186	1.7475	0.3070	3.0657	0.6370	4.3666	1.0360
0.4217	0.1250	1.8655	0.3164	3.0873	0.6436	4.3790	1.0463
0.4613	0.1286	1.9125	0.3307	3.1144	0.6444	4.4365	1.0874
0.5353	0.1363	1.9575	0.3351	3.1979	0.6425	4.5064	1.0467
0.5706	0.1395	1.9859	0.3382	3.2532	0.6420	4.6758	1.0420
0.5836	0.1408	2.0629	0.3659	3.2757	0.6420	4.7636	1.0228
0.6150	0.1442	2.1800	0.3883	3.3067	0.6474	4.7878	1.0249
0.6299	0.1459	2.2062	0.3911	3.3485	0.6789	4.8038	1.0257
0.6435	0.1476	2.2639	0.3979	3.3681	0.6813	4.8409	1.0516
0.7430	0.1592	2.2962	0.4014	3.4395	0.7016	4.8663	1.0582
0.7887	0.1639	2.3420	0.4084	3.5105	0.7264	4.9128	1.0196
0.8431	0.1708	2.3628	0.4156	3.5221	0.7312	4.9232	1.0250
0.9015	0.1770	2.3747	0.4210	3.5542	0.7435	4.9655	1.0038
0.9439	0.1813	2.4530	0.4435	3.5891	0.7771	4.9887	1.0294
1.0210	0.1890	2.5139	0.4724	3.6518	0.8165	5.0134	1.0285
1.0832	0.1965	2.5731	0.4771	3.7019	0.8148	5.0395	1.0602
1.1297	0.2020	2.6075	0.4841	3.7341	0.8423	5.0880	1.0946
1.1755	0.2058	2.6375	0.4914	3.7742	0.8624	5.1296	1.1411
1.1875	0.2081	2.6626	0.4933	3.8050	0.8634	5.1513	1.1391
1.2179	0.2117	2.7197	0.5040	3.8839	0.8838	5.1937	1.1208
1.2557	0.2167	2.7400	0.5117	3.9350	0.8918	5.2511	1.1385
1.3112	0.2255	2.8296	0.5549	3.9700	0.9335	5.3077	1.0958
1.3560	0.2319	2.8422	0.5544	3.9941	0.9424	5.3187	1.1050
1.3793	0.2358	2.8926	0.5663	4.0110	0.9417	5.4071	1.1002
1.4551	0.2468	2.9041	0.5706	4.0568	0.9273	5.4329	1.1186

Take $k = 20$ as the sliding window for estimating μ_t and σ_t . Then, for each index j with $2 \leq j \leq 101$, we choose (t_i, x_{t_i}) , $i = j - 1, j, \dots, j + 19$ as the observational data for estimating unknown parameters $\mu_{t_{j-1}}$ and $\sigma_{t_{j-1}}$. By substituting (t_i, x_{t_i}) , $i = j - 1, j, \dots, j + 19$ into (4), we can obtain

$$h_i(\mu_{t_{j-1}}, \sigma_{t_{j-1}}) = \frac{x_{t_i} - x_{t_{i-1}} - \mu_{t_{j-1}} x_{t_{i-1}} (t_i - t_{i-1})}{\sigma_{t_{j-1}} x_{t_{i-1}} (t_i - t_{i-1})} = \frac{1}{\sigma_{t_{j-1}}} \left(\frac{x_{t_i} - x_{t_{i-1}}}{x_{t_{i-1}} (t_i - t_{i-1})} - \mu_{t_{j-1}} \right) \quad (20)$$

with $i = j - 1, j, \dots, j + 19$. Then, it follows from (14) that the least squares estimation of $(\mu_{t_{j-1}}, \sigma_{t_{j-1}})$ solves

$$\min_{\mu_{t_{j-1}}, \sigma_{t_{j-1}}} \sum_{i=j}^{j+19} \left(\left(1 + \exp \left(\frac{-\pi}{\sqrt{3} \sigma_{t_{j-1}}} \left(\frac{x_{t_i} - x_{t_{i-1}}}{x_{t_{i-1}} (t_i - t_{i-1})} - \mu_{t_{j-1}} \right) \right) \right)^{-1} \right. \\ \left. - \frac{1}{20} \sum_{l=j}^{j+19} \mathbb{I} \left(\frac{x_{t_l} - x_{t_{l-1}}}{x_{t_{l-1}} (t_l - t_{l-1})} \leq \frac{x_{t_i} - x_{t_{i-1}}}{x_{t_{i-1}} (t_i - t_{i-1})} \right) \right)^2.$$

By using MATLAB to solve the above minimization problem, we can obtain the least squares estimation $(\hat{\mu}_{t_{j-1}}, \hat{\sigma}_{t_{j-1}})$. Iterating over j from 2 to 101, and repeating the above operation, we can obtain a set of least squares estimations $(\hat{\mu}_{t_{j-1}}, \hat{\sigma}_{t_{j-1}})$, $j = 2, 3, \dots, 101$, as shown in Table 4 and Figures 2 and 3.

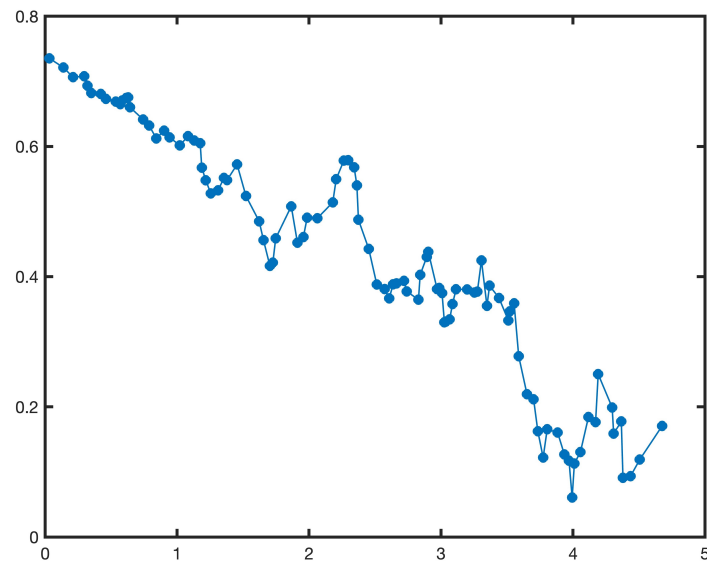


Figure 2. The estimated $\hat{\mu}_{t_j}$ with $t_j, j = 1, 2, \dots, 100$ in Example 3.

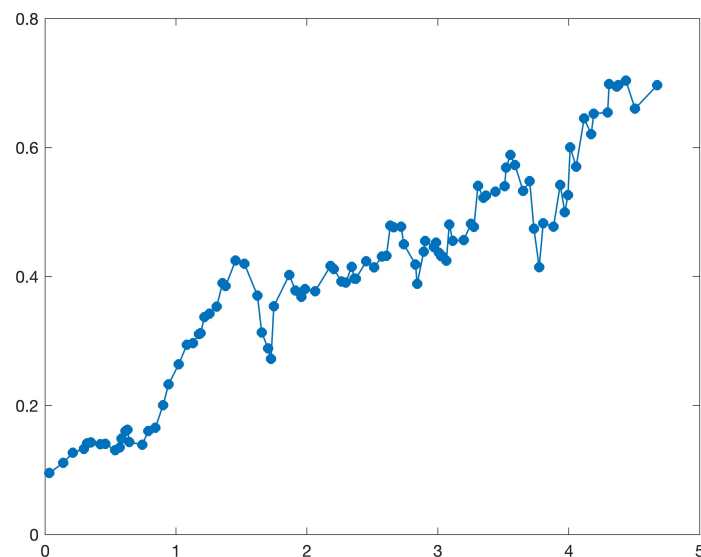


Figure 3. The estimated $\hat{\sigma}_{t_j}$ with $t_j, j = 1, 2, \dots, 100$ in Example 3.

Table 4. The estimated $\hat{\mu}_{t_j}$ and $\hat{\sigma}_{t_j}$ with $j = 1, 2, \dots, 100$ in Example 3.

μ_{t_j}	σ_{t_j}	μ_{t_j}	σ_{t_j}	μ_{t_j}	σ_{t_j}	μ_{t_j}	σ_{t_j}
0.7354	0.0954	0.6158	0.2942	0.4897	0.3772	0.3810	0.4458
0.7213	0.1111	0.6091	0.2969	0.5141	0.4165	0.3827	0.4527
0.7064	0.1270	0.6049	0.3109	0.5499	0.4118	0.3748	0.4369
0.7080	0.1327	0.5675	0.3122	0.5783	0.3924	0.3300	0.4316
0.6932	0.1413	0.5481	0.3371	0.5790	0.3908	0.3310	0.4313
0.6821	0.1427	0.5277	0.3425	0.5680	0.4152	0.3343	0.4248
0.6808	0.1401	0.5327	0.3534	0.5402	0.3968	0.3578	0.4808
0.6731	0.1406	0.5518	0.3900	0.4874	0.3967	0.3806	0.4555
0.6690	0.1308	0.5484	0.3853	0.4425	0.4238	0.3803	0.4565
0.6649	0.1350	0.5723	0.4248	0.3879	0.4142	0.3757	0.4818
0.6705	0.1488	0.5238	0.4200	0.3811	0.4311	0.3769	0.4771
0.6746	0.1605	0.4850	0.3706	0.3666	0.4323	0.4249	0.5407

Table 4. *Cont.*

μ_{t_j}	σ_{t_j}	μ_{t_j}	σ_{t_j}	μ_{t_j}	σ_{t_j}	μ_{t_j}	σ_{t_j}
0.6755	0.1626	0.4562	0.3133	0.3883	0.4791	0.3550	0.5222
0.6601	0.1433	0.4164	0.2885	0.3897	0.4765	0.3862	0.5257
0.6415	0.1393	0.4217	0.2724	0.3935	0.4773	0.3673	0.5317
0.6323	0.1606	0.4588	0.3540	0.3774	0.4503	0.3327	0.5401
0.6121	0.1656	0.5081	0.4026	0.3647	0.4188	0.3467	0.5691
0.6244	0.2007	0.4521	0.3784	0.4027	0.3888	0.3590	0.5886
0.6139	0.2330	0.4609	0.3687	0.4304	0.4387	0.2778	0.5731
0.6017	0.2639	0.4905	0.3807	0.4380	0.4553	0.2195	0.5328

Next, we should choose appropriate regression models to fit the estimated $\hat{\mu}_{t_j}$ and $\hat{\sigma}_{t_j}$ with $j = 1, 2, \dots, 100$. It is easy to see from Figures 2 and 3 that we can choose the linear regression model to fit them. At first, we assume that the estimated $\hat{\mu}_{t_j}$, $j = 1, 2, \dots, 100$ can be fitted by

$$\mu_t = \beta_1 + \beta_2 t.$$

Then, β_1 and β_2 can be obtained by solving

$$\min_{\beta_1, \beta_2} \sum_{j=1}^{100} \left(\hat{\mu}_{t_j} - \beta_1 - \beta_2 t_j \right)^2,$$

which are

$$\beta_2 = \frac{\sum_{j=1}^{100} \left(\hat{\mu}_{t_j} - \frac{1}{100} \sum_{j=1}^{100} \hat{\mu}_{t_j} \right) \left(t_j - \frac{1}{100} \sum_{j=1}^{100} t_j \right)}{\sum_{j=1}^{100} \left(t_j - \frac{1}{100} \sum_{j=1}^{100} t_j \right)^2} = -0.1337$$

and

$$\beta_1 = \frac{1}{100} \sum_{j=1}^{100} \hat{\mu}_{t_j} - \beta_2 \frac{1}{100} \sum_{j=1}^{100} t_j = 0.7484.$$

That is, the time-varying parameter μ_t can be fitted by

$$\hat{\mu}_t = 0.7484 - 0.1337t.$$

Similarly, we can also assume that the estimated $\hat{\sigma}_{t_j}$, $j = 1, 2, \dots, 100$ can be fitted by

$$\sigma_t = \beta_3 + \beta_4 t,$$

and β_3 and β_4 can be obtained as

$$\beta_3 = \frac{\sum_{j=1}^{100} \left(\hat{\sigma}_{t_j} - \frac{1}{100} \sum_{j=1}^{100} \hat{\sigma}_{t_j} \right) \left(t_j - \frac{1}{100} \sum_{j=1}^{100} t_j \right)}{\sum_{j=1}^{100} \left(t_j - \frac{1}{100} \sum_{j=1}^{100} t_j \right)^2} = 0.1170$$

and

$$\beta_4 = \frac{1}{100} \sum_{j=1}^{100} \hat{\sigma}_{t_j} - \beta_3 \frac{1}{100} \sum_{j=1}^{100} t_j = 0.1230.$$

Thus, the time-varying parameter σ_t can be fitted by

$$\hat{\sigma}_t = 0.1230 + 0.1170t,$$

and the estimated exponential uncertain differential equation is

$$dX_t = (0.7484 - 0.1337t)X_t dt + (0.1230 + 0.1170t)X_t dC_t \quad (21)$$

immediately.

By substituting the estimated $\hat{\mu}_t$ and $\hat{\sigma}_t$ into (20), we can obtain 119 samples $h_2, h_3, \dots, h_{120}$ of the normal uncertainty distribution $\mathcal{N}(0, 1)$, as shown in Table 2 and Figure 1. In order to test whether the estimated time-varying parameters $\hat{\mu}_t$ and $\hat{\sigma}_t$ are appropriate, it follows from the analysis of Example 2 that the test should be

$$W = \{(h_2, h_3, \dots, h_{120}) : \text{there are at least 6 of indexes } i\text{'s with } 2 \leq i \leq 120 \\ \text{such that } h_i < \Phi^{-1}\left(\frac{\alpha}{2}\right) \text{ or } h_i > \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\},$$

for a given significance level $\alpha = 0.05$, where

$$\Phi^{-1}(\alpha) = \frac{\sqrt{3}}{\pi} \ln \frac{\alpha}{1 - \alpha}$$

is the inverse uncertainty distribution of $\mathcal{N}(0, 1)$. As shown in Table 5 and Figure 4, it is easy to see that all the points in $h_2, h_3, \dots, h_{120}$ are within the interval

$$\left[\Phi^{-1}\left(\frac{\alpha}{2}\right), \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\right] = [-2.0198, 2.0198].$$

Thus, we obtain

$$(h_2, h_3, \dots, h_{120}) \notin W.$$

That is, the obtained time-varying parameters $\hat{\mu}_t$ and $\hat{\sigma}_t$ are indeed a set of samples of normal uncertainty distribution $\mathcal{N}(0, 1)$. Therefore, the estimated time-varying parameters $\hat{\mu}_t$ and $\hat{\sigma}_t$ are appropriate.

Table 5. The samples of normal uncertainty distribution $\mathcal{N}(0, 1)$ obtained in Example 3.

0.2887	0.2486	0.6226	0.4367	0.2968	0.3572	0.1748	0.7016	−0.0654
0.2730	0.5156	0.5907	0.9889	0.6417	−0.0053	0.6515	−0.1051	−0.2377
−0.3043	0.1318	−0.0391	−0.7073	1.2839	−0.0764	0.1271	0.5802	0.2210
0.5554	0.1718	1.1114	1.3088	0.8498	0.3145	−0.8332	−0.2165	−0.7801
1.3694	−0.5808	−0.4625	1.6414	0.1351	−0.4722	−0.3969	−0.4480	−0.1641
1.0534	1.6631	0.6235	1.5862	−0.5863	0.0538	0.2453	−0.5593	−0.0336
0.8264	1.2629	−0.9426	0.1213	0.6404	1.3149	1.1298	0.8047	0.6414
−0.1174	−0.4655	0.2891	−0.6024	−0.7542	−0.6718	−0.6198	−0.0749	1.6776
−0.2347	0.2306	0.3992	0.5437	0.4653	1.8935	0.9971	−0.5501	1.4302
0.6127	−0.3598	0.1038	−0.0875	1.9075	0.3020	−0.4414	−0.9219	−0.2796
1.2611	−0.6454	0.5686	0.0511	−0.1551	0.9998	0.8180	−1.0748	−0.2652
−0.4971	−0.0400	−0.0842	0.8385	0.2111	−1.2752	0.5928	−0.8269	1.4410
−0.1673	1.5591	0.8324	1.3284	−0.1970	−0.6050	0.3042	−0.9614	0.9620
−0.1154	0.8258							

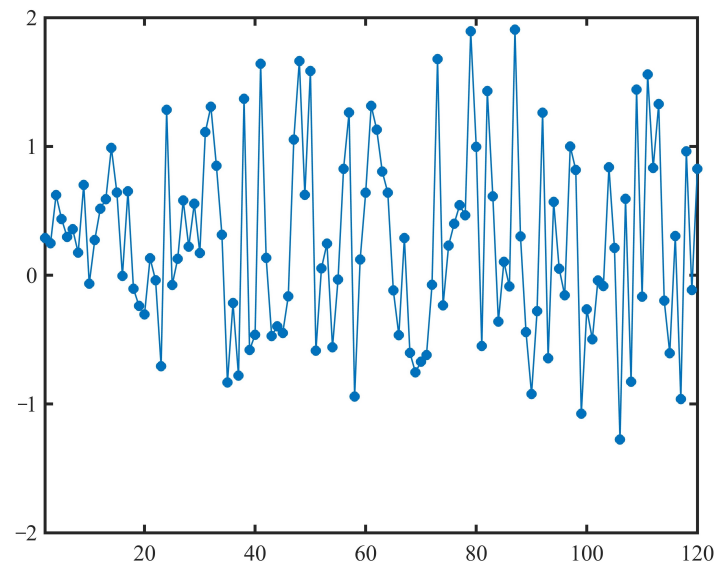


Figure 4. The samples of normal uncertainty distribution $\mathcal{N}(0, 1)$ obtained in Example 3.

8. The Application in Economic Field

In this section, we will apply the method proposed in this paper to model stock prices and compare it with the method based on stochastic differential equations.

8.1. Data Source

Anhui Gujinggong Liquor Co., Ltd. (Bozhou, China) is the core enterprise of Gujing Group and is the first liquor enterprise in China to issue both A shares and B shares. The stock prices of this company can to some extent reflect the fluctuations of the Chinese liquor market. In order to model the stock prices of Anhui Gujinggong Liquor Co., Ltd. (stock code: 000596.SZ), listed on the Shenzhen Stock Exchange, we collected the closing prices of the company's stocks for all trading days from 1 January 2024 to 21 June 2024, with the currency unit being RMB. The data were sourced from the official historical trading records of Anhui Gujinggong Liquor Co., Ltd. provided by Sina Finance, which can be accessed via the webpage link (<https://finance.sina.com.cn/realstock/company/sz000596/nc.shtml>, accessed on 12 September 2025). The data were obtained on 12 September 2025, and a total of 112 valid data samples were obtained. The specific details of the closing prices can be seen in Table 6, while the trend of the closing prices can be seen in Figure 5.

Table 6. The closing prices of Anhui Gujinggong Liquor Co., Ltd. from 1 January 2024 to 21 June 2024, accessed by Sina Finance <https://finance.sina.com.cn/realstock/company/sz000596/nc.shtml> (accessed on 12 September 2025).

224.34	217.28	211.55	212.77	213.75	211.34	211.21	212.89	214.76
212.43	213.98	208.2	210.97	214.14	207.02	207.2	207.83	212.5
211.74	210.3	200.79	199.12	200.06	204.65	210.22	219.6	224.5
220	221.33	218.11	226.56	224.88	224.36	222	219.5	216.65
219.13	221.51	224.26	231.1	229.1	231.99	228.94	234.7	242.53
240.19	241.75	240.87	247	256.89	257.19	266.45	261.5	257.92
265.6	262.6	262.44	260	269	266.8	267	253.76	256.64
255.15	257.13	252.18	262.65	253.76	255.19	254.85	251.78	259.1
262.39	261.99	258.47	264.5	266.06	271.24	279.12	280.38	277.48
277.88	275.32	265	261.8	263.13	263.68	267.15	264.3	262.47
259.24	255.67	249.58	250.6	247.77	246.18	244.54	250.7	253.65
256.9	251.45	247.56	243.44	236.2	236.11	234.17	233.34	237
230.68	231.59	225.21	220.64					

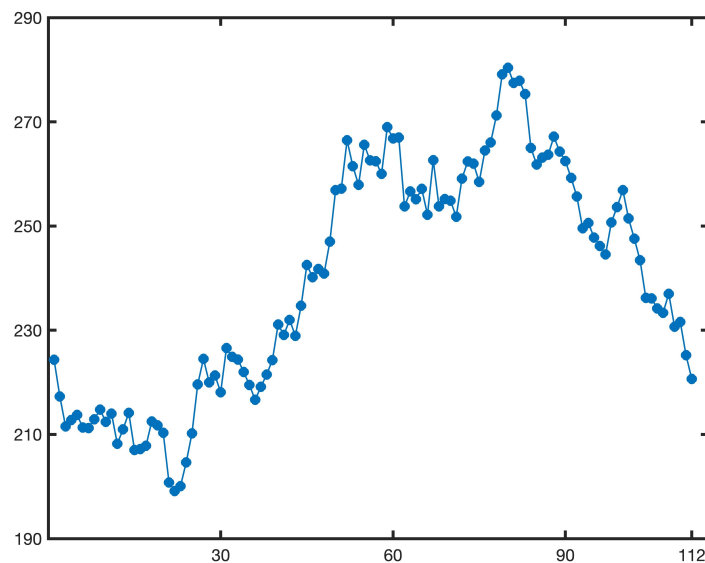


Figure 5. The closing prices of Anhui Gujinggong Liquor Co., Ltd. from 1 January 2024 to 21 June 2024 accessed by Sina Finance <https://finance.sina.com.cn/realstock/company/sz000596/nc.shtml> (accessed on 12 September 2025).

8.2. Uncertain Stock Prices Model

Let $i = 1, 2, \dots, 112$ denote the trading days from 1 January 2024 to 21 June 2024, and let $x_i, i = 1, 2, \dots, 112$ denote the closing prices of Anhui Gujinggong Liquor Co., Ltd., respectively. Then, we choose the following uncertain mean reversion equation,

$$dX_t = (m - aX_t)dt + \sigma dC_t,$$

to model the closing prices $x_i, i = 1, 2, \dots, 112$.

Next, we need to estimate the unknown parameters m, a , and σ . By substituting the observational data $(i, x_i), i = 1, 2, \dots, 112$ into (4), we can obtain

$$h_i(m, a, \sigma) = \frac{x_i - x_{i-1} - m + ax_{i-1}}{\sigma} = \frac{1}{\sigma}(x_i - x_{i-1} - m + ax_{i-1}) \quad (22)$$

for $i = 2, 3, \dots, 112$. Then, it follows from (10) that the least squares estimation of (m, a, σ) solves

$$\min_{m, a, \sigma} \sum_{i=2}^{112} \left(\left(1 + \exp \left(\frac{-\pi}{\sqrt{3}\sigma} (x_i - x_{i-1} - m + ax_{i-1}) \right) \right)^{-1} - \frac{1}{111} \sum_{j=2}^{112} \mathbb{I}(x_j - x_{j-1} + ax_{j-1} \leq x_i - x_{i-1} + ax_{i-1}) \right)^2.$$

By using MATLAB to solve the above minimization problem, the optimal solution can be obtained as

$$\hat{m} = 23.2099, \hat{a} = 0.0980, \hat{\sigma} = 5.0665.$$

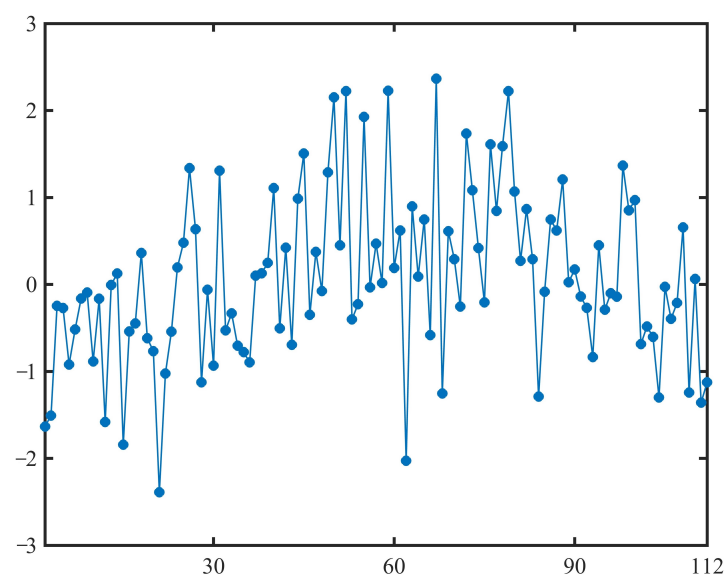
Thus, the estimated uncertain stock prices model is

$$dX_t = (23.2099 - 0.0980X_t)dt + 5.0665dC_t. \quad (23)$$

By substituting the estimated $\hat{m}$, $\hat{a}$, and $\hat{\sigma}$ into (22), we can obtain 111 samples $h_2, h_3, \dots, h_{112}$ of the normal uncertainty distribution $\mathcal{N}(0, 1)$, as shown in Table 7 and Figure 6.

Table 7. The samples of normal uncertainty distribution $\mathcal{N}(0, 1)$ obtained in uncertain stock prices model (23).

−1.6333	−1.5074	−0.2465	−0.2703	−0.9204	−0.5170	−0.1623	−0.0923	−0.8851
−0.1644	−1.5811	−0.0054	0.1271	−1.8425	−0.5395	−0.4472	0.3624	−0.6189
−0.7679	−2.3886	−1.0252	−0.5423	0.1963	0.4785	1.3383	0.6356	−1.1249
−0.0613	−0.9336	1.3074	−0.5285	−0.3320	−0.7052	−0.7785	−0.8960	0.1009
0.1291	0.2482	1.1087	−0.5038	0.4227	−0.6938	0.9861	1.5061	−0.3497
0.3748	−0.0766	1.2900	2.1507	0.4493	2.2236	−0.4020	−0.2273	1.9258
−0.0335	0.4690	0.0159	2.2266	0.1902	0.6213	−2.0276	0.8979	0.0911
0.7472	−0.5823	2.3654	−1.2532	0.6117	0.2901	−0.2554	1.7360	1.0822
0.4175	−0.2060	1.6108	0.8452	1.5899	2.2231	1.0689	0.2722	0.8674
0.2910	−1.2902	−0.0846	0.7476	0.6194	1.2063	0.0261	0.1723	−0.1395
−0.2691	−0.8356	0.4499	−0.2902	−0.1003	−0.1409	1.3669	0.8525	0.9688
−0.6854	−0.4830	−0.6037	−1.2992	−0.0281	−0.3950	−0.2134	0.6567	−1.2423
0.0625	−1.3588	−1.1250						

**Figure 6.** The samples of normal uncertainty distribution $\mathcal{N}(0, 1)$ obtained in uncertain stock prices model (23).

If the estimated parameters $\hat{m}$, $\hat{a}$, and $\hat{\sigma}$ are appropriate, then the obtained $h_2, h_3, \dots, h_{112}$ should indeed be a set of samples of normal uncertainty distribution $\mathcal{N}(0, 1)$. That is, the test of the estimated parameters $\hat{m}$, $\hat{a}$, and $\hat{\sigma}$ is equivalent to testing whether $h_2, h_3, \dots, h_{112}$ is a set of samples of $\mathcal{N}(0, 1)$, whose test for a given significance level $\alpha = 0.01$ is

$$W = \{(h_2, h_3, \dots, h_{112}) : \text{there are at least 2 of indexes } i\text{'s with } 2 \leq i \leq 112 \\ \text{such that } h_i < \Phi^{-1}\left(\frac{\alpha}{2}\right) \text{ or } h_i > \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\},$$

where

$$\Phi^{-1}(\alpha) = \frac{\sqrt{3}}{\pi} \ln \frac{\alpha}{1 - \alpha}$$

is the inverse uncertainty distribution of $\mathcal{N}(0, 1)$. As shown in Table 7 and Figure 6, it is easy to see that all the points in $h_2, h_3, \dots, h_{112}$ are within the interval

$$\left[\Phi^{-1}\left(\frac{\alpha}{2}\right), \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\right] = [-2.9184, 2.9184].$$

Thus, we obtain

$$(h_2, h_3, \dots, h_{112}) \notin W.$$

That is, the obtained $h_2, h_3, \dots, h_{112}$ is indeed a set of samples of normal uncertainty distribution $\mathcal{N}(0, 1)$. Therefore, the estimated parameters $\hat{m}$, $\hat{a}$, and $\hat{\sigma}$ are appropriate.

8.3. Stochastic Stock Prices Model

In this subsection, we retain the closing prices of Anhui Gujinggong Liquor Co., Ltd. from 1 January 2024 to 21 June 2024 shown in Table 6 and Figure 5, as described in the previous subsection, and choose the following stochastic mean reversion equation:

$$dX_t = (m - aX_t)dt + \sigma dW_t,$$

to model the closing prices x_i , $i = 1, 2, \dots, 112$, where W_t is a Wiener process.

Before modeling the closing prices, we first derive the statistical invariant based on the above stochastic mean reversion equation. By writing the above stochastic mean reversion equation into the Euler difference form, we obtain

$$X_i = X_{i-1} + (m - aX_{i-1}) + \sigma(W_i - W_{i-1}), \quad i = 2, 3, \dots, 112.$$

Based on the above Euler difference form, we can obtain

$$\frac{X_i - X_{i-1} - (m - aX_{i-1})}{\sigma} = W_i - W_{i-1}, \quad i = 2, 3, \dots, 112,$$

which are independent and identically distributed standard normal random variables according to the definition of the Wiener process and its property of independent increments. That is, we have

$$\frac{X_i - X_{i-1} - (m - aX_{i-1})}{\sigma} = W_i - W_{i-1} \sim N(0, 1).$$

We substitute the closing prices x_i , $i = 1, 2, \dots, 112$ into $\frac{X_i - X_{i-1} - (m - aX_{i-1})}{\sigma}$ and write

$$h_i(m, a, \sigma) = \frac{x_i - x_{i-1} - m + ax_{i-1}}{\sigma}, \quad i = 2, 3, \dots, 112, \quad (24)$$

which are a set of samples of the standard normal random distribution function $N(0, 1)$.

Similar to the least squares estimation based on the uncertainty distribution, we can also derive that the least squares estimation of (m, a, σ) based on the random distribution function is the solution to the optimization problem,

$$\min_{m, a, \sigma} \sum_{i=2}^{112} \left(\Psi \left(\frac{x_i - x_{i-1} - m + ax_{i-1}}{\sigma} \right) - \frac{1}{111} \sum_{j=2}^{112} \mathbb{I}(x_j - x_{j-1} + ax_{j-1} \leq x_i - x_{i-1} + ax_{i-1}) \right)^2,$$

where Ψ is the random distribution function of the standard normal random distribution function $N(0, 1)$. By using MATLAB to solve the above minimization problem, the optimal solution can be obtained as

$$\hat{m} = 36.3820, \hat{a} = 0.1529, \hat{\sigma} = 5.4680.$$

Thus, the estimated stochastic stock prices model is

$$dX_t = (36.3820 - 0.1529X_t)dt + 5.4680dW_t. \quad (25)$$

By substituting the estimated $\hat{m}$, $\hat{a}$, and $\hat{\sigma}$ into (24), we can obtain 111 samples $h_2, h_3, \dots, h_{112}$ of the standard normal random distribution function $N(0,1)$, as shown in Table 8 and Figure 7.

Table 8. The samples of the standard normal random distribution function $N(0,1)$ obtained in stochastic stock prices model (25).

−1.6729	−1.6271	−0.5162	−0.5260	−1.1186	−0.7690	−0.4416	−0.3599	−1.0757
−0.4313	−1.7285	−0.3264	−0.1758	−1.9691	−0.8331	−0.7457	0.0107	−0.8518
−0.9974	−2.5135	−1.3456	−0.9149	−0.2211	0.0864	0.9389	0.3819	−1.2003
−0.2599	−1.0548	0.9894	−0.6269	−0.4618	−0.8128	−0.9044	−1.0383	−0.1432
−0.0922	0.0420	0.8669	−0.5585	0.2798	−0.7257	0.8003	1.3399	−0.3012
0.3467	−0.0560	1.2014	2.0605	0.5831	2.2301	−0.1098	0.0024	1.9616
0.2231	0.6586	0.2372	2.2611	0.4645	0.8419	−1.6105	0.9674	0.2488
0.8417	−0.3703	2.3114	−0.9366	0.7023	0.4185	−0.0902	1.7241	1.1917
0.6089	0.0271	1.6752	1.0263	1.7319	2.3706	1.3802	0.6546	1.1770
0.6469	−0.8439	0.1697	0.9087	0.8033	1.3527	0.2939	0.4007	0.0935
−0.0590	−0.6196	0.5104	−0.1652	−0.0175	−0.0711	1.3095	0.8947	1.0320
−0.4682	−0.3353	−0.4861	−1.1719	−0.0667	−0.4075	−0.2587	0.5392	−1.1837
−0.0381	−1.3459	−1.1932						

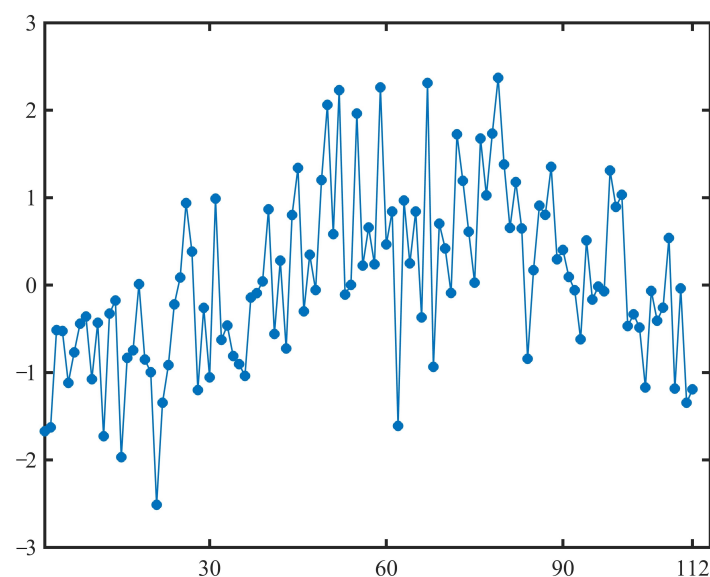


Figure 7. The samples of the standard normal random distribution function $N(0,1)$ obtained in stochastic stock prices model (25).

As shown in Table 8 and Figure 7, if we select two non-overlapping parts from the 111 samples of the standard normal random distribution function $N(0,1)$, i.e.,

$$h_2, h_3, \dots, h_{30}$$

and

$$h_{62}, h_{63}, \dots, h_{90},$$

and use the two-sample Kolmogorov–Smirnov test to test them via the function “kstest2” in MATLAB, then the p value is 1.4798×10^{-7} , which shows that the above two parts do not come from the same population. Thus, the samples $h_2, h_3, \dots, h_{112}$ are not white noise in the sense of probability theory, nor do they follow the standard normal random distribution function $N(0,1)$. Thus, the least squares estimation based on the uncertainty distribution

for uncertain differential equations is more appropriate than the least squares estimation based on the random distribution function for stochastic differential equations.

9. Conclusions

For the purpose of overcoming the disadvantage that the current parameter estimation study of uncertain differential equations is strictly dependent on the solvability of residuals and cannot be applied to general uncertain differential equations, a symmetrical statistical invariant based on the difference scheme of uncertain differential equations was constructed. Further, based on the statistical invariant and the principle of least squares, the least squares estimation for the fixed constant parameters and time-varying parameters of the uncertain differential equations were investigated, and the corresponding numerical solutions were also designed to facilitate the acquisition of the numerical solution of the above least squares estimations. Finally, two numerical examples and an empirical study were provided to illustrate the effects of the methods proposed in this paper.

Author Contributions: Conceptualization, X.X.; methodology, X.X. and Y.D.; software, X.X.; validation, X.X. and Y.D.; formal analysis, Y.D.; investigation, X.X.; resources, X.X. and Y.D.; data curation, X.X.; writing—original draft preparation, X.X.; writing—review and editing, Y.D.; supervision, X.X. and Y.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The original contributions presented in this study are included in the article, and the research data mentioned in this article have been presented in Tables 1, 3 and 6. Further inquiries can be directed to the corresponding author.

Acknowledgments: The authors especially thank the editors and anonymous referees for their kindly review and helpful comments. Any remaining errors are ours.

Conflicts of Interest: We declare that we have no relevant or material financial interests that relate to the research described in this paper. Neither the entire paper nor any part of its content has been published or accepted elsewhere. It is also not being submitted to any other journal.

References

1. Kolmogorov, A. *Grundbegriffe der Wahrscheinlichkeitsrechnung*; Springer: Berlin/Heidelberg, Germany, 1933.
2. Liu, B. *Uncertainty Theory*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2007.
3. Yang, X.; Ke, H. Uncertain interest rate model for Shanghai interbank offered rate and pricing of American swaption. *Fuzzy Optim. Decis. Mak.* **2023**, *22*, 447–462. [\[CrossRef\]](#)
4. Yang, L.; Liu, Y. Solution method and parameter estimation of uncertain partial differential equation with application to China's population. *Fuzzy Optim. Decis. Mak.* **2024**, *23*, 155–177. [\[CrossRef\]](#)
5. Liu, Y.; Qin, Z.; Li, X. Are the queueing systems in practice random or uncertain? Evidence from online car-hailing data in Beijing. *Fuzzy Optim. Decis. Mak.* **2024**, *23*, 497–511. [\[CrossRef\]](#)
6. Liu, B. *Uncertainty Theory*, 5th ed.; Springer: Berlin/Heidelberg, Germany, 2025. Available online: <https://cloud.tsinghua.edu.cn/d/df71e9ec330e49e59c9c/files/?p=%2FUncertainty.Theory.Book.pdf> (accessed on 19 September 2025).
7. Yao, K. *Uncertain Differential Equations*; Springer: Berlin/Heidelberg, Germany, 2016.
8. Zhu, Y. *Uncertain Optimal Control*; Springer: Singapore, 2019.
9. Yao, K. *Uncertain Renewal Processes*; Springer: Singapore, 2019.
10. Zhang, B.; Peng, J. *Uncertain Graph and Network Optimization*; Springer: Singapore, 2022.
11. Liu, B. Fuzzy process, hybrid process and uncertain process. *J. Uncertain Syst.* **2008**, *2*, 3–16.
12. Chen, X.; Liu, B. Existence and uniqueness theorem for uncertain differential equations. *Fuzzy Optim. Decis. Mak.* **2010**, *9*, 69–81. [\[CrossRef\]](#)
13. Liu, B. Some research problems in uncertainty theory. *J. Uncertain Syst.* **2009**, *3*, 3–10.
14. Yao, K.; Gao, J.; Gao, Y. Some stability theorems of uncertain differential equation. *Fuzzy Optim. Decis. Mak.* **2013**, *12*, 3–13. [\[CrossRef\]](#)

15. Yao, K.; Chen, X. A numerical method for solving uncertain differential equations. *J. Intell. Fuzzy Syst.* **2013**, *25*, 825–832. [[CrossRef](#)]
16. Yao, K.; Liu, B. Parameter estimation in uncertain differential equations. *Fuzzy Optim. Decis. Mak.* **2020**, *19*, 1–12. [[CrossRef](#)]
17. Liu, Z. Generalized moment estimation for uncertain differential equations. *Appl. Math. Comput.* **2021**, *392*, 125724. [[CrossRef](#)]
18. Sheng, Y.; Yao, K.; Chen, X. Least squares estimation in uncertain differential equations. *IEEE Trans. Fuzzy Syst.* **2020**, *28*, 2651–2655. [[CrossRef](#)]
19. Yang, X.; Liu, Y.H.; Park, G. Parameter estimation of uncertain differential equation with application to financial market. *Chaos Solitons Fract.* **2020**, *139*, 110026. [[CrossRef](#)]
20. Liu, Y.; Liu, B. Estimating unknown parameters in uncertain differential equation by maximum likelihood estimation. *Soft Comput.* **2022**, *26*, 2773–2780. [[CrossRef](#)]
21. Liu, Y.; Liu, B. Residual analysis and parameter estimation of uncertain differential equations. *Fuzzy Optim. Decis. Mak.* **2022**, *21*, 513–530. [[CrossRef](#)]
22. Liu, Y.; Liu, B. A modified uncertain maximum likelihood estimation with applications in uncertain statistics. *Commun. Stat.—Theory Methods* **2024**, *53*, 6649–6670. [[CrossRef](#)]
23. Liu, Y.; Liu, B. Estimation of uncertainty distribution function by the principle of least squares. *Commun. Stat.—Theory Methods* **2024**, *53*, 7624–7641. [[CrossRef](#)]
24. Yao, L.; Sheng, Y. Moments estimation for multi-factor uncertain differential equations based on residuals. *Soft Comput.* **2023**, *27*, 11193–11203. [[CrossRef](#)]
25. Liu, Z.; Liu, Y.; Gao, R. Maximum likelihood estimation for multi-factor uncertain differential equations. *J. Comput. Appl. Math.* **2025**, *454*, 116205. [[CrossRef](#)]
26. Wu, N.; Liu, Y. Least squares estimation of multifactor uncertain differential equations with applications to the stock market. *Symmetry* **2024**, *16*, 904. [[CrossRef](#)]
27. Yang, L.; Liu, Y. Least squares estimation of uncertain partial differential equations. *Math. Method. Appl. Sci.* **2025**, *48*, 3482–3495. [[CrossRef](#)]
28. Ye, T.; Liu, B. Uncertain hypothesis test with application to uncertain regression analysis. *Fuzzy Optim. Decis. Mak.* **2022**, *21*, 157–174. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.