



OPEN ACCESS

EDITED BY

Shuvojit Paul,
Indian Institute of Science Education and
Research Kolkata, India

REVIEWED BY

N. Narinder,
Technical University Dresden, Germany
Prasanta Panigrahi,
Indian Institute of Science Education and
Research Kolkata, India

*CORRESPONDENCE

Houshyar Noshad,
✉ hnoshad@aut.ac.ir

RECEIVED 17 March 2024

ACCEPTED 18 June 2024

PUBLISHED 18 July 2024

CITATION

Movahed AA and Noshad H (2024), Introducing
a new approach for modeling stock market
prices using the combination of jump-
drift processes.

Front. Phys. 12:1402593.

doi: 10.3389/fphy.2024.1402593

COPYRIGHT

© 2024 Movahed and Noshad. This is an open-
access article distributed under the terms of the
[Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).
The use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in this
journal is cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

Introducing a new approach for modeling stock market prices using the combination of jump-drift processes

Ali Asghar Movahed and Houshyar Noshad*

Department of Physics and Energy Engineering, Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran

The stock price data are sampled at discrete times (e.g., hourly, daily, weekly, etc). When data are sampled at discrete times, they appear as a sequence of discontinuous jump events, even if they have been sampled from a continuous process. On the other hand, distinguishing between discontinuities due to finite sampling of the continuous stochastic process and real jump discontinuities in the sample path is often a challenging task. Such considerations, led us to the question: Can discrete data (e.g., stock price) be modeled using only jump-drift processes, regardless of whether the sampled time series originally belongs to the class of continuous processes or discontinuous processes? To answer this question, we built a stochastic dynamical equation in the general form $dy(t) = \bar{\mu}dt + \sum_{i=1}^N \xi_i dJ_i(t)$, which includes a deterministic drift term ($\bar{\mu}dt$) and a combination of stochastic terms with jumpy behaviors ($\xi_i dJ_i(t)$), and used it to model the log-price time series $y(t)$. In this article, we first introduce this equation in its simplest form, including a drift term and a stochastic term, and show that such a jump-drift equation is capable of reconstructing stock prices in Black-Scholes diffusion markets. Afterwards, we extend the equation by considering two jump processes, and show that such a drift-jump-jump equation enables us to reconstruct stock prices in jump-diffusion markets more accurately than the old jump-diffusion model. To demonstrate the practical applications of the proposed method, we analyze real-world data, including the daily stock price of two different shares and gold price data with two different time horizons (hourly and weekly). Our analysis supports the practical applicability of the methodology. It should be noted that the presented approach is expandable and can be used even in non-financial research fields.

KEYWORDS

stock prices modeling, stochastic dynamical equation, Black-Scholes model, poisson jump process, jump-diffusion model, jump-drift process

1 Introduction

The stock price is known as a highly volatile variable in a stock market. Price fluctuations, which occur randomly and frequently and sometimes include sudden jumps, increase investment risk and cause concern for investors and company owners who want to increase their capital. Therefore, researchers are propelled to study the fluctuating behavior of the market to find a way to model prices (or improve existing models) to advise investors looking for the best investments [1–5]. So far, significant

progress has been made in this field, the most important of which is stock price modeling via continuous stochastic processes and discontinuous jump processes. The “arithmetic Brownian motion” model was the first mathematical model of stock prices, presented by Louis Bachelier in [6]. In his proposed model, Bachelier assumed that the discount rate is zero and the stochastic differential equation (SDE) governing the stock price is as follows:

$$dS(t) = \sigma dW(t) \tag{1}$$

where $S(t)$ is the spot stock price at time t , σ is diffusion coefficient (known as volatility), and $\{W(t), t \geq 0\}$ is a scalar Wiener process (a standard Brownian motion). Integration of Eq. 1 over $(t, t + \Delta t)$ yields the stochastic solution of the equation:

$$\Delta S(t) = \sigma \Delta W(t)$$

where $\Delta S(t) = S(t + \Delta t) - S(t)$ is the relative change in the price during a time lag Δt , and $\Delta W(t) = W(t + \Delta t) - W(t)$ is the increment of the wiener process which is computed as $\Delta W(t) = \eta \sqrt{\Delta t}$, where η is a random variable that follows a normal (Gaussian) distribution with zero mean and unit variance, i.e., $\eta \sim N(0, 1)$. Therefore, the following can be written:

$$S(t + \Delta t) = S(t) + \sigma \eta \sqrt{\Delta t} \tag{2}$$

The main shortcoming of Bachelier’s model is that it assumes that the future value of the assets follows a normal distribution. Based on this assumption, Eq. 2 can lead to a negative stock price with a positive probability, which is not possible in reality. In [7] Osborn demonstrated that the future value of the stock should follow a log-normal distribution, but the log-return of the stock follows a normal distribution. Shortly, the Bachelier model was modified by Samuelson in [8], where he introduced the “geometric Brownian motion” model (also known as Black-Scholes model). In this model, it is assumed that the price of the risky stock evolves according to the following SDE:

$$dS(t) = \mu S(t)dt + \sigma S(t)dW(t) \tag{3}$$

where μ and σ are the drift and diffusion coefficients, and again $\{W(t), t \geq 0\}$ is a scalar Wiener process. The field of mathematical finance has gained significant attention since Black and Scholes published their work in [9, 10]. They contributed to the world of finance via the introduction of Itô calculus to financial mathematics, and also the Black-Scholes formula. By choosing $y(t) = \ln[S(t)]$ and applying Itô’s lemma [11, 12], Eq. 3 becomes:

$$dy(t) = \left(\mu - \frac{\sigma^2}{2} \right) dt + \sigma dW(t) \tag{4}$$

Integration of Eq. 4 over $(t, t + \Delta t)$ gives us:

$$\Delta y(t) = \left(\mu - \frac{\sigma^2}{2} \right) \Delta t + \sigma \Delta W(t)$$

where $\Delta y(t) = y(t + \Delta t) - y(t) = \ln \left[\frac{S(t + \Delta t)}{S(t)} \right]$ represents the logarithmic increment of stock price data (known as log-return), Δt is the length of time interval between two consecutive trading periods, and $\Delta W(t) = \eta \sqrt{\Delta t}$, $\eta \sim N(0, 1)$. Therefore, the following can be written:

$$\Delta y(t) = \left(\mu - \frac{\sigma^2}{2} \right) \Delta t + \sigma \eta \sqrt{\Delta t} \tag{5}$$

In turn, the stock price can be determined from Eq. 5 as:

$$S(t + \Delta t) = S(t) e^{\left[\left(\mu - \frac{\sigma^2}{2} \right) \Delta t + \sigma \eta \sqrt{\Delta t} \right]} \tag{6}$$

Eq. 6 enables one to simulate the possible stock price trajectories with time step Δt , through the Black-Scholes model. For this purpose, one must first find the parameters μ and σ^2 from historical log-returns data based on the following relations:

$$M_1 = \langle \Delta y(t) \rangle = \left(\mu - \frac{\sigma^2}{2} \right) \Delta t$$

$$M_2 = \langle (\Delta y(t) - \langle \Delta y(t) \rangle)^2 \rangle = \sigma^2 \Delta t \tag{7}$$

where $\langle \dots \rangle$ denotes averaging over the data, so that M_1 and M_2 in Eq. 7 are the mean and variance of the historical log-returns data, respectively. Having M_1 and M_2 , first σ^2 is obtained:

$$\sigma^2 = \frac{1}{\Delta t} M_2$$

once σ^2 is identified, the parameter μ is obtained from the first moment M_1 .

The main disadvantage of the Black-Scholes model is its constant volatility assumption, while it is widely believed and empirically confirmed that stock prices do not have constant volatility, rather it varies during time [13–15]. This shortcoming and unsatisfactory performance of the Black-Scholes model caused researchers look for better alternatives and improve the classic Black-Scholes model in two directions:

- 1- Adding a term with jumpy behavior to the Black-Scholes equation to allow for random jumps in the stock price process (jump-diffusion model e.g., Merton model [16])
- 2- Considering stochastic volatility for the stock price (e.g., Heston model [17] or GARCH model [18]).

Here we only focus on the first option and describe the jump-diffusion model. Merton in [16] presented one of the first models in which jump processes were used in financial modeling. To take into account price discontinuities, Merton added a Poisson jump process to the log-price while preserving the independence and stationarity of log-returns. A jump-diffusion equation is generally written as:

$$dy(t) = \bar{\mu} dt + \sigma dW(t) + \xi dJ(t) \tag{8}$$

where $\bar{\mu}$ and σ are the drift and diffusion coefficients, $W(t)$ is a wiener process, and $J(t)$ is a Poisson jump process with rate λ and distributed size ξ which Merton assumed follows a Gaussian distribution with zero mean and variance σ_ξ^2 , i.e., $\xi \sim N(0, \sigma_\xi^2)$. It was also assumed that Poisson process, jump size ξ and Wiener process in Eq. 8 are three independent processes.

Integration of Eq. 8 over $(t, t + \Delta t)$, leads to:

$$\Delta y(t) = \bar{\mu} \Delta t + \sigma \Delta W(t) + \xi \Delta J(t)$$

here $\Delta J(t) = J(t + \Delta t) - J(t)$ follows a Poisson distribution with mean $\lambda \Delta t$, and $\Delta W(t) = \eta \sqrt{\Delta t}$, $\eta \sim N(0, 1)$. Therefore, the following can be written:

$$\Delta y(t) = \bar{\mu}\Delta t + \sigma\eta\sqrt{\Delta t} + \xi\Delta J(t) \tag{9}$$

In turn, the stock price can be determined from Eq. 9 as:

$$S(t + \Delta t) = S(t) e^{(\bar{\mu}\Delta t + \sigma\eta\sqrt{\Delta t} + \xi\Delta J(t))} \tag{10}$$

Eq. 10 enables one to simulate the possible stock price trajectories with time step Δt , via the jump-diffusion model. For this purpose, one must find the parameters $\bar{\mu}$, σ^2 , σ_ξ^2 and $\lambda\Delta t$ from the historical log-returns data based on the following relations:

$$\begin{aligned} M_1 &= \langle \Delta y(t) \rangle = \bar{\mu}\Delta t \\ M_2 &= \langle (\Delta y(t) - \bar{\mu}\Delta t)^2 \rangle = \sigma^2\Delta t + \sigma_\xi^2\lambda\Delta t \\ M_4 &= \langle (\Delta y(t) - \bar{\mu}\Delta t)^4 \rangle = 3\sigma_\xi^4\lambda\Delta t \\ M_6 &= \langle (\Delta y(t) - \bar{\mu}\Delta t)^6 \rangle = 15\sigma_\xi^6\lambda\Delta t \end{aligned} \tag{11}$$

where M_1 , M_2 , M_4 and M_6 are the statistical moments of the historical log-returns data. Having these moments, first the jump characteristics σ_ξ^2 and $\lambda\Delta t$ are obtained from Eq. 11:

$$\begin{aligned} \sigma_\xi^2 &= \frac{M_6}{5M_4} \\ \lambda\Delta t &= \frac{M_4}{3\sigma_\xi^4} \end{aligned}$$

once σ_ξ^2 and $\lambda\Delta t$ are identified, the parameter σ^2 is identified from the second moment M_2 and the parameter $\bar{\mu}$ is obtained from the first moment M_1 .

The main shortcoming of the jump-diffusion model is that the jumps reconstructed by the model have larger amplitudes than the jumps in the actual data. Let us demonstrate how this problem occurs. Suppose we want to model the daily prices of a stock via jump-diffusion model. As mentioned, first we need to determine the parameters $\bar{\mu}$, σ^2 , σ_ξ^2 and $\lambda\Delta t$ from the historical log-returns data. Since in the Poisson jump process, the probability of occurrence of more than one jump in any small time interval Δt is zero, so ΔJ in Eq. 9 takes only the values of one (one jump) or 0 (no jump) with the probabilities $\lambda\Delta t$ and $1 - \lambda\Delta t$, respectively [19]. Given these probabilities, the data points can be reconstructed by one of the following sub-equations:

If $\Delta J = 0$, meaning that no jump occurs at that Δt , then the data point is reconstructed by:

$$\Delta y(t) = \bar{\mu}\Delta t + \sigma\eta\sqrt{\Delta t} \tag{12}$$

If $\Delta J = 1$, meaning that a jump occurs at that Δt , then the data point is reconstructed by:

$$\Delta y(t) = \bar{\mu}\Delta t + \sigma\eta\sqrt{\Delta t} + \xi \tag{13}$$

As can be seen from Eqs 12, 13, the diffusion term $\sigma\eta\sqrt{\Delta t}$ appears in both equations and is involved in the reconstruction of all data points, even jumpy data points. Since the random variables $(\sigma\eta\sqrt{\Delta t})$ and (ξ) are two independent zero mean normally distributed variables with variances $\sigma^2\Delta t$ and σ_ξ^2 , respectively, their sum in Eq. 13 is also a normally distributed variable, i.e., $(\sigma\eta\sqrt{\Delta t} + \xi) \sim N(0, \sigma^2\Delta t + \sigma_\xi^2)$. The variance of this distribution $(\sigma^2\Delta t + \sigma_\xi^2)$ represents the amplitude of the reconstructed jumps, which is larger than the amplitude of the jumps in the historical data (σ_ξ^2) that was originally obtained. Obviously, if $\sigma^2\Delta t \ll \sigma_\xi^2$, so that $\sigma^2\Delta t$ can be neglected compared

to σ_ξ^2 , then the data reconstructed by the jump-diffusion model will be similar to the original data in the statistical sense, otherwise the model will fail. This shortcoming led us to modify the jump-diffusion equation in such a way that, if necessary, we can discard the contribution of the diffusion term in Eq. 9 so that it does not interfere with the reconstruction of the jumps. For this purpose, we replace the diffusion term in Eq. 9 by a term with jumpy behavior. This idea is supported by the fact that when data are sampled at discrete times, they appear as a sequence of discontinuous jump events, even if they have been sampled from a continuous diffusion process [20]. This is precisely why distinguishing between discontinuities due to discrete sampling of continuous process and real discontinuities in a jump-diffusion process is itself a challenging task [21]. Based on the above considerations, we modify Eq. 9 by considering two jump process with different distributed sizes ξ_1 and ξ_2 and different rates $\lambda_1\Delta t$ and $\lambda_2\Delta t$ as follows:

$$\Delta y(t) = \bar{\mu}\Delta t + \xi_1\Delta J_1(t) + \xi_2\Delta J_2(t)$$

where $\xi_1\Delta J_1(t)$ has replaced diffusion term in Eq. 9, and $\xi_2\Delta J_2(t)$ has the same role as $\xi\Delta J$. Each of $\Delta J_1(t)$ and $\Delta J_2(t)$ take the values of 1 and 0, but to avoid their simultaneous occurrence, we stipulate that if $\Delta J_1(t) = 1$, then $\Delta J_2(t) = 0$ and *vice versa*. Applying this condition causes each data point to be reconstructed by only one of the jump events. The procedure is as follows:

If $\Delta J_1(t) = 1$, and $\Delta J_2(t) = 0$, then the data point is reconstructed by:

$$\Delta y(t) = \bar{\mu}\Delta t + \xi_1$$

If $\Delta J_1(t) = 0$, and $\Delta J_2(t) = 1$, then the data point is reconstructed by:

$$\Delta y(t) = \bar{\mu}\Delta t + \xi_2$$

With this modification, the shortcoming of the jump-diffusion model can be solved. In this model, we assume that ξ_1 and ξ_2 are two zero mean Gaussian random variables with variances $\sigma_{\xi_1}^2$ and $\sigma_{\xi_2}^2$, i.e., $\xi_1 \sim N(0, \sigma_{\xi_1}^2)$ and $\xi_2 \sim N(0, \sigma_{\xi_2}^2)$. These two random variables produce fluctuations that are additively superimposed on the trajectory generated by the deterministic dynamics. In the following, we will describe the model in detail and demonstrate that all the unknown parameters of this modeling can be derived directly from the historical stock price.

2 Model description

In [22] we have introduced a general dynamical stochastic equation as follows, which includes a deterministic drift term ($\bar{\mu}dt$) and a combination of stochastic terms with jumpy behaviors ($\xi_i dJ_i(t)$):

$$dy(t) = \bar{\mu}dt + \sum_{i=1}^N \xi_i dJ_i(t) \tag{14}$$

where $\bar{\mu}dt$ indicates the deterministic part of the process and $J_1(t), J_2(t), etc$ are Poisson jump processes. The jumps have rates $\lambda_1, \lambda_2, etc$ and sizes ξ_1, ξ_2, etc , which we assume they have zero mean Gaussian distributions with variances $\sigma_{\xi_1}^2, \sigma_{\xi_2}^2, etc$, respectively. In this article, we intend to use this equation

specifically to simulate asset prices. For this purpose, we first start with the simplest form of Eq. 14, which includes a drift term and only a jump process. We will demonstrate that such a jump-drift equation is able to describe the discrete-time evolution of price time series in the Black-Scholes markets. Since the real markets are usually jump-diffusion markets, in the second section, we extend modeling by considering two jump processes with different rates (λ_1, λ_2) and different distributed sizes (ξ_1, ξ_2) and use it to model prices in actual markets. In each stage, we will demonstrate that all unknown parameters involved in the model can be derived non-parametrically from the historical price data. It should be noted that due to the small number of data points in the price time series, or the lack of diversity in the distributed sizes of fluctuations, we will model prices only by considering two jump processes. However, depending on the number of available data points and the variety of amplitudes, one can extend the proposed model.

2.1 Jump-drift modeling

In the first step, we consider Eq. 14 in its simplest form including a drift term and a stochastic term with jumpy behavior, and show that it can be used to reconstruct prices data of the diffusion markets (e.g., Black-Scholes markets). The general form of a jump-drift equation is as follows:

$$dy(t) = \bar{\mu}dt + \xi dJ(t) \tag{15}$$

where $y(t) = \ln(S(t))$ is the log-price and $\bar{\mu}dt$ denotes the deterministic drift part of the dynamics and $J(t)$ is a Poisson jump process characterized by the rate λ and the size ξ . We assume that ξ is a random variable with zero mean Gaussian distribution, i.e., $\xi \sim N(0, \sigma_\xi^2)$. Also, we assume that Poisson-distributed jumps $dJ(t)$ and jump size ξ are two independent processes.

Integration of Eq. 15 over $(t, t + \Delta t)$ gives us:

$$\Delta y(t) = \bar{\mu}\Delta t + \xi\Delta J(t) \tag{16}$$

where $\Delta y(t) = y(t + \Delta t) - y(t) = \ln[\frac{S(t+\Delta t)}{S(t)}]$ is the log-return, Δt is the length of time interval between two consecutive points and $\Delta J(t) = J(t + \Delta t) - J(t)$ follows a Poisson distribution with mean $\lambda\Delta t$.

In turn, the stock price can be determined as:

$$S(t + \Delta t) = S(t) e^{(\bar{\mu}\Delta t + \xi\Delta J(t))}$$

To reconstruct prices data with the above relation, we must find three parameters $\bar{\mu}$, σ_ξ^2 and $\lambda\Delta t$. We now show that all these parameters can be estimated directly from the log-return time series $\Delta y(t)$. For this purpose, we derive the statistical moments of $\Delta y(t)$ from Eq. 16 (note that $\Delta J(t)$ and ξ are two independent processes):

$$\begin{aligned} M_1 &= \langle \Delta y(t) \rangle = \langle \bar{\mu}\Delta t \rangle + \langle \xi \rangle \langle \Delta J(t) \rangle \\ M_2 &= \langle (\Delta y(t) - \bar{\mu}\Delta t)^2 \rangle = \langle \xi^2 \rangle \langle (\Delta J(t))^2 \rangle \\ M_4 &= \langle (\Delta y(t) - \bar{\mu}\Delta t)^4 \rangle = \langle \xi^4 \rangle \langle (\Delta J(t))^4 \rangle \end{aligned}$$

where $\langle \dots \rangle$ denotes averaging over the data, so that M_1 is the mean of log-returns, and M_2, M_4 and M_6 are the other statistical moments of log-returns about the mean. Since for small Δt , all of the statistical moments of jumps are proportional to $\lambda\Delta t$,

i.e., $\langle (\Delta J(t))^m \rangle = \lambda\Delta t$ [19, 20], as well as for a zero mean Gaussian random variable ξ with variance σ_ξ^2 , all of the even order statistical moments are obtained by $\langle \xi^{2l} \rangle = \frac{2^l l!}{2^l l!} \langle \xi^2 \rangle^l$, the above relations become (note that $\langle \xi \rangle = 0$ and $\langle \xi^2 \rangle = \sigma_\xi^2$):

$$\begin{aligned} M_1 &= \bar{\mu}\Delta t \\ M_2 &= \sigma_\xi^2 \lambda\Delta t \\ M_4 &= 3\sigma_\xi^4 \lambda\Delta t \end{aligned} \tag{17}$$

According to the first relation in Eq. 17, the mean of log-returns (M_1) gives us the drift parameter $\bar{\mu}$, and the second and fourth-order moments (M_2, M_4) identify the jump characteristics, namely:

$$\begin{aligned} \bar{\mu} &= \frac{1}{\Delta t} M_1 \\ \sigma_\xi^2 &= \frac{M_4}{3M_2} \\ \lambda\Delta t &= \frac{M_2}{\sigma_\xi^2} \end{aligned} \tag{18}$$

We claim that the proposed jump-drift dynamics enable us to model diffusion processes such as the Black-Scholes process. We will check the validity of this claim by reconstructing a Black-Scholes process via the new dynamics using the parameters determined from Eq. 18. But before that, let us provide the following two criteria for evaluating the reconstructed process:

- 1) We know from Wick's theorem that for the time series of the Black-Scholes process, the statistical moments of the data satisfy the relation $\frac{M_4}{3M_2^2} \approx 1$, which follows from the fact that the short-time propagator of the Black-Scholes dynamics is a Gaussian distribution. Therefore, if the proposed jump-drift dynamics be capable of reconstructing a time series which is statistically similar to the original Black-Scholes time series, then the statistical moments of the reconstructed data should satisfy the Wick's relation, i.e., $(\frac{M_4}{3M_2^2})_{rec} \approx 1$.
- 2) In continuation of the previous point, we find the ratio $\frac{M_4}{3M_2^2}$ from relations (17):

$$\frac{M_4}{3M_2^2} = \frac{3\sigma_\xi^4 \lambda\Delta t}{3(\sigma_\xi^2 \lambda\Delta t)^2} = \frac{1}{\lambda\Delta t}$$

by comparing this relation with Wick's relation, i.e., $\frac{M_4}{3M_2^2} \approx 1$, we expect that $\lambda\Delta t = 1$. On the other, if $\lambda\Delta t = 1$, then the second moment in Eq. 17 becomes:

$$M_2 = \sigma_\xi^2$$

this is while, the second moment in original Black-Scholes process is $M_2 = \sigma^2\Delta t$ (Eq. 7). Therefore, it can be concluded that if the new model works correctly, the estimation of jumps amplitude (σ_ξ^2) should be equal to the variance of the original data ($\sigma^2\Delta t$), namely,

$$\sigma_\xi^2 = \sigma^2\Delta t$$

In the following, we reconstruct a Black-Scholes process with known drift and volatility parameters via the jump-drift equation, and then evaluate the reconstructed data.

Example 1. First, we generate a synthetic time series $\Delta y(t)$ with 10^6 data points via Black-Scholes dynamics (Eq. 5) and using preset

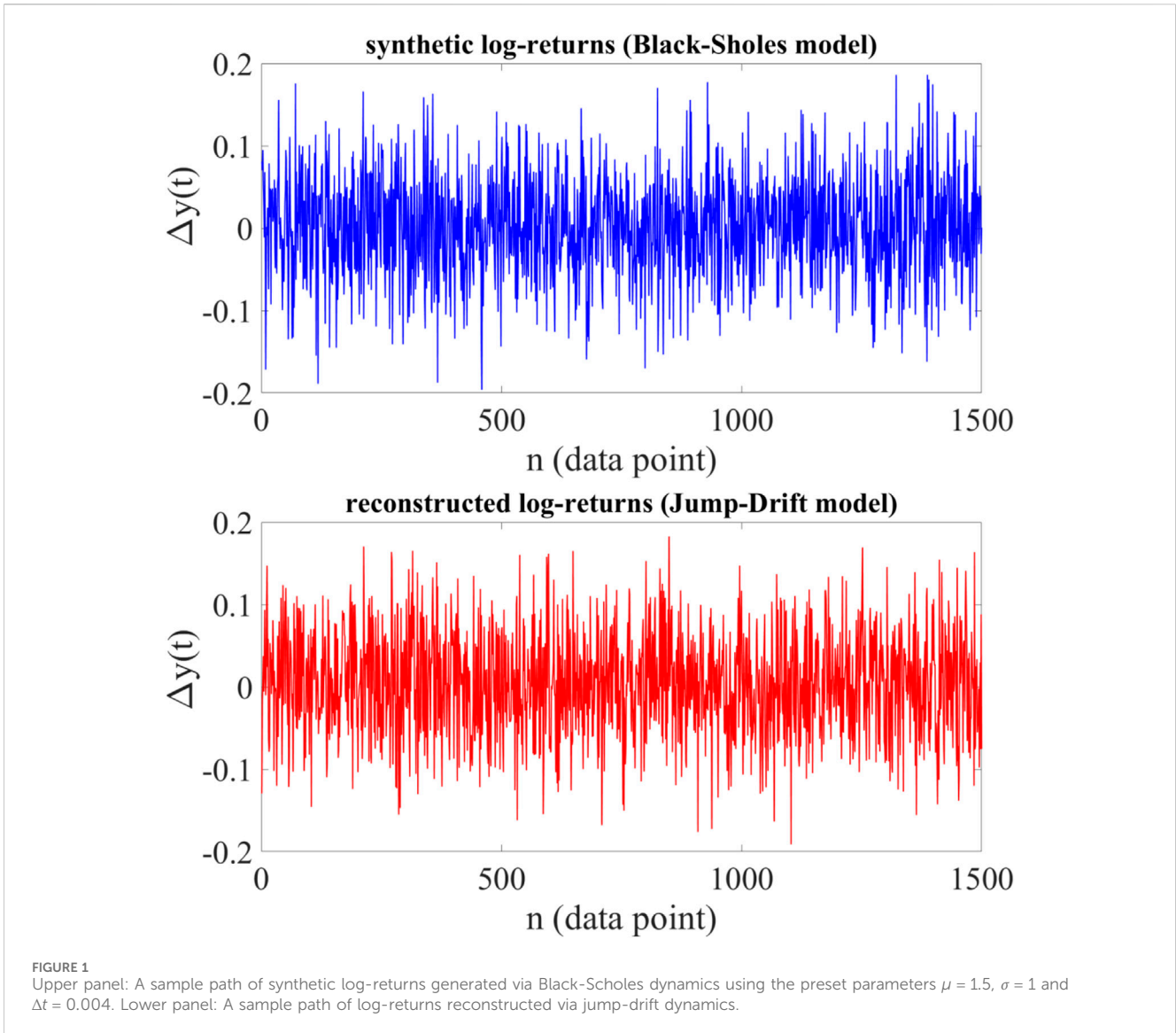


TABLE 1 A scale of judgment of forecast accuracy.

Mean absolute percentage error (MAPE) (%)	Judgment of forecast accuracy
Less than 10%	Highly Accurate
11% to 20%	Good Accurate
21% to 50%	Reasonable Forecast
more than 50%	Inaccurate Forecast

parameters $\mu = 1.5$ and $\sigma = 1$ with $\Delta t = 0.004$. In Figure 1, we have shown the trajectory of 1,500 data points out of 10^6 generated data points so that the fluctuations can be clearly seen (blue graph). By obtaining the statistical moments M_n for $n = 1, 2, 4$ from the generated data, and substituting in relations (18), we determine the parameters required for the new modeling. The results are as follows:

TABLE 2 values of the drift, jump amplitudes, and jump rates obtained from historical daily prices of Apple and IBM stocks using the jump-jump-drift modeling.

Stock	$\bar{\mu}$	$\sigma_{\xi_1}^2$	$\lambda_1 \Delta t$	$\sigma_{\xi_2}^2$	$\lambda_2 \Delta t$
APPLE	0.2772	2.5345×10^{-4}	0.8248	0.0011	0.1752
IBM	0.1283	8.8795×10^{-5}	0.6790	3.7351×10^{-4}	0.3210

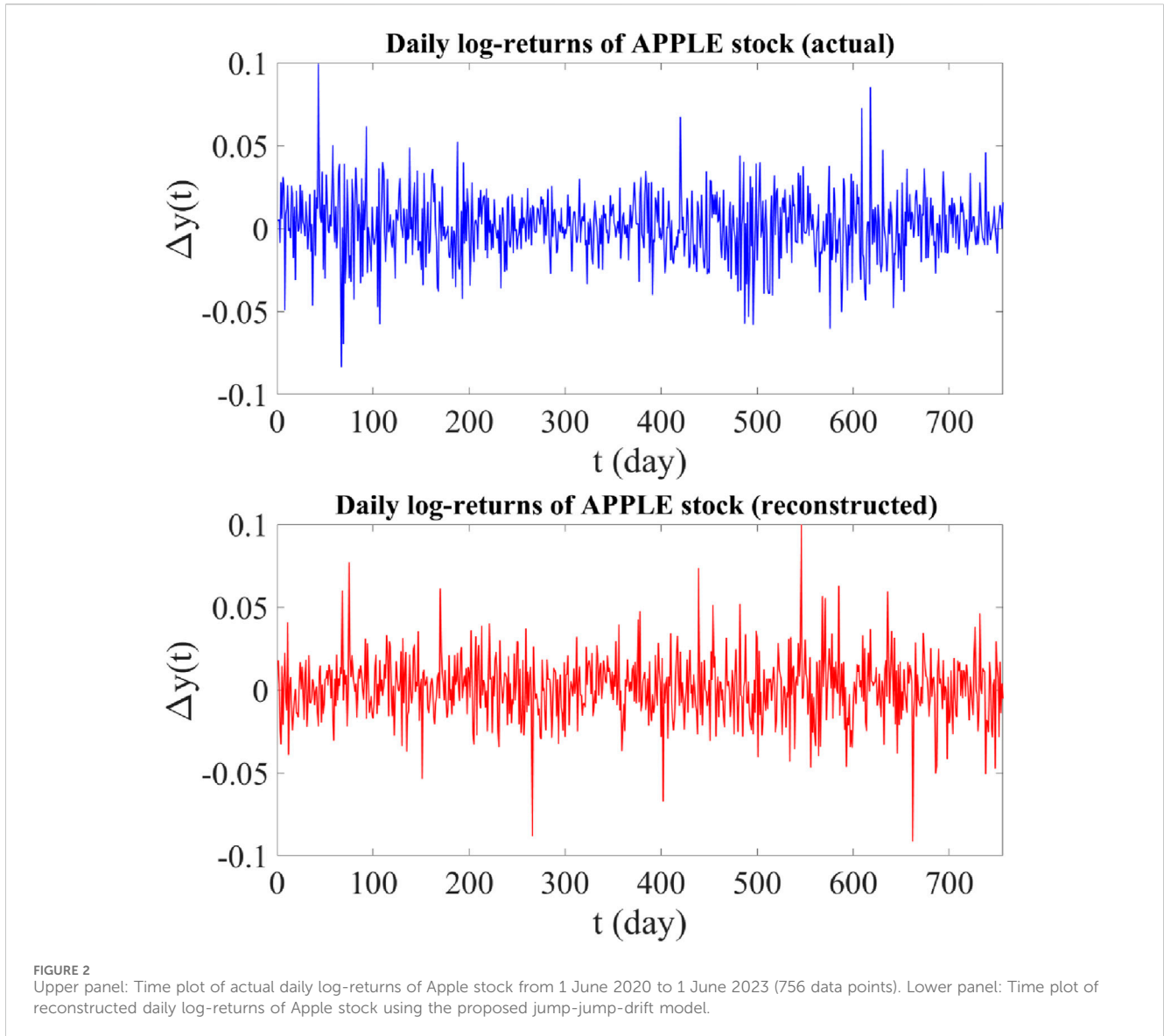
Statistical moments determined from generated data:

$$M_1 = 0.004, M_2 = 0.004, M_4 = 4.8121 \times 10^{-5}, \frac{M_4}{3M_2^2} = 1.002$$

Required parameters for new modeling:

$$\bar{\mu} = \frac{1}{\Delta t} M_1 = 1$$

$$\sigma_{\xi}^2 = \frac{M_4}{3M_2} = 0.00401 \text{ (In agreement with } \sigma^2 \Delta t = 0.004)$$



$$\lambda\Delta t = \frac{M_2}{\sigma_\xi^2} = 0.997 \approx 1$$

In the second step, we reconstruct a time series $\Delta y(t)$ via the jump-drift equation (Eq. 16) with 10^6 data points. For comparison with the original data, a sample path including 1,500 reconstructed data points is shown in Figure 1 (red graph). Finally, to ensure that the two time series (generated and reconstructed) are statistically equivalent, we obtain the statistical moments of the reconstructed data, and check the establishment of $(\frac{M_4}{3M_2^2})_{rec} \approx 1$. The results are as follows:

Statistical moments of reconstructed data:

$$M_1 = 0.004, M_2 = 0.004, M_4 = 4.8172 \times 10^{-5},$$

$$\left(\frac{M_4}{3M_2^2}\right)_{rec} = 0.9985 \approx 1$$

As can be seen, the reconstructed data are statistically similar to original data with high accuracy, and there is a very good agreement between these results and the theory.

2.2 Jump-jump-drift modeling

In the previous section we modeled a continuous diffusion process through the jump-drift equation. Since real markets are usually jump-diffusion markets, the generalizing of jump-drift modeling to a jump-jump-drift modeling improves the characterization of real markets dynamics beyond a continuous process. The general form of a jump-jump-drift equation is as follows:

$$dy(t) = \bar{\mu}dt + \xi_1 dJ_1(t) + \xi_2 dJ_2(t) \tag{19}$$

where $\bar{\mu}dt$ indicates the deterministic part of the process and $J_1(t)$ and $J_2(t)$ are Poisson jump processes. The jumps have rates λ_1 and λ_2 , and sizes ξ_1 and ξ_2 , which we assume have zero mean Gaussian distributions, i.e., $\xi_1 \sim N(0, \sigma_{\xi_1}^2)$ and $\xi_2 \sim N(0, \sigma_{\xi_2}^2)$. We call $\sigma_{\xi_1}^2$ and $\sigma_{\xi_2}^2$ the jump amplitudes.

Integration of Eq. 19 over $(t, t + \Delta t)$ gives us:

$$\Delta y(t) = \bar{\mu}\Delta t + \xi_1 \Delta J_1(t) + \xi_2 \Delta J_2(t) \tag{20}$$

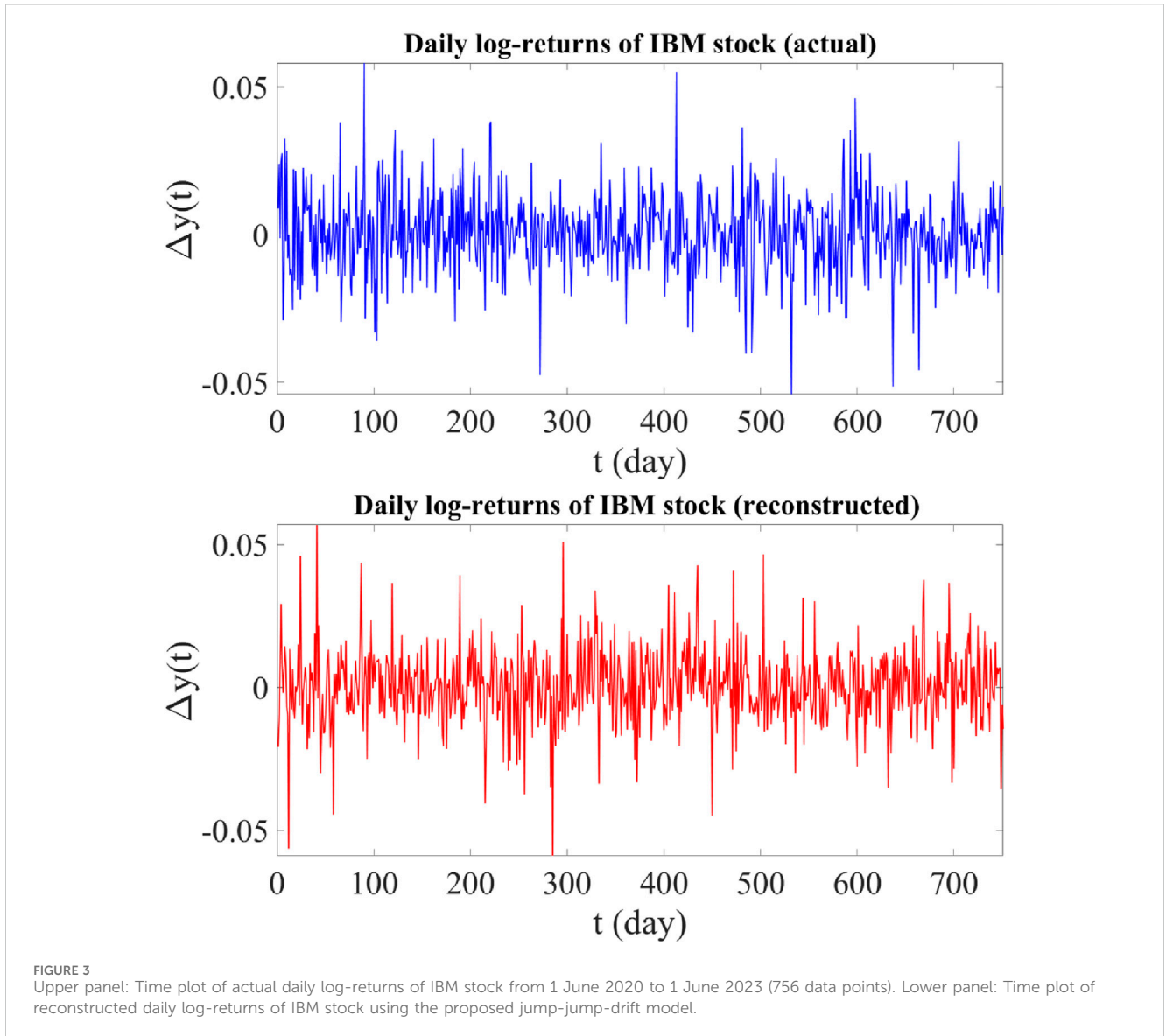


FIGURE 3 Upper panel: Time plot of actual daily log-returns of IBM stock from 1 June 2020 to 1 June 2023 (756 data points). Lower panel: Time plot of reconstructed daily log-returns of IBM stock using the proposed jump-jump-drift model.

Furthermore, the stock price can be determined from Eq. 20 as:

$$S(t + \Delta t) = S(t) e^{(\bar{\mu}\Delta t + \xi_1 \Delta J_1(t) + \xi_2 \Delta J_2(t))} \tag{21}$$

In modeling the stock price via Eq. 21, we also assume that two jumps do not occur simultaneously, which means that in the time interval $(t, t + \Delta t]$, if, for example, $\Delta J_1(t)$ occurs and takes the value of 1, $\Delta J_2(t)$ does not occur and its value is 0, and *vice versa*. Let $\lambda_1 \Delta t$ and $\lambda_2 \Delta t$ be the probabilities of occurrence of $\Delta J_1(t)$ and $\Delta J_2(t)$ in a small time step Δt , if we assume only one of the jumps ($\Delta J_1(t)$ or $\Delta J_2(t)$) occurs in each time step, then we can write:

$$\lambda_1 \Delta t + \lambda_2 \Delta t = 1 \tag{22}$$

According to this condition, we can discard one of the jump events at each time step, and reconstruct the corresponding data point by another jump event.

To model the stock prices via Eq. 20, we must find the five unknown parameters $\bar{\mu}$, $\lambda_1 \Delta t$, $\lambda_2 \Delta t$, $\sigma_{\xi_1}^2$ and $\sigma_{\xi_2}^2$. We now show that all these parameters can be estimated directly from the log-returns

time series $\Delta y(t)$. For this purpose, we derive the statistical moments of $\Delta y(t)$ from Eq. 20 (note that ξ_1 and ξ_2 are two Gaussian random variables independent from the jumps, and $\Delta J_1(t)$ and $\Delta J_2(t)$ do not occur simultaneously):

$$M_1 = \langle \Delta y(t) \rangle = \langle \bar{\mu} \Delta t \rangle + \langle \xi_1 \rangle \langle \Delta J_1(t) \rangle + \langle \xi_2 \rangle \langle \Delta J_2(t) \rangle$$

$$M_2 = \langle (\Delta y(t) - \bar{\mu} \Delta t)^2 \rangle = \langle \xi_1^2 \rangle \langle (\Delta J_1(t))^2 \rangle + \langle \xi_2^2 \rangle \langle (\Delta J_2(t))^2 \rangle$$

$$M_4 = \langle (\Delta y(t) - \bar{\mu} \Delta t)^4 \rangle = \langle \xi_1^4 \rangle \langle (\Delta J_1(t))^4 \rangle + \langle \xi_2^4 \rangle \langle (\Delta J_2(t))^4 \rangle$$

$$M_6 = \langle (\Delta y(t) - \bar{\mu} \Delta t)^6 \rangle = \langle \xi_1^6 \rangle \langle (\Delta J_1(t))^6 \rangle + \langle \xi_2^6 \rangle \langle (\Delta J_2(t))^6 \rangle$$

By using the relations $\langle (\Delta J_1(t))^m \rangle = \lambda_1 \Delta t$ and $\langle (\Delta J_2(t))^m \rangle = \lambda_2 \Delta t$ for the statistical moments of jump processes, and the relations $\langle \xi_1^{2l} \rangle = \frac{2^l l!}{2^l l!} \langle \xi_1^2 \rangle^l$ and $\langle \xi_2^{2l} \rangle = \frac{2^l l!}{2^l l!} \langle \xi_2^2 \rangle^l$ for the even order statistical moments of zero mean Gaussian random variables ξ_1 and ξ_2 with variances $\sigma_{\xi_1}^2$ and $\sigma_{\xi_2}^2$, we will have (note that $\langle \xi_1 \rangle = \langle \xi_2 \rangle = 0$, $\langle \xi_1^2 \rangle = \sigma_{\xi_1}^2$, and $\langle \xi_2^2 \rangle = \sigma_{\xi_2}^2$):

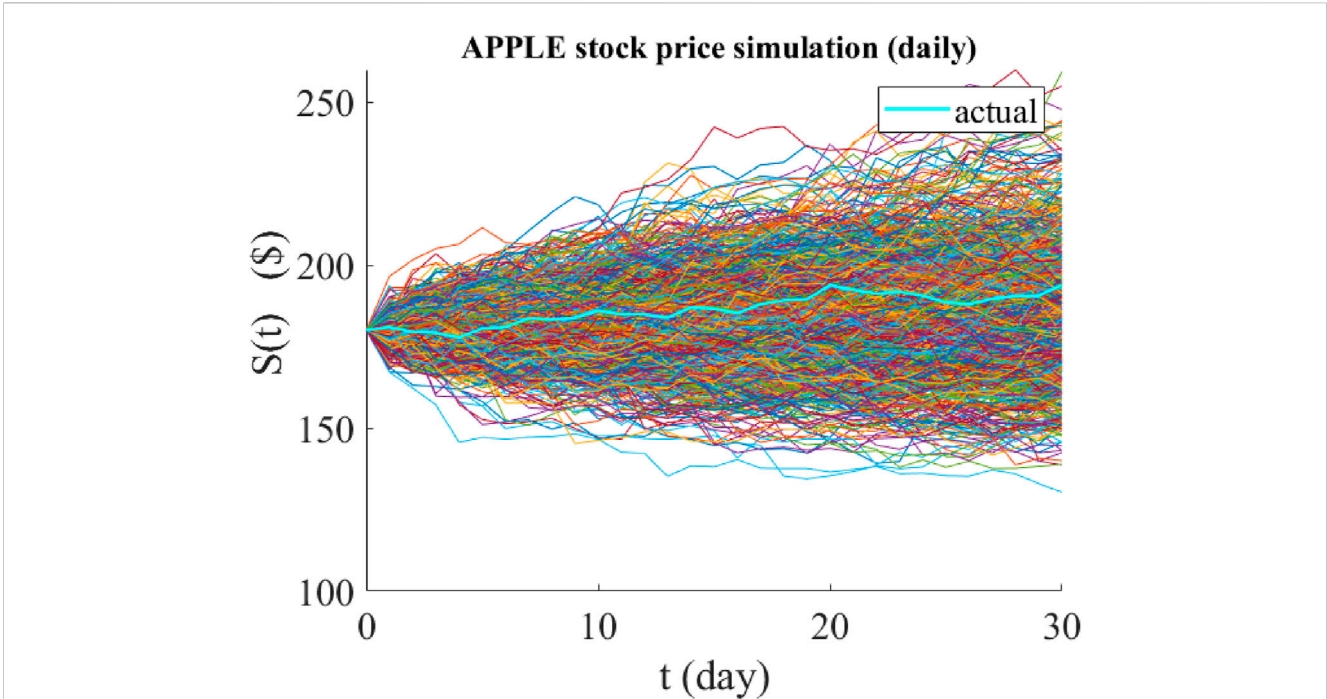


FIGURE 4 Graphical representation of the predicted paths of the daily price of Apple stock using jump-jump-drift modeling. The time period of all predictions is 30 days and their starting point is 1 June 2023. The cyan graph is the actual price path realized over the same 30 days, and the colored graphs are the 1,000 possible paths predicted by the model.

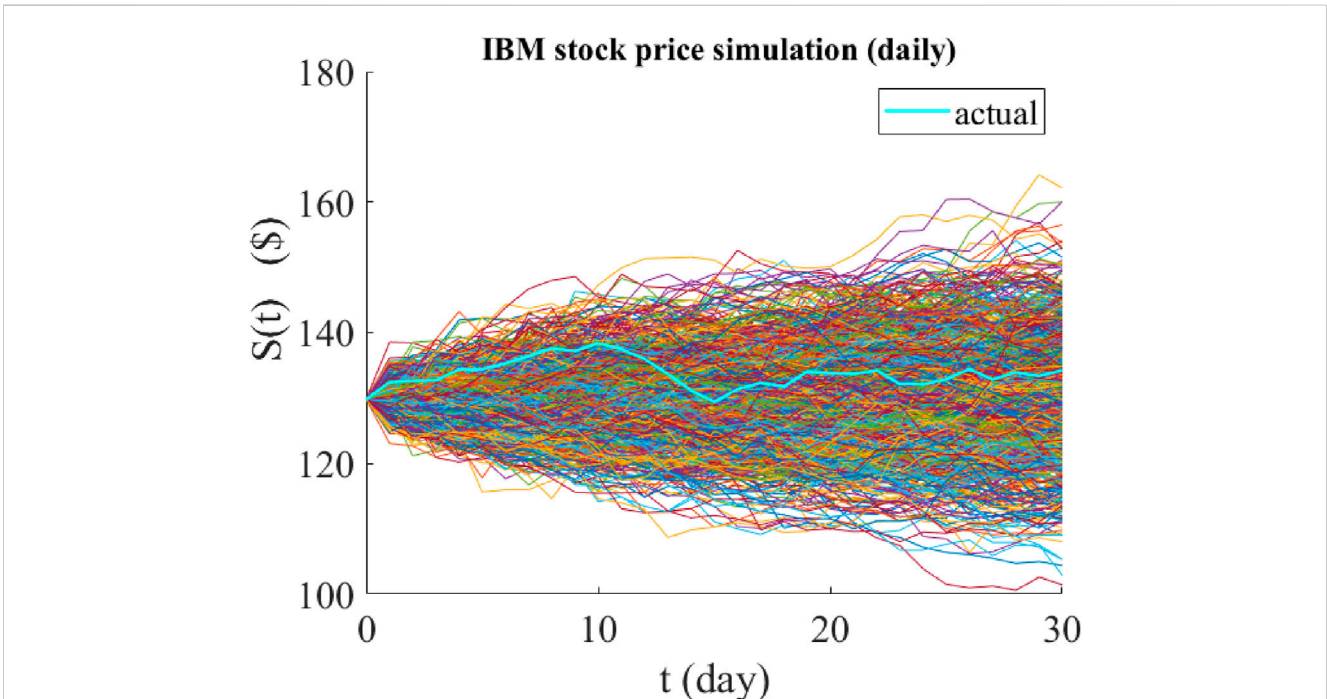


FIGURE 5 Graphical representation of the predicted paths of the daily price of IBM stock using jump-jump-drift modeling. The time period of all predictions is 30 days and their starting point is 1 June 2023. The cyan graph is the actual price path realized over the same 30 days, and the colored graphs are the 1,000 possible paths predicted by the model.

TABLE 3 values of the drift, jump amplitudes, and jump rates obtained from historical gold prices (weekly and hourly) using jump-jump-drift modeling.

Time horizon	$\bar{\mu}$	$\sigma_{\xi_1}^2$	$\lambda_1\Delta t$	$\sigma_{\xi_2}^2$	$\lambda_2\Delta t$
weekly	0.0728	2.0482×10^{-4}	0.54	9.5290×10^{-4}	0.46
hourly	-0.2114	2.0184×10^{-6}	0.9228	2.3761×10^{-5}	0.0772

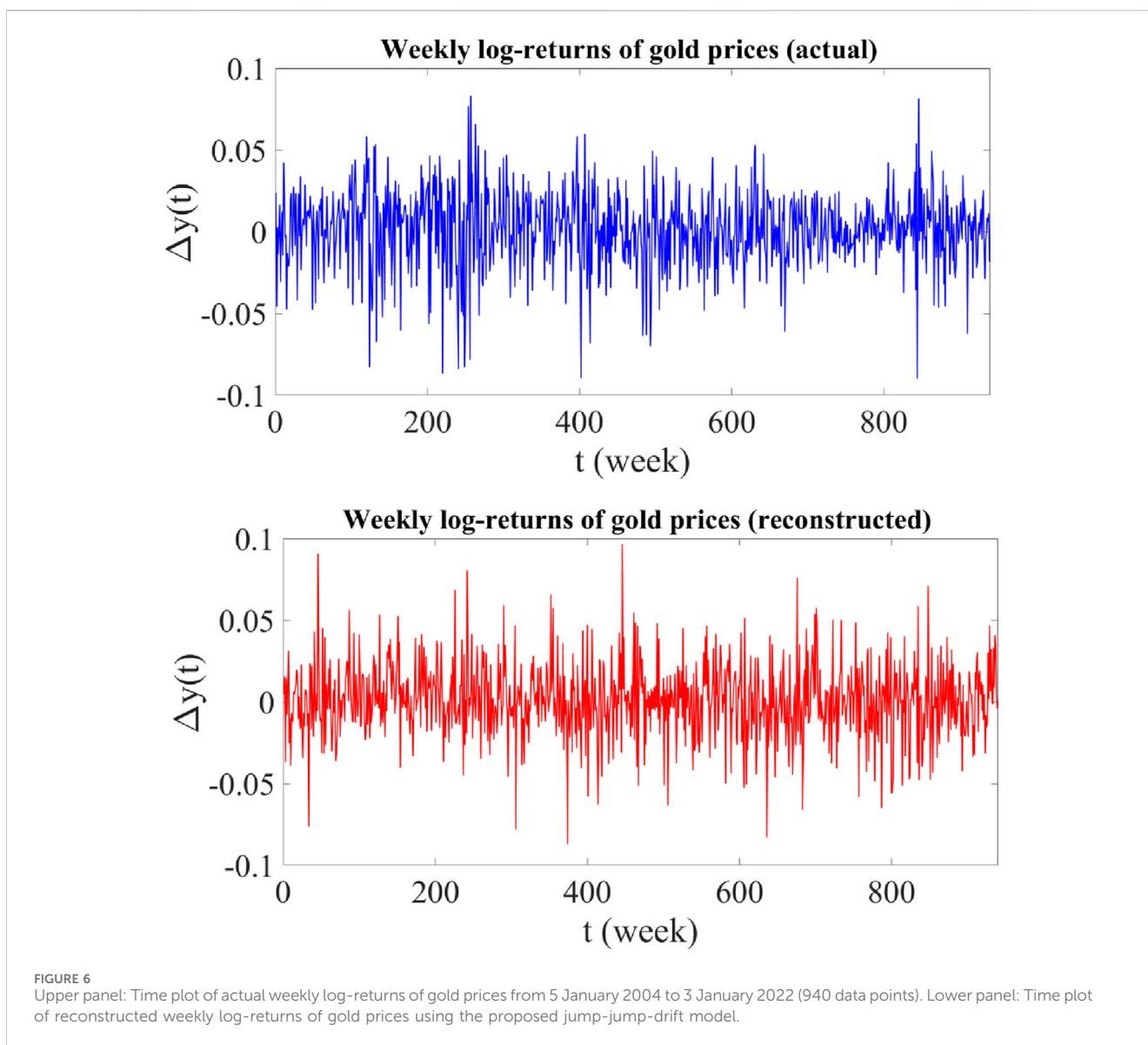
$$\begin{aligned}
 M_1 &= \bar{\mu}\Delta t \\
 M_2 &= \sigma_{\xi_1}^2 \lambda_1\Delta t + \sigma_{\xi_2}^2 \lambda_2\Delta t \\
 M_4 &= 3\sigma_{\xi_1}^4 \lambda_1\Delta t + 3\sigma_{\xi_2}^4 \lambda_2\Delta t \\
 M_6 &= 15\sigma_{\xi_1}^6 \lambda_1\Delta t + 15\sigma_{\xi_2}^6 \lambda_2\Delta t
 \end{aligned}$$

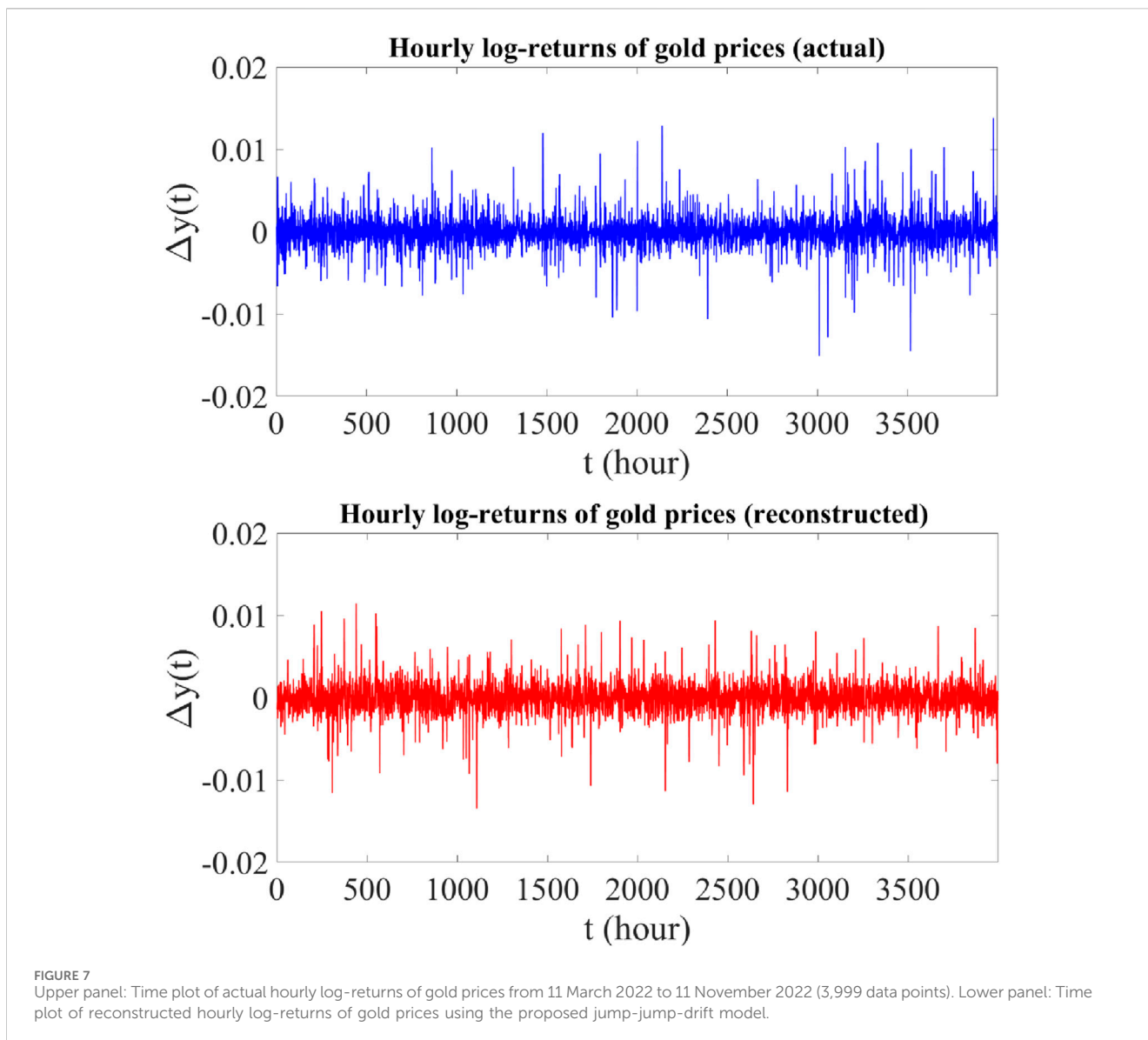
To find the five unknowns $\bar{\mu}, \lambda_1\Delta t, \lambda_2\Delta t, \sigma_{\xi_1}^2$ and $\sigma_{\xi_2}^2$, we need to add one more equation to the above relations. For this purpose, we use Eq. 22 as $\lambda_1\Delta t = 1 - \lambda_2\Delta t$, and reduce the number of unknowns, so we will have:

$$\begin{aligned}
 M_1 &= \bar{\mu}\Delta t \\
 M_2 &= \sigma_{\xi_1}^2 + (\sigma_{\xi_2}^2 - \sigma_{\xi_1}^2) \lambda_2\Delta t \\
 M_4 &= 3\sigma_{\xi_1}^4 + 3(\sigma_{\xi_2}^4 - \sigma_{\xi_1}^4) \lambda_2\Delta t \\
 M_6 &= 15\sigma_{\xi_1}^6 + 15(\sigma_{\xi_2}^6 - \sigma_{\xi_1}^6) \lambda_2\Delta t
 \end{aligned} \tag{23}$$

Having the statistical moments M_1, M_2, M_4 and M_6 from the log-return time series and solving the above system of equations numerically, the four unknown parameters $\bar{\mu}, \lambda_2\Delta t, \sigma_{\xi_1}^2$ and $\sigma_{\xi_2}^2$ are determined. Once $\lambda_2\Delta t$ is identified, $\lambda_1\Delta t$ is obtained from Eq. 22.

We claim that the proposed dynamics enables us to model time series with jump discontinuities more accurately than the classic jump-diffusion dynamics. We will check the validity of this claim by reconstructing a jump-diffusion process via the jump-jump-drift equation. But before that, let us prove this claim by showing that the new relations in Eq. 23 are generalizations of the old jump-diffusion relations in Eq. 11. For this purpose, we consider the case in which





$\sigma_{\xi_1}^2 \ll \sigma_{\xi_2}^2$, so that $\sigma_{\xi_1}^2$ can be ignored compared to $\sigma_{\xi_2}^2$, and at the same time $\sigma_{\xi_1}^2$ be so small that $\sigma_{\xi_1}^4 = \sigma_{\xi_1}^6 = 0$. Under these special conditions, relations (23) can be written as follows:

$$\begin{aligned}
 M_1 &= \bar{\mu}\Delta t \\
 M_2 &= \sigma_{\xi_1}^2 + \sigma_{\xi_2}^2 \lambda_2 \Delta t \\
 M_4 &= 3\sigma_{\xi_2}^4 \lambda_2 \Delta t \\
 M_6 &= 15\sigma_{\xi_2}^6 \lambda_2 \Delta t
 \end{aligned}$$

As can be seen, these relations are similar to relations of jump-diffusion model (Eq. 11), so that $\sigma_{\xi_1}^2$ has replaced $\sigma^2 \Delta t$, and identifies the diffusion part, and $\sigma_{\xi_2}^2 \lambda_2 \Delta t$ has the same role as $\sigma_{\xi}^2 \lambda \Delta t$. This means that under these special conditions ($\sigma_{\xi_1}^2 \ll \sigma_{\xi_2}^2$ and $\sigma_{\xi_1}^4 = \sigma_{\xi_1}^6 = 0$), the new model works like the jump-diffusion model and the parameters obtained from the data are the same in both models. But if the data fluctuations are such that these conditions are not satisfied, it is clear that the proposed model will lead to more accurate estimates than the jump-diffusion model. By analyzing stock prices, we found that although the release of exciting news in the

market causes sudden jumps in log-returns, the amplitude of these jumps is not so much larger than the amplitude of the fluctuations in normal days. Therefore, it seems that the new model has a better performance for modeling and forecasting prices.

In the following, to demonstrate the reliability of the new model, we test it on synthetic data. Furthermore, to ensure the effectiveness of the proposed approach in different conditions, we test the model with different synthetic data.

Example 2: First, we test the model with data generated through the Black-Scholes process in Example 1. By obtaining the statistical moments M_n for $n = 1, 2, 4, 6$ from the generated data, and replacing them in relations (23), we determine the parameters required for the new modeling via the numerical solution of the obtained system of equations. Since the data generated in example 1 are diffusive data, and we have already modeled it through the jump-drift equation, we expect the occurrence rate of one of the jumps to be zero when we model the same data through the jump-jump-drift equation. The following results, confirm our opinion:

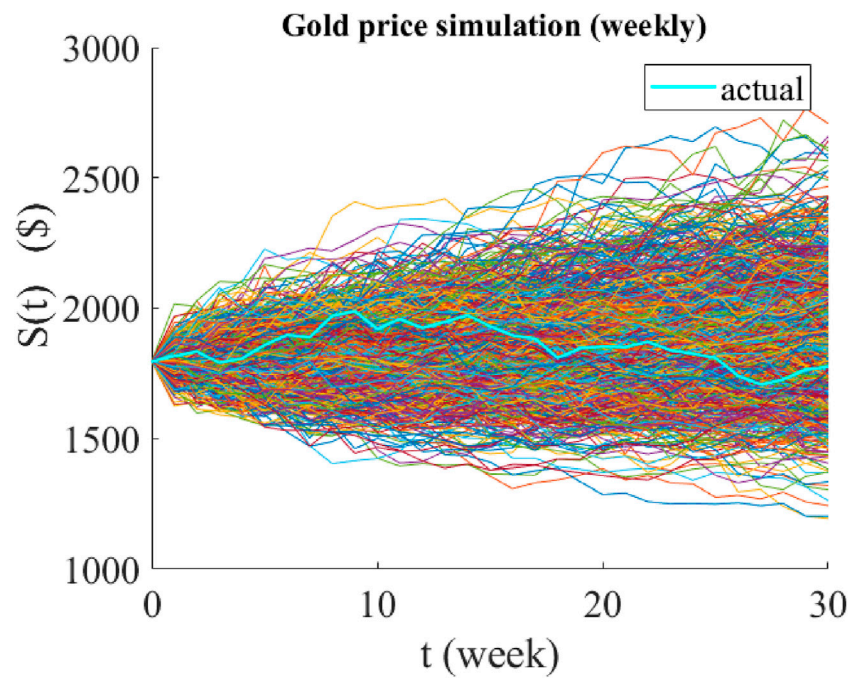


FIGURE 8 Graphical representation of the predicted paths of the weekly price of gold using jump-jump-drift modeling. The time period of all predictions is 30 weeks and their starting point is 3 January 2022. The cyan graph is the actual price path realized over the same 30 weeks, and the colored graphs are the 1,000 possible paths predicted by the model.

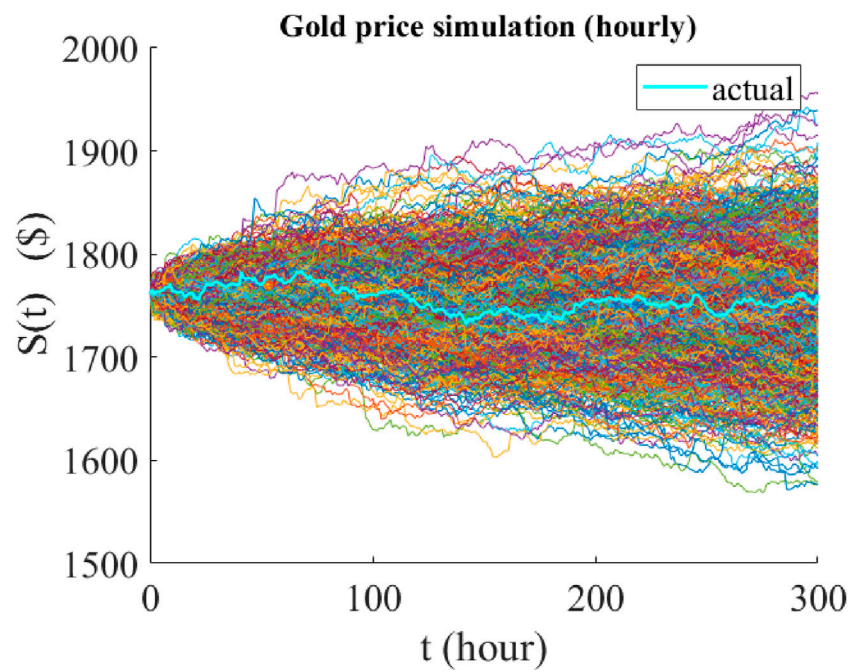


FIGURE 9 Graphical representation of the predicted paths of the hourly price of gold using jump-jump-drift modeling. The time period of all predictions is 300 h and their starting point is 11 November 2022. The cyan graph is the actual price path realized over the same 300 h, and the colored graphs are the 1,000 possible paths predicted by the model.

$$\begin{aligned} \bar{\mu} &= 1 \\ \sigma_{\xi_1}^2 &= 0.004 \\ \lambda_1 \Delta t &= 0.9999 \approx 1 \\ \sigma_{\xi_2}^2 &= 0.0007 \\ \lambda_2 \Delta t &= 0.0001 \approx 0 \end{aligned}$$

The value of $\lambda_2 \Delta t \approx 0$ show that when the time series belongs to the class of continuous diffusion processes (e.g., Black-Scholes process), the jump-jump-drift dynamics, models it by using only one jump process and completely omitting the second jump process. In the next step, we test the model on two synthetic log-return time series generated via jump-diffusion Equation 9 with preset parameters. Each time series contains 3×10^6 data points which generated by considering $\bar{\mu} = 5$ and $\sigma = 2$ with a sampling interval $\Delta t = 0.0001$, so that $\sigma^2 \Delta t = 0.0004$. The jumps in both time series have the same jump rate $\lambda \Delta t = 0.3$ (jump rate per data point), but the amplitude of the jumps are $\sigma_{\xi}^2 = 0.1$ and $\sigma_{\xi}^2 = 0.001$, respectively. We deliberately choose these jump amplitudes with different orders of magnitude to observe the effect of their amplitude in retrieving the coefficients, Note that in the first case $\frac{\sigma_{\xi}^2}{\sigma^2 \Delta t} = 250$ and in the second case $\frac{\sigma_{\xi}^2}{\sigma^2 \Delta t} = 2.5$, that is, in the first case, the variance of diffusion part ($\sigma^2 \Delta t$) is negligible compared to the amplitude of jumps (σ_{ξ}^2), and as mentioned earlier, we expect both models show almost the same results, but in the second case, we expect the estimates of the new model to be more accurate than the jump-diffusion model.

By obtaining the statistical moments M_n for $n = 1, 2, 4, 6$ from the generated data, and substituting in relations (11) and (23), we determine the parameters of the two models. The following results are estimated from the numerical solution of the corresponding system of equations:

Case1:

Preset parameters:

$$\bar{\mu} = 1, \sigma^2 \Delta t = 0.0004, \sigma_{\xi}^2 = 0.1, \lambda \Delta t = 0.3$$

Estimated parameters via jump-diffusion model:

$$\bar{\mu} = 1, \sigma^2 \Delta t = 0.00031, \sigma_{\xi}^2 = 0.1005, \lambda \Delta t = 0.299$$

Estimated parameters via jump-jump-drift model:

$$\bar{\mu} = 1, \sigma_{\xi_1}^2 = 0.00045, \sigma_{\xi_2}^2 = 0.1005, \lambda_2 \Delta t = 0.299, \lambda_1 \Delta t = 0.701$$

Case2:

Preset parameters:

$$\bar{\mu} = 1, \sigma^2 \Delta t = 0.0004, \sigma_{\xi}^2 = 0.001, \lambda \Delta t = 0.3$$

Estimated parameters via jump-diffusion model:

$$\bar{\mu} = 1, \sigma^2 \Delta t = 0.00013, \sigma_{\xi}^2 = 0.0012, \lambda \Delta t = 0.5$$

Estimated parameters via jump-jump-drift model:

$$\bar{\mu} = 1, \sigma_{\xi_1}^2 = 0.00040, \sigma_{\xi_2}^2 = 0.0014, \lambda_2 \Delta t = 0.302, \lambda_1 \Delta t = 0.698$$

The above results show that in the first case, both models lead to almost the same results, but in the second case, the proposed model

leads to more accurate results (note that in the new model, $\sigma_{\xi_1}^2$ is an estimate for the variance of the diffusive data, i.e., $\sigma_{\xi_1}^2 = \sigma^2 \Delta t$, and $\sigma_{\xi_2}^2$ is an estimate for the variance of the jumpy data, i.e., $\sigma_{\xi_2}^2 = \sigma_{\xi}^2$).

3 Data and methodology

Our dataset comprises the daily closing prices of the Apple and IBM stocks, as well as gold prices with two different time horizons (weekly and hourly). For Apple and IBM stocks, the historical data that will be used are daily closing prices from 1 June 2020 to 1 June 2023, which are obtained from Yahoo Finance source. For gold, the historical data that will be used are weekly gold prices from 5 January 2004 to 3 January 2022, as well as hourly gold prices from 11 March 2022 to 11 November 2022, which are obtained from dukascopy historical data source.

For each of the collected data, we will obtain log-returns time series $\Delta y(t)$ by:

$$\Delta y(t) = \ln \left[\frac{S(t + \Delta t)}{S(t)} \right], t = 1, 2, \dots$$

where $S(t)$ and $S(t + 1)$ are consecutive prices in the price time series. Afterwards, we will determine the statistical moments of log-returns data as follows:

$$\begin{aligned} M_1 &= \overline{\Delta y} = \frac{1}{N} \sum_{t=1}^N \Delta y(t) \\ M_n &= \frac{1}{N} \sum_{t=1}^N (\Delta y(t) - \overline{\Delta y})^n, \text{ for } n = 2, 4, 6 \end{aligned}$$

where N is the number of log-returns data points. By determining these statistical moments and replacing them in relations (23), we will identify the required parameters of the model, i.e., $\bar{\mu}$, $\sigma_{\xi_1}^2$, $\sigma_{\xi_2}^2$, $\lambda_1 \Delta t$, $\lambda_2 \Delta t$. Using these parameters, we will reconstruct the log-returns data by the following equation:

$$\Delta y(t) = \bar{\mu} \Delta t + \xi_1 \Delta J_1(t) + \xi_2 \Delta J_2(t)$$

In addition, we will use the following equation to forecast prices for several time steps after the chosen historical period:

$$S(t + 1) = S(t) e^{(\bar{\mu} \Delta t + \xi_1 \Delta J_1(t) + \xi_2 \Delta J_2(t))}$$

To determine the forecasts accuracy, we will use ‘‘Mean Absolute Percentage Error’’ (MAPE) calculation as follows:

$$MAPE = \frac{1}{N} \sum_{t=1}^N \frac{|F(t) - S(t)|}{S(t)}$$

where $F(t)$ is the forecasted price at time t , $S(t)$ is the actual stock price at time t , and N is the number of predicted data points. We will use MAPE values to evaluate our forecasting method. A scale for judging model accuracy based on the MAPE criterion was presented by Lawrence et al. [23], and is shown in Table 1.

3.1 Research output and discussion

In the following, considering the elements described in the methodology, we first model Apple and IBM stocks and predict

their prices for a period of 30 days. Historical data used are daily closing stock prices from 1 June 2020 to 1 June 2023. For daily prices, the trading period is $\Delta t = \frac{1}{252}$ years (based on a year with an average of 252 stock trading days). The parameters obtained from the model are presented in Table 2. Based on this table, in both stocks, $\sigma_{\xi_2}^2$ is not more than one order of magnitude larger than $\sigma_{\xi_1}^2$, so that for Apple stock we have $\frac{\sigma_{\xi_2}^2}{\sigma_{\xi_1}^2} = 4.34$, and for IBM stock this ratio is $\frac{\sigma_{\xi_2}^2}{\sigma_{\xi_1}^2} = 4.2$. As mentioned earlier, in this situation, the new model has better performance. Furthermore, for the Apple stock, the jump rates are $\lambda_1\Delta t = 0.8248$ and $\lambda_2\Delta t = 0.1752$, which means that, in the data reconstruction stage, 82.48% of the data points are reconstructed by using a Gaussian random variable with smaller variance $\sigma_{\xi_1}^2$, and 17/52% of the data points are reconstructed by using a Gaussian random variable with larger variance $\sigma_{\xi_2}^2$. This is while, these rates for the IBM stock are $\lambda_1\Delta t = 0.6790$ and $\lambda_2\Delta t = 0.3210$, respectively.

To reconstruct log-return data through proposed model, we use the equation $\Delta y(t) = \bar{\mu}\Delta t + \xi_1\Delta J_1(t) + \xi_2\Delta J_2(t)$ and reconstruct a time series for $\Delta y(t)$ that is statistically similar to the original ones. Figures 2, 3 show the actual and reconstructed log-returns of Apple and IBM stocks, respectively.

To predict stock prices, we use parameters estimated from historical data. The forecast period is 30 days and is related to the days after the selected historical period. Simulation of predictions is done by 1,000 realization of the trajectory. Each trajectory is realized using the equation $S(t+1) = S(t)e^{(\bar{\mu}\Delta t + \xi_1\Delta J_1(t) + \xi_2\Delta J_2(t))}$ with 30 iterations. Figures 4, 5 show the daily forecasts of Apple and IBM stock prices, respectively. In order to compare the predictions with the actual prices, the graph of realized prices in the same 30 days is also shown in each figure (cyan graph). As can be seen, in both stocks, the actual prices are located within the predicted trajectories. In addition, the data analysis shows that all 1,000 predicted trajectories of Apple stock price have MAPE values less than 20% (with the smallest MAPE = 1.4%, the largest MAPE = 19.8% and the average MAPE = 5.84%). Meanwhile, the corresponding values obtained through the jump-diffusion model are as follows:

The smallest MAPE = 1.51%, the largest MAPE = 24.3%, and the average MAPE = 6.32%. These results show that the jump-jump-drift model has a better performance than the jump-diffusion model, and if the time period of the forecasts becomes larger (e.g., in annual forecasts), the difference between the forecasts of the two models becomes more visible.

The results of IBM stock price predictions are even more surprising than Apple stock. Analysis of IBM stock simulation outputs shows that all 1,000 predicted trajectories have MAPE values less than 15% (with the smallest MAPE = 1.22%, the largest MAPE = 14.02% and the average MAPE = 4.62%), indicating good accuracy of the model predictions. The corresponding values obtained through jump-diffusion model are as follows:

The smallest MAPE = 1.62%, the largest MAPE = 20.7%, and the average MAPE = 5.12%.

Finally, to see the effectiveness of the proposed approach for different time horizons, we simulate gold prices with two different time horizons. Historical data used are weekly gold price from 5 January 2004 to 3 January 2022, as well as hourly gold prices from 11 March 2022 to 11 November 2022. For weekly prices, the trading

period is $\Delta t = \frac{1}{52}$ years (based on a year with an average of 52 trading weeks), while for hourly prices, the trading period is $\Delta t = \frac{1}{5916}$ years (based on 2022 with 5,916 trading hours). The parameters obtained from the model are presented in Table 3. Based on this table, in both cases, $\sigma_{\xi_2}^2$ is not more than one order of magnitude larger than $\sigma_{\xi_1}^2$, so that for weekly data we have $\frac{\sigma_{\xi_2}^2}{\sigma_{\xi_1}^2} = 4.6$, and for hourly data this ratio is $\frac{\sigma_{\xi_2}^2}{\sigma_{\xi_1}^2} = 11.7$.

To reconstruct log-return data through proposed model, we use the equation $\Delta y(t) = \bar{\mu}\Delta t + \xi_1\Delta J_1(t) + \xi_2\Delta J_2(t)$ and reconstruct a time series for $\Delta y(t)$ that is statistically similar to the original ones. Figures 6, 7 show the actual and reconstructed log-returns of weekly and hourly prices of gold, respectively.

To predict gold prices, we use parameters estimated from historical data. The forecast period for weekly price is 30 weeks and for hourly price is 300 h and related to the times after historical periods. Simulation of predictions is done by 1,000 realization of the trajectory. Each trajectory is realized using the equation $S(t+1) = S(t)e^{(\bar{\mu}\Delta t + \xi_1\Delta J_1(t) + \xi_2\Delta J_2(t))}$ with 30 iterations for weekly gold price and 300 iterations for hourly gold price. Figures 8, 9 show the weekly and hourly gold price forecasts, respectively. As can be seen, in both cases, the actual prices are located within the trajectories predicted by the model. Furthermore, the data analysis shows that all 1,000 predicted trajectories of weekly gold price have MAPE values less than 30% (with the smallest MAPE = 2.1%, the largest MAPE = 29.43% and the average MAPE = 7.57%), which are acceptable forecasts. The corresponding values obtained by jump-diffusion model are as follows:

The smallest MAPE = 2.3%, the largest MPPE = 35.7%, and the average MAPE = 7.85%.

Analysis of hourly gold prices shows that all 1,000 predicted paths have MAPE values less than 10% (with the smallest MAPE = 0.51%, the largest MPPE = 7.17% and the average MAPE = 1.82%), indicating very high accuracy of the model predictions for hourly time horizons. The corresponding values obtained by jump-diffusion model are as follows:

The smallest MAPE = 0.8%, the largest MPPE = 10.3%, and the average MAPE = 2.15%.

4 Conclusion

We discussed that when data are sampled at discrete times (e.g., stock prices), they appear as a sequence of discontinuous jump events, even if they have been sampled from a continuous process. This issue gave us the idea to propose a new modeling in which random variations in the sample path of a measured time series are attributed to jump events, even if the time series belongs to the class of diffusion processes. Based on this, we introduced a new dynamical stochastic equation including a deterministic drift term and a combination of several stochastic terms with jumpy behaviors. The general form of this equation is as follows:

$$dy(t) = \bar{\mu}dt + \sum_{i=1}^N \xi_i dJ_i(t)$$

In this modeling we also assumed that the jump events do not occur simultaneously so that the jumps have no overlap. We started with the simplest form of equation including a deterministic drift term

and a jump process as the stochastic component, and argued that it can be used to describe the discrete-time evolution of a diffusion process, e.g., Black-Scholes process. Afterwards, we extended the equation by considering two jump processes with different distributed sizes, and used it to model assets such as stock prices and gold prices with different time horizons. We also demonstrated that in all cases the proposed model works better than the old jump model. It should be noted that, due to the small number of available price data and the lack of diversity in the amplitudes of jumps, in this article we modeled prices data only by considering two jump processes. However, depending on the number of data points and variation in the amplitudes of fluctuations, more stochastic terms can be kept in the equation to increase the accuracy of the modeling. But on the other hand, the more the number of terms in the equation, the need to solve the system of equations with more unknowns, the cost of which must be paid in the form of longer runtime.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

AM: Conceptualization, Data curation, Formal Analysis, Funding acquisition, Investigation, Methodology, Project

administration, Resources, Software, Supervision, Validation, Visualization, Writing—original draft, Writing—review and editing. HN: Conceptualization, Data curation, Formal Analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing—original draft, Writing—review and editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Reddy K, Clinton V. Simulating stock prices using geometric Brownian motion: evidence from Australian companies. *Australas Account Business Finance J* (2016) 10(3):23–47. doi:10.14453/aabf.v10i3.3
- Synowicz D. Jump-diffusion models with constant parameters for financial log-return processes. *Comput Math Appl* (2008) 56(8):2120–7. doi:10.1016/j.camwa.2008.02.051
- Azizah M, Irawan M, Putri E. Comparison of stock price prediction using geometric Brownian motion and multilayer perceptron. In: AIP conference proceedings: Depok, Indonesia. AIP Publishing (2020).
- Mota PP, Esquivel ML. Model selection for stock prices data. *J Appl Stat* (2016) 43(16):2977–87. doi:10.1080/02664763.2016.1155205
- Benninga S. *Financial modeling, fourth edition* By Simon Benninga Hardcover. The MIT press (2014).
- Bachelier L. Théorie de la spéculation. In: *Annales scientifiques de l'École normale supérieure* (1900).
- Osborne MF. Brownian motion in the stock market. *Operations Res* (1959) 7(2):145–73. doi:10.1287/opre.7.2.145
- Samuelson PA. Economic theory and mathematics--an appraisal. *Am Econ Rev* (1952) 42(2):56–66.
- Black F, Scholes M. The pricing of options and corporate liabilities. *J Polit economy* (1973) 81(3):637–54. doi:10.1086/260062
- Black F, Karasinski P. Bond and option pricing when short rates are lognormal. *Financial Analysts J* (1991) 47(4):52–9. doi:10.2469/faj.v47.n4.52
- Risken H. *Springer series in synergetics* (1996).
- Bouchaud J-P, Cont R. A Langevin approach to stock market fluctuations and crashes. *Eur Phys J B-Condensed Matter Complex Syst* (1998) 6:543–50. doi:10.1007/s100510050582
- Hull J, White A. The pricing of options on assets with stochastic volatilities. *J Finance* (1987) 42(2):281–300. doi:10.1111/j.1540-6261.1987.tb02568.x
- Mercurio D, Spokoiny V. *Estimation of time dependent volatility via local change point analysis* (2005).
- Goldentayer L, Klebaner F, Liptser RS. Tracking volatility. *Probl Inf Transm* (2005) 41:212–29. doi:10.1007/s11122-005-0026-2
- Merton RC. Option pricing when underlying stock returns are discontinuous. *J financial Econ* (1976) 3(1-2):125–44. doi:10.1016/0304-405x(76)90022-2
- Heston SL. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev financial Stud* (1993) 6(2):327–43. doi:10.1093/rfs/6.2.327
- Nelson DB. ARCH models as diffusion approximations. *J Econom* (1990) 45(1-2):7–38. doi:10.1016/0304-4076(90)90092-8
- Anvari M, Tabar MRR, Peinke J, Lehnertz K. Disentangling the stochastic behavior of complex time series. *Scientific Rep* (2016) 6(1):35435. doi:10.1038/srep35435
- Tabar R. *Analysis and data-based reconstruction of complex nonlinear dynamical systems*, 730. Springer (2019).
- Lehnertz K, Zabawa L, Tabar MRR. Characterizing abrupt transitions in stochastic dynamics. *New J Phys* (2018) 20(11):113043. doi:10.1088/1367-2630/aaf0d7
- Movahed AA, Noshad H. Introducing a new approach for modeling a given time series based on attributing any random variation to a jump event: jump-jump modeling. *Scientific Rep* (2024) 14(1):1234. doi:10.1038/s41598-024-51863-5
- Excel FOFU. *Fundamentals of forecasting using excel* (2019).