ARTICLE IN PRESS

Int. J. Production Economics xxx (xxxx) xxx

Contents lists available at ScienceDirect

International Journal of Production Economics

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

Data-driven prediction for volatile processes based on real option theories[☆]

Abdullah AlShelahi[a,c,*], Jingxing Wang[b], Mingdi You[b], Eunshin Byon[a], Romesh Saigal[a]

[a] Industrial and Operations Engineering, University of Michigan, 1205 Beal Avenue, Ann Arbor MI 48109-2117, United States of America

[b] Ford Motor Company, 22001 Michigan Ave, Dearborn, MI 48124, United States of America

[c] Department of Industrial and Management Systems Engineering, College of Engineering and Petroleum, Kuwait University, Kuwait

ARTICLE INFO

Keywords:
Predictive modeling
Time-varying geometric brownian motion
Real options
Manufacturing

ABSTRACT

This paper presents a new prediction model for time series data by integrating a time-varying Geometric Brownian Motion model with a pricing mechanism used in financial engineering. Typical time series models such as Auto-Regressive Integrated Moving Average assumes a linear correlation structure in time series data. When a stochastic process is highly volatile, such an assumption can be easily violated, leading to inaccurate predictions. We develop a new prediction model that can flexibly characterize a time-varying volatile process without assuming linearity. We formulate the prediction problem as an optimization problem with unequal overestimation and underestimation costs. Based on real option theories developed in finance, we solve the optimization problem and obtain a predicted value, which can minimize the expected prediction cost. We evaluate the proposed approach using multiple datasets obtained from real-life applications including manufacturing, and finance. The numerical results demonstrate that the proposed model shows competitive prediction capability, compared with alternative approaches.

1. Introduction

In many applications including manufacturing, energy, and finance, accurate prediction is required to support strategic, tactical and/or operational decisions of organization (Chatfield, 2000). When physical information about the underlying mechanism that generates the time series data is limited, data-driven methods can be useful for predicting future observations (Zhang, 2003). In general, data-driven forecasting methods predict future observations based on past observations (Box et al., 2015). Several data-driven methods have been proposed in the literature for modeling time series data, among which Auto-Regressive Integrated Moving Average (ARIMA) and its variants such as the ARIMA-General Auto Regressive Conditional Heteroskedasticity (ARIMA-GARCH) have been widely used in many applications due to their flexibility and statistical properties (Sohn and Lim, 2007; Hahn et al., 2009; Lu et al., 2014; Ruppert, 2015). ARIMA assumes a constant standard deviation of stochastic noises, whereas ARIMA-GARCH extends it by allowing the standard deviation to vary over time. Some studies modify the original ARIMA model to update the parameters using new observations (Ledolter, 1981; Tran and Reed, 2004). The basic idea of these ARIMA-based models is that the future observation can be predicted by using a linear combination of past observations (and estimated noises). Therefore they assume a linear correlation structure between consecutive observations (Brooks, 2002). However, when the underlying dynamics exhibits a highly volatile process, such a simple linear structure may provide poor prediction performance (Kantz and Schreiber, 2004).

This study aims to provide accurate predictions for a highly volatile and time-varying stochastic process whose underlying dynamics is complicated and possibly nonlinear. As an example, let us consider a prediction problem faced by a contract manufacturer (CM) located in Michigan in the U.S, which motivates this study. The CM is a manufacturing company that produces various automotive parts, such as front and rear bumper beams, for several large automotive companies worldwide. The CM deals with a large number of orders for bumper beams from several automotive companies and the order sizes are time-varying. The CM should plan its production capacity carefully so that it can deliver products promptly when it gets orders. When an actual order size is greater than expected (i.e., when an order size is underestimated), overtime wages must be paid to workers to meet demands. On the other hand, when an order size is smaller than predicted (i.e., when an order size is overestimated), workers and equipment become idle.

As such, CM wants to predict future order sizes accurately, so that it can reduce its operating costs resulting from the discrepancy between its predicted value and actual sizes. Currently, CM uses its

☆ This work was supported by the National Science Foundation, United States under Grant IIS-1741166.

* Corresponding author.
E-mail addresses: shelahi@umich.edu (A. AlShelahi), jeffwjx@umich.edu (J. Wang), mingdyou@umich.edu (M. You), ebyon@umich.edu (E. Byon), rsaigal@umich.edu (R. Saigal).

https://doi.org/10.1016/j.ijpe.2019.107605
Received 3 July 2019; Received in revised form 15 November 2019; Accepted 23 December 2019
Available online 24 December 2019
0925-5273/

own proprietary prediction model, but its prediction performance is not satisfactory. The details of CM's proprietary model are confidential, so we cannot find reasons for its unsatisfactory performance. When we apply the ARIMA and ARIMA-GARCH models to CM's datasets, we also do not obtain significantly better prediction results (detailed results will be provided in Section 3). We believe such poor performance of ARIMA-based approaches is because they cannot fully characterize the underlying volatile dynamics. In addition to historical data, the future order size may depend on other factors which possibly make the order process behave nonlinearly. A new prediction approach that can adapt to such time-varying, and possibly nonlinear, dynamics is needed for providing better forecasts.

To this end we develop a new method for predicting future values in highly volatile processes, based on real option pricing theories typically used in financial engineering. In valuing investment projects, standard Discount Cash Flow (DCF) methods may fail to capture future opportunities (Benninga and Tolkowsky, 2002). The research on valuing investments under uncertainty was motivated by the pioneering work of Black and Scholes (Black and Scholes, 1973) and Marton (Merton et al., 1973) in financial options theory. As a result, options pricing theory has emerged in capital budgeting, and valuing projects or real options. The real options approach has an advantage over traditional methods because it includes the value of various managerial options (Hahn and Dyer, 2008; Bengtsson, 2001). Unlike standard financial options, real options reflect "real" and not traded assets or investments. Other investment options include deferring building a product, abandoning a project upon completion, or expanding a product to a new market (Trigeorgis et al., 1996). Consider a project with two stages: a pilot stage for research and development (R&D) and a second stage for commercialization. At the end of the first stage, project managers can decide whether to proceed to the second stage or to terminate the project and avoid potential poor outcomes. Real options can be used to evaluate such decision (Datar and Mathews, 2004), in which a call option represents the pilot stage. To justify the project, the investments in the first stage should not exceed the option's value

One of the popularly used stochastic process models for pricing real options is the Geometric Brownian Motion (GBM) model. Brownian motion is a continuous-time stochastic process, describing random movements in time series variables. The GBM, which is a stochastic differential equation, incorporates the idea of Brownian motion and consists of two terms: a deterministic term to characterize the main trend over time and a stochastic term to account for random variations. In GBM the random variations are represented by Brownian Motion (Björk, 2009). GBM is useful to model a positive quantity whose changes over equal and non-overlapping time intervals are identically distributed and independent.

The GBM and its variants have represented various real processes in finance, physics, etc. (Gardiner, 1986; Wu and Wu, 2015; Zhai and Ye, 2017). In particular, it is fundamental to many asset pricing models (Björk, 2009), and recently it has been applied to facilitate the use of a rich area of options theory to solve various pricing problems (see, for example, Whitt, 1981; Thorsen, 1999; Benninga and Tolkowsky, 2002; Nembhard et al., 2002; Boomsma et al., 2012; Chiu et al., 2017). However, most of the current studies in real options have been limited to solving pricing problems and have not been used to make forecasts (Xiao et al., 2015).

In this study, by utilizing the full power of real options theory, we present a new approach for predicting future observations when the system's underlying dynamics follows the GBM process. In particular, we allow the GBM parameters to adaptively change over time in order to characterize time-varying dynamics. The inhomogeneous GBM offers significantly more flexibility in characterizing the non-stationary nature of the stochastic process. Moreover, GBM, which includes the volatility term that changes dynamically, can capture highly volatile and nonlinear processes.

We formulate the prediction problem as an optimization problem and provide a solution using real option theories. Our approach provides extra flexibility by allowing overestimation (or over-prediction) to be handled differently from underestimation (or under-prediction). The overestimation and underestimation costs are determined in real life applications, depending on a decision-maker's (or organization's) preference. For example, in the aforementioned CM case, overestimation and underestimation of order sizes could cause different costs. The CM may want to put a larger penalty on the demand underestimation than on the overestimation, so that it can avoid extra overtime wages. We incorporate unequal overestimation and underestimation costs into the optimization problem and find the optimal forecast that minimizes the expected prediction cost. To the best of our knowledge, our study is the first attempt to incorporate options theory in the prediction problem.

The proposed approach with unequal over and underestimation penalties is closely related to the bad news principle (Bernanke, 1983). This principle states that an investment decision is only affected by the severity of future expected bad news. The choice of the parameter ω introduced in our approach, which denotes the ratio of overestimation penalty to the underestimation penalty, is highly dependent on the expected severity of future outcomes. In the CM example, ω reflects the expected severity of overtime and inventory costs.

Another way to see the role of ω is through good and bad economic uncertainties. The changes in economic activities depend on the type of uncertainty (Segal et al., 2015). Good uncertainty represents positive shocks that translate into higher asset prices and investments. On the other hand, bad uncertainty is associated with the volatility that impacts asset prices and investments negatively. The two types of volatility can be illustrated in two well-known financial incidents, namely, the high tech revolution in the 1990s and the collapse of Lehman Brothers. The former incident has introduced positive volatility and hence growth to the economy. The latter collapse has sent negative shocks through the markets. The variations of good and bad uncertainty have been demonstrated in Segal et al. (2015) where each has significant but opposing impacts on the economy. The key finding in Segal et al. (2015) is that good uncertainty predicts future growth in the economy while bad uncertainty predicts falling in asset prices. In our model, one may think the introduced parameter ω is a proxy to risk preference or type of uncertainty. In situations where positive shocks are expected, overestimation ($\omega > 1$) is preferred. In this case, the predicted value will be biased in that direction. Conversely, bad uncertainty would be associated with underestimation ($\omega < 1$). The predicted value, therefore, will be biased towards that preference.

To evaluate the prediction performance, we use two datasets collected from different applications, including the demand for bumper beams in CM (manufacturing), and stock prices (finance). We compare the performance of our model with ARIMA and ARIMA-GARCH models (and the proprietary prediction model in the CM case study) with different combinations of overestimation and underestimation costs. In most cases, our model outperforms those alternative models. In particular, we find that when the process is highly time-varying such as stock prices, the proposed approach provides much stronger prediction capability than ARMA and ARIMA-GARCH.

The remainder of the paper is organized as follows. The mathematical formulation and solution procedure are discussed in Section 2. Section 3 provides numerical results in two different applications. Section 4 concludes the paper.

2. Methodology

2.1. Problem formulation

Consider a real-valued variable $S(t)$ which represents a system state at time t. For example, the state variable can be a stock market index price, or a manufacturer's order size. This state variable is assumed to follow an inhomogeneous GBM with time-varying parameters.

Let us consider a filtered probability space $(\Omega, \mathcal{F}, P, \mathcal{F}_t)$, where the filtration $\mathcal{F}_t$ is generated by the Brownian motion W, i.e. $\mathcal{F}_t = \mathcal{F}_t^W$ so that $\mathcal{F}_t$ contains all information generated by $W(t)$, up to and including time t. With GBM, the stochastic process $S(t)$ is modeled by the following dynamics.

$$dS(t) = \mu(t)S(t)dt + \sigma(t)S(t)dW(t), \tag{1}$$

where $\sigma(t)$ denotes the volatility of $S(t)$ and $\mu(t)$ represents a drift process. The stochastic process $W(t)$ represents the Brownian motion where the increment $W(t+\Delta t) - W(t)$ during the time interval Δt is normally distributed with mean 0 and variance Δt, denoted by $\mathcal{N}(0, \Delta t)$, and $W(t)$ is assumed to be stationary.

Our objective is to predict $S(T)$ in the future time at $T(> t)$ when the current time is t. Solving (1) by using Itô's lemma (Shreve, 2004), we obtain

$$S(T) = S(t)\, exp\Big(\int_t^T \Big(\mu(s) - \frac{1}{2}\sigma^2(s)\Big)ds + \int_t^T \sigma(s)dW(s)\Big), \tag{2}$$

and

$$\mathbb{E}(S(T)|\mathcal{F}_t) = S(t)\, exp\Big(\int_t^T \mu(s)ds\Big). \tag{3}$$

Let K be the predicted value of $S(T)$ at time T. When the overestimation and underestimation is penalized equally, the quantity that represent the variable's central tendency, such as mean and median, is commonly used for prediction. But we consider a more general case where overestimation needs to penalized differently from underestimation, as discussed in Section 1. When the observed value is $S(T)$, the overestimated quantity becomes $\max\{K - S(T), 0\}$, while the underestimated quantity is $\max\{S(T) - K, 0\}$.

Let p_o and p_u denote the penalties for over/underestimation, respectively. We formulate the optimization problem for estimating $S(K)$ that can minimize the expected prediction cost,

$$\min_{K \in R^+} \mathbb{E}\Big[P_o \max\{K - S(T), 0\} + P_u \max\{S(T) - K, 0\}|\mathcal{F}_t\Big]. \tag{4}$$

Note that

$$\max\{K - S(T), 0\} = K - S(T) + \max\{S(T) - K, 0\}. \tag{5}$$

If we substitute (5) into (4), the optimal predicted value, denoted by K^*, can be obtained by solving the following objective function.

$$K(T)^* = argmin_{K \in R^+}\mathbb{E}\Big[(P_o + P_u)\max\{S(T) - K, 0\} + P_o(K - S(T))|\mathcal{F}_t\Big], \tag{6}$$

or equivalently,

$$K(T)^* = argmin_{K \in R^+}\mathbb{E}\Big[P_o\Big(\frac{P_o + P_u}{P_o}\max\{S(T)-K, 0\} + (K-S(T))\Big)|\mathcal{F}_t\Big]. \tag{7}$$

In the next section we will present a solution procedure to obtain $K^*(T)$, based on the option theory.

2.2. Real option based solution procedure

The optimization problem in (7) can be reformulated by employing the financial pricing theories. Suppose that we want to predict a state at the future time T. In pricing theories, T can be viewed as the date to maturity, or the expiration date.

A real option, with the date to maturity T, can be constructed on the state variable $S(t)$. A real option is a stochastic variable $\mathcal{X} \in \mathcal{F}_T^W$ that can be expressed as

$$\mathcal{X} = \Phi(S(T)), \tag{8}$$

where $\Phi(\cdot)$ is a function $\Phi(\cdot)$ is typically set to the payoff of the real option at time T.

When the predicted value is K, K can be viewed as the strike value in the option theory, while $\max\{S(T) - K, 0\}$ is the payoff. Therefore, we get

$$\Phi(S(T)) = \max\{S(T) - K, 0\} \tag{9}$$

It is required that $\mathcal{X} \in \mathcal{F}_T^W$ ensures that the value of the payoff of the real option $\mathcal{X}$ is determined at time T.

Let the price process $\Pi(t; \mathcal{X})$ for the real option at time t be given by a function $F(t, S(t)) \in [t, T] \times R_+$, i.e.,

$$\Pi(t; \mathcal{X}) = F(t, S(t)). \tag{10}$$

Here $F(\cdot)$ is a function which is assumed to be once continuously differentiable in t, and twice in $S(t)$.

For a short-term prediction, the time interval Δt between the current time t and the future time T is small, so we can assume that $\mu(t)$ and $\sigma(t)$ are constants during $[t, T]$. Then $F(t, S(t))$ can be obtained by solving the following Partial Differential Equation (PDE),

$$\frac{\partial F(t, S(t))}{\partial t} + \mu(t)S(t)\frac{\partial F(t, S(t))}{\partial S} + \frac{1}{2}S(t)^2\sigma^2(t)\frac{\partial F(t, S(t))}{\partial S^2} = 0 \tag{11}$$

with

$$F(T, S(T)) = \Phi(S(T)), \tag{12}$$

The PDE in (11)–(12) can be solved numerically. But alternatively, we solve it using the Feynman–Kač stochastic representation formula (Shreve, 2004), to obtain

$$F(t, S(t)) = \mathbb{E}_S\big[\Phi(S(T)) \mid \mathcal{F}_t\big]. \tag{13}$$

Next, we derive $F(t, S(t))$ in a closed form, given K. Letting $y = ln\big[S(T)/S(t)\big]$ and using the fact that $S(T) = S(t)exp((\mu(t) - \frac{1}{2}\sigma^2(t))\Delta t + \sigma(t)\Delta W(t))$, it follows that $y \sim \mathcal{N}\big((\mu(t) - \frac{1}{2}\sigma^2(t))\Delta t, \sigma^2(t)\Delta t\big)$. Thus, the probability density function $f(y)$ of y is given by

$$f(y) = \frac{1}{\sigma(t)\sqrt{2\pi\Delta t}}e^{-\left(\frac{(y - (\mu(t) - \frac{1}{2}\sigma^2(t))\Delta t)^2}{2\sigma(t)^2\Delta t}\right)}. \tag{14}$$

Consequently, we obtain

$$\mathbb{E}_S\big[\Phi(S(T)) \mid \mathcal{F}_t\big]$$
$$= \mathbb{E}_S\big[\max\{S(T) - K, 0\}|\mathcal{F}_t\big] \tag{15}$$
$$= \mathbb{E}_S\big[\max\{S(t)e^y - K, 0\}\big] \tag{16}$$
$$= \int_{ln\frac{K}{S(t)}}^{\infty} S(t)e^y f(y)dy - \int_{ln\frac{K}{S(t)}}^{\infty} Kf(y)dy \tag{17}$$

To solve (17), let I_1 and I_2, respectively, denote the first and second terms in (17). We also let $z = y - (\mu(t) - 0.5\sigma^2(t))\Delta t/\sigma(t)\sqrt{\Delta t}$. First, I_2 becomes

$$I_2 = \int_{ln\frac{K}{S(t)}}^{\infty} Kf(y)dy \tag{18}$$
$$= K\int_{-d2}^{\infty}\frac{1}{\sqrt{2\pi}}e^{\frac{-z^2}{2}}dz \tag{19}$$
$$= K\int_{-\infty}^{d2}\frac{1}{\sqrt{2\pi}}e^{\frac{-z^2}{2}}dz = K\mathcal{N}(d_2) \tag{20}$$

where $d_2 = ln(\frac{S(t)}{K}) + (\mu(t) - \frac{1}{2}\sigma^2(t))\Delta t/\sigma(t)\sqrt{\Delta t}$ and $\mathcal{N}(\cdot)$ denotes the cumulative distribution function (CFD) for the standard normal distribution. Next, we obtain I_1 as

$$I_1 = \int_{ln\frac{K}{S(t)}}^{\infty} S(t)e^y f(y)dy \tag{21}$$
$$= S(t)\int_{-d2}^{\infty}\frac{1}{\sqrt{2\pi}}e^{\frac{-z^2}{2} + z\sigma(t)\sqrt{\Delta t} + (\mu(t) - \frac{1}{2}\sigma^2(t))\Delta t}dz \tag{22}$$

$$=S(t)\int_{-d2}^{\infty}\frac{1}{\sqrt{2\pi}}e^{\frac{-1}{2}(z-\sigma(t)\sqrt{\Delta t})^2}e^{(\mu(t)\Delta t)}dz \tag{23}$$

$$=S(t)e^{\mu(t)\Delta t}\int_{-d_2-\sigma\sqrt{\Delta t}}^{\infty}\frac{1}{\sqrt{2\pi}}e^{\frac{-v^2}{2}}dv \tag{24}$$

$$=S(t)e^{\mu(t)\Delta t}\mathcal{N}(d_1) \tag{25}$$

where we use $v = z - \sigma(t)\sqrt{\Delta t}$ in (24) and $d_1 = d_2 + \sigma(t)\sqrt{\Delta t}$ in (25).

Then, $F(t, S(t))$ in (13) becomes:

$$F(t, S(t)) = e^{\mu(t)\Delta t}\,\mathcal{N}(d_1)S(t) - \mathcal{N}(d_2)\,K. \tag{26}$$

Note that given $\mu(t)$ and $S(t)$ at the current time t and K, we can obtain $F(t, S(t))$.

With the obtained expected payoff $\mathbb{E}_S\left[\Phi(S(T)) \mid \mathcal{F}_t\right]$ where $\Phi(S(T)) = \max\{S(T)-K, 0\}$, we can find the optimal $K^*(T)$ in (7). Let ω denote the ratio of overestimation cost to underestimation cost, i.e.,

$$\omega = \frac{P_u}{P_o} \tag{27}$$

Given the price of the real option, defined in (13), we can reformulate the optimization problem in (7) as

$$K^*(T) = argmin_{K\in R^+}\,\mathbb{E}\left[(1+\omega)\max\{S(T)-K, 0\} \mid \mathcal{F}_t\right] + \mathbb{E}\left[(K - S(T)) \mid \mathcal{F}_t\right] \tag{28}$$

$$=argmin_{K\in R^+}\left[(1+\omega)F(t, S(t)) + K - S(t)\,e^{\mu(t)\Delta t}\right] \tag{29}$$

$$=argmin_{K\in R^+}\left[(1+\omega)(e^{\mu(t)\Delta t}\,\mathcal{N}(d_1)S(t) - \mathcal{N}(d_2)\,K) + K - S(t)\,e^{\mu(t)\Delta t}\right] \tag{30}$$

with $d_2 = ln(\frac{S(t)}{K}) + (\mu(t) - \frac{1}{2}\sigma^2(t))\Delta t/\sigma(t)\sqrt{\Delta t}$ and $d_1 = d_2 + \sigma(t)\sqrt{\Delta t}$. We use (13) in the first term in the second equality and the last term in the second equality is obtained using (3). By plugging $F(t, S(t))$ in (26), we get the last equality.

The predictor K^* prefers overestimation when $\omega > 1$ or underestimation when $\omega < 1$. When overestimation and underestimation are equally penalized, the optimal K^* can be obtained with $w = 1$ in (28). It is worthwhile to mention that one way to see the advantage of using the proposed pricing method is to look at Equation (30). The solution of the pricing problem is implicitly a function of the volatility, making it more robust to changes in the intensity of the volatility. The optimization function in (28) is a convex optimization problem that can be solved efficiently by existing numerical optimization softwares. In our implementation, we use Scipy's (Scientific Python) optimization library in Python.

As a remark, Black–Scholes model is considered one of the most tractable framework for valuing options. The Black–Schole's pricing model requires a few assumptions including: the underlying asset is traded, follows GBM, and the contingent claim is replicable (Shreve, 2004). When any contingent claim is replicable, the market is referred to as complete. The market completeness is often determined through a set of market price of risk (Kaido and White, 2009). There is a single market price of risk when a market is complete. In valuing financial options, the probability space is usually transformed from P-space to the risk-neutral Q-space. The transformation helps determine the unique and no-arbitrage price of options.

In this paper, we take a slightly different approach to price over and underestimation options. In our approach, we formulate the underlying process of the stochastic variable as GBM. The close-form price of options is obtained by a PDE, which turns out to be similar to the Black–Schole's pricing model with risk-free interest rate equal to zero but under the P probability space. It should be noted that the drift term μ in our model appears in the derived PDE, emphasizing the pricing in the P-space (real-world space) which is important to reflect the prediction space. Therefore, our approach is valid as long as the underlying assumption follows GBM. Other assumptions in Black–Schole's model are not required. In Section 3, the Shapiro–Wilk normality test (Shapiro and Wilk, 1965) is performed to verify the GBM assumption.

2.3. Parameters estimation

For a volatile stochastic process, the parameters $\mu(t)$ and $\sigma(t)$ can be time-varying. We estimate the nonstationary parameters using recent observations. Consider the n most recent observations at the current time t, i.e., $S(t-(n-1)\Delta t), S(t-(n-2)\Delta t), \ldots, S(t)$. Because $S(t)$ follows geometric Brownian motion and $\mu(t)$ and $\sigma(t)$ are assumed to be constant during the short interval Δt, the discretization scheme of (2) is given by

$$\ln\left(\frac{S(t+\Delta t)}{S(t)}\right) = \left(\mu(t) - \frac{1}{2}\sigma^2(t)\right)\Delta t + \sigma(t)\big(W(t+\Delta t) - W(t)\big). \tag{31}$$

Noting that under GBM $\ln\big(S(t+\Delta t)/S(t)\big)$ is normally distributed with mean $\left[\mu(t) - \frac{1}{2}\sigma^2(t)\right]\Delta t$ and variance $\sigma^2(t)$, we estimate $\mu(t)$ and $\sigma(t)$ using maximum likelihood method as

$$\hat{\sigma}(t) = \left(\frac{1}{n}\sum_{i=2}^{n}\left(\ln\left(\frac{S(t-(n-i)\Delta t)}{S(t-(n-i+1)\Delta t)}\right) - \frac{1}{n}\sum_{i=1}^{n-1}\left[\ln\left(\frac{S(t-(n-i)\Delta t)}{S(t-(n-i+1)\Delta t)}\right)\right]\right)^2\right)^{\frac{1}{2}}, \tag{32}$$

$$\hat{\mu}(t) = \frac{1}{n}\sum_{i=1}^{n}\left[\ln\left(\frac{S(t-(n-i)\Delta t)}{S(t-(n-i+1)\Delta t)}\right)\right] + \frac{1}{2}\hat{\sigma}(t)^2, \tag{33}$$

respectively.

The estimated parameters $\hat{\mu}(t)$ and $\hat{\sigma}(t)$ are plugged into (26) and we obtain the optimal predicted value K^* for $S(t+\Delta t)$ by solving (30).

2.4. Implementation details

We refer our proposed model to as the *option prediction model*. Fig. 1 summarizes the overall procedure of the proposed approach. We also summarize the procedure of the proposed approach in Algorithm 1 below. We set the time step $\Delta t = 1$ to make the one-ahead step prediction. The data is divided into three sets: training, validation, and testing. The training set starts at $t = 1$ and ends at $t = N_1$, consisting of about 50% of the entire data set, is used to determine the model parameters as shown in Fig. 1. The validation set, consisting of about 20% of the data set, is used for determining the window size n. Lastly the testing set consists of the last 30% of the data set and it starts at $t = N_2$ and ends at N_3.

Algorithm 1 Option prediction model

1: **Initialization:**
2: Choose a window size n by the validation technique as shown in Fig. 1.
3: Obtain initial estimates for the model parameters $\hat{\sigma}(N_2)$ and $\hat{\mu}(N_2)$ in (32) and (33), respectively.
4: Determine $F(N_2, S(N_2))$ in (13).
5: **for** $k = N_2 + 1$ to N_3 **do**
6: **Prediction:**
7: Obtain K^* by solving (30) to obtain the one-step ahead state prediction.
8: **Update:**
9: Observe $S(k)$.
10: Obtain $\hat{\sigma}(k)$ and $\hat{\mu}(k)$ in (32) and (33), respectively, by using n recent observations.
11: Determine $F(k, S(k))$ in (13).
12: **end for**

In Algorithm 1, we determine the window size n for obtaining the parameters $\hat{\mu}(t)$ and $\hat{\sigma}(t)$ using the validation technique (Friedman et al., 2009). Specifically we fit the model with a different window size n and

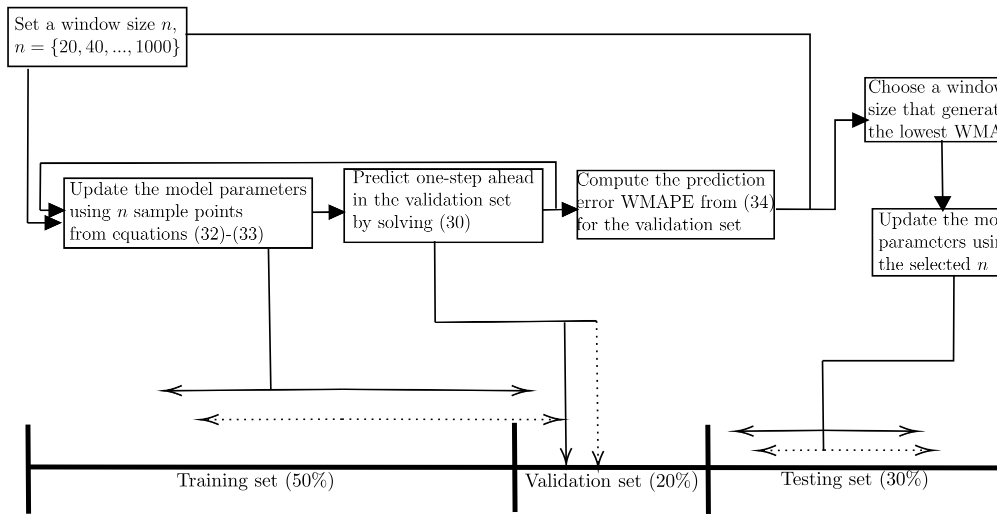

Fig. 1. Overall procedure of the proposed approach (the dotted lines imply that the model parameters are updated in a rolling-horizon manner using the most n recent observations).

evaluate the prediction performance using data in the validation set and choose the best window size that generates the lowest prediction error in the validation set. The performance of our approach is evaluated using data in the testing set (See Fig. 1). We report the prediction performance in the testing set in Section 3.

In evaluating the prediction performance, we consider that the overestimated and underestimated prediction results need to be evaluated differently for $\omega \neq 1$. As such we employ the following two performance measures, namely, Weighted Mean Absolute Percentage Error (WMAPE), defined by

$$\text{WMAPE} = \frac{1}{N}\sum_{t=1}^{N}\Big(\frac{\mathbb{1}_{(S(t)>K^*(t))}\omega|S(t)-K^*(t)|}{S(t)} + \frac{\mathbb{1}_{(S(t)<K^*(t))}|S(t)-K^*(t)|}{S(t)}\Big), \tag{34}$$

where N denotes the number of data points in the testing set and $K^*(t)$ is the predicted value at time t.

3. Case studies

This section implements the proposed prediction model using multiple datasets obtained from real-life applications. Specifically we examine the performance of the predictive model in predicting the size of a manufacturer's order, and a stock market index price.

3.1. Alternative methods

We compare our model with two standard time series models, namely, the ARIMA and the ARIMA-GARCH. We use the Akaika Information Criteria (AIC) to select the model order in both models. For fair comparison, we update the model parameters in a rolling horizon manner, similar to the procedure discussed in Section 2.4. That is, we determine the window size n using the validation technique and update the model parameters using the most recent n observations whenever a new observation is obtained.

With underestimation penalties, Pourhabib et al. (2016) suggest using quantile of the predictive state density. Specifically, with $\omega(= p_u/p_o)$ denoting the ratio of underestimation cost to overestimation cost, we use the $(\omega/1+\omega)$-quantile, given by

$$\text{Quantile prediction} = \hat{\mu}_a(t) + \hat{\sigma}_a(t)\mathcal{N}^{-1}\Big(\frac{\omega}{1+\omega}\Big), \tag{35}$$

where $\hat{\mu}_a(t)$ denotes the estimated predicted mean, $\hat{\sigma}_a$ is the estimated standard deviation in ARIMA (or ARIMA-GARCH) model, and $\mathcal{N}^{-1}(\cdot)$ denotes the inverse of the standard normal CDF. Note that large (small) w puts more penalty on p_u (p_o) and the quantile prediction provides a larger (smaller) predicted value, so underestimation (overestimation) can be avoided.

3.2. Manufacturing data

We first study the prediction problem faced by our industry partner, CM. The historical data obtained from CM includes orders of 10 different types of bumper beams. We use monthly data on those 10 types of bumper beams ordered over a period of 29 consecutive months (the order size varies from 0 to over 36,000 items). As indicated in Section 2.3, $\ln\Big(S(t+\Delta t)/S(t)\Big)$ is assumed to be normally distributed with mean $\big[\mu(t) - \frac{1}{2}\sigma^2(t)\big]\Delta t$ and variance $\sigma^2(t)$. The Shapiro–Wilk normality test (Shapiro and Wilk, 1965) is used to examine the normality and demonstrate the validity of the GBM assumption in the presented datasets. The test statistic of the ten products ranges between 0.924 and 0.967. Its associated p-value is between 0.287 and 0.859, indicating that we fail to reject the null hypothesis that $\ln\Big(S(t+\Delta t)/S(t)\Big)$ is normal.

When applying the proposed model to this problem, the choice of weight ω affects the final prediction. By changing the weight, we are able to show a preference for over-capacity (overestimation) or under-capacity (underestimation). We consider different cases for choosing the weight parameter ω.

Let us first look at the case when ω is set to be less than one (i.e, $p_u \leq p_o$). According to CM, workers and equipment can be shifted from one type of bumper beam to another, but doing so incurs 15% loss of production efficiency. In other words, if one type of bumper beam is overestimated, causing over-capacity, available resources can be assigned to other bumper beam production, but with a reduced efficiency. In this case, underestimation is favored and we set $w = 1/1.15$.

Next the weight parameter can be set to be greater than 1 (i.e, $p_u \geq p_o$) when the prediction is preferred to be more than the actual order size. According to the labor law in Michigan in the U.S., overtime rate is higher than the regular salary. In this case we set $w = 1.15$ to emphasize the preference of overestimation to underestimation. Finally we also consider $w = 1$, which reflects equal penalties.

The errors in terms of WMAPE for all ten types of bumper beams are presented in Tables 1–3 with three different weights. Overall our option prediction model performs better than the CM's own prediction, ARIMA and ARIMA-GARCH in both criteria. With $w = 1/1.15$ the proposed approach provides lower WMAPEs for 5 types of bumper beams out

Table 1
CM Prediction Results for ten types of bumper beams with $\omega = 1/1.15$ in the Testing Set (The values in bold indicate the lowest prediction error for each product).

Weighted Mean Absolute Percent Error (WMAPE)

Product no.	Option prediction	ARIMA	ARIMA-GARCH	CM prediction
1	**0.19**	1.14	0.94	0.52
2	0.43	0.62	0.42	**0.40**
3	0.44	28.89	**0.41**	0.69
4	0.69	1.20	3.92	**0.66**
5	**0.11**	0.12	0.12	0.15
6	**5.30**	45.23	11.01	7.93
7	**3.66**	215.17	408.98	9.63
8	0.26	**0.24**	0.76	0.60
9	0.22	**0.19**	0.66	0.58
10	**0.20**	0.36	2.63	0.99

Table 2
CM Prediction Results for ten types of bumper beams with $\omega = 1$ in the Testing Set (The values in bold indicate the lowest prediction error for each product).

Weighted Mean Absolute Percent Error (WMAPE)

Product no.	Option prediction	ARIMA	ARIMA-GARCH	CM Prediction
1	**0.24**	1.09	0.94	0.59
2	0.45	0.61	**0.40**	0.44
3	0.50	25.76	**0.39**	0.79
4	**0.72**	1.05	3.41	0.75
5	0.12	0.11	**0.11**	0.17
6	**6.17**	39.61	9.59	7.95
7	**3.87**	187.36	356.23	9.67
8	0.28	**0.22**	0.66	0.60
9	0.23	**0.18**	0.58	0.58
10	**0.21**	0.31	2.29	0.99

Table 3
CM Prediction Results for ten types of bumper beams with $\omega = 1.15$ in the Testing Set (The values in bold indicate the lowest prediction error for each product).

Weighted Mean Absolute Percent Error (WMAPE)

Product no.	Option prediction	ARIMA	ARIMA-GARCH	CM prediction
1	**0.19**	1.14	0.94	0.52
2	0.43	0.62	0.42	**0.40**
3	0.44	28.89	**0.41**	0.69
4	0.69	1.20	3.92	**0.66**
5	**0.11**	0.12	0.12	0.15
6	**5.30**	45.23	11.01	7.93
7	**3.66**	215.17	408.98	9.63
8	0.26	**0.24**	0.76	0.60
9	0.22	**0.19**	0.66	0.58
10	**0.20**	0.36	2.63	0.99

of 10 types. Similarly, with other w values, our approach outperforms the alternative models in most cases.

Although ARIMA and ARIMA-GARCH provide the lowest errors for some products, their prediction performance is not consistent. For example, for 1st, 4th and 7th product, WMAPEs from ARMA are much higher than the proposed approach, whereas ARIMA-GARCH results in pretty poor performance for predicting order sizes for $7\text{th} - 10\text{th}$ products. On the contrary, our approach provides more stable results. Even when WMAPEs from our approach are higher than other approaches, they are close to the lowest errors. Therefore, we can conclude that our approach is more accurate and reliable. The CM's proprietary model does not account for unequal weights on overestimation and underestimation. If the company wants to minimize the excess inventory due to overestimation, a small (less than 1) weight parameter should be assigned. If the company goal is to meet customer satisfaction, overestimation should be preferred with a large (larger than 1) weight parameter. In this sense our approach can reflect the company's management preference more flexibly.

In our approach, it is assumed that the stochastic process is stationary during $\Delta t = [t, T]$, implying that its statistical properties, characterized by $\mu(t)$ and $\sigma(t)$, do not change, while $\mu(t)$ and $\sigma(t)$ are allowed to change every period. In other words, the period by period volatility is captured by varying $\mu(t)$ and $\sigma(t)$, whereas the inter-period (or within-period) volatility is captured by $\sigma(t)$. Therefore, the period duration Δ needs to be defined to satisfy such assumption. In this case study, automotive companies make orders monthly and the underlying demand process does not dramatically change within a short period of time (here, a month), although it can change month by month. For that reason, we believe that the stationarity assumption during each month is satisfied in this CM case study.

3.3. Stock market index data

To evaluate the performance of our approach in a highly volatile process, we consider stock market index price time series data. We analyze the daily closing price of the Dow Jones index in three time periods between 2010 and 2015. The Shapiro–Wilk normality test is employed to verify the GBM assumption in this case study. The test statistic of the three different time periods is between 0.941 and 0.985 and its associated p-value is between 0.103 and 0.432, concluding that we fail to reject the null hypothesis that $\ln\left(S(t+\Delta t)/S(t)\right)$ is normal.

Risk averse and risk seeking investors have different preferences in terms of overestimation and underestimation. That being said, in a bull market, stock prices are expected to increase. In such a case, risk seeking investors with aggressive investment strategies would prefer biasing their prediction to overestimation. On the contrary, risk averse investors tend to be less optimistic, making them conservative, preferring underestimation. To reflect different investment preferences, we consider three values of the weight parameter ω, $1/1.15$, 1, or 1.15, to represent the underestimation preference, neutral/no preference, and overestimation preference, respectively.

Table 4 summarizes the results with three testing periods. Each testing period includes 100 days. Clearly, our option prediction performs better than ARIMA-GARCH and ARIMA in all cases, alerting for the possibility of a profitable trading strategy. The ARMA and ARMA-GARCH models generate 2 to 11 times higher WMAPEs.

4. Conclusion

In this study, we present a new prediction methodology for the time series data, based on option theories in finance when the underlying dynamics is assumed to follow the GBM process. To characterize time-varying patterns, we allow the GBM model parameters to vary over time and update the parameter values using recent observations. We formulate the prediction problem with unequal overestimation and underestimation penalties as the stochastic optimization problem and provide its solution procedure. We demonstrate the prediction capability of the proposed approach in various applications. Our approach appears to work well in the manufacturing application, when the order size varies over time. For more highly volatile processes such as stock prices, the proposed model exhibits much stronger prediction capability, compared to alternative ARIMA-based models.

In the future, we plan to investigate other parameter updating schemes. In this study, we update parameters in a rolling horizon manner using the maximum likelihood estimations. Another possibility is to use the Kalman filtering or its variants. Long-term (or multi-step ahead) predictions are beyond the scope of this study, but we plan to extend the approach presented in this study for obtaining accurate long-term predictions. We will also incorporate prediction results into managerial decision-making in several applications such as power grid operation with renewable energy (Bouffard and Galiana, 2008).

In economics, different factors potentially affect the product price and quantity where the relationship between supply and demand could affect the forecasting mechanism. Because this study is solely concerned with univariate analysis, we do not make a distinction between whether a stochastic variable is a price or quantity and the demand-supply relationship is not taken into consideration. However, our framework can be extended to include other affecting factors, which is a subject of our ongoing work (Wang et al., 2018).

Table 4
Dow Jones index price prediction results in the testing set (The values in bold indicate the lowest prediction error for each testing period and weight).

Testing period	Weight (ω)	Method	WMAPE
Oct 2010–Mar 2011	1/1.15	Option Prediction	**0.0043**
		ARIMA-GARCH	0.04998
		ARIMA	0.0615
	1	Option Prediction	**0.0046**
		ARIMA-GARCH	0.0499
		ARIMA	0.0541
	1.15	Option Prediction	**0.0050**
		ARIMA-GARCH	0.0572
		ARIMA	0.0548
Aug 2013–Dec 2013	1/1.15	Option Prediction	**0.0046**
		ARIMA-GARCH	0.0237
		ARIMA	0.0309
	1	Option Prediction	**0.0049**
		ARIMA-GARCH	0.0209
		ARIMA	0.0276
	1.15	Option Prediction	**0.0052**
		ARIMA-GARCH	0.0211
		ARIMA	0.0284
Oct 2014–Mar 2015	1/1.15	Option Prediction	**0.0056**
		ARIMA-GARCH	0.0195
		ARIMA	0.0109
	1	Option Prediction	**0.0060**
		ARIMA-GARCH	0.0184
		ARIMA	0.0099
	1.15	Option Prediction	**0.0065**
		ARIMA-GARCH	0.0200
		ARIMA	0.0104

CRediT authorship contribution statement

Abdullah AlShelahi: Conceptualization, Methodology, Software, Formal analysis, Writing - original draft, Writing - review & editing. **Jingxing Wang:** Conceptualization, Methodology, Formal analysis. **Mingdi You:** Data curation, Validation. **Eunshin Byon:** Conceptualization, Methodology, Writing - review & editing, Funding acquisition, Supervision. **Romesh Saigal:** Conceptualization, Methodology, Writing - review & editing, Supervision, Resources.

References

Bengtsson, J., 2001. Manufacturing flexibility and real options: a review. Int. J. Prod. Econ. 74 (1–3), 213–224.
Benninga, S., Tolkowsky, E., 2002. Real options—an introduction and an application to R&D valuation. Eng. Econ. 47 (2), 151–168.
Bernanke, B.S., 1983. Irreversibility, uncertainty, and cyclical investment. Q. J. Econ. 98 (1), 85–106.
Björk, T., 2009. Arbitrage Theory in Continuous Time. Oxford university press.
Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. J. Political Econ. 81 (3), 637–654.
Boomsma, T.K., Meade, N., Fleten, S.-E., 2012. Renewable energy investments under different support schemes: a real options approach. European J. Oper. Res. 220 (1), 225–237.
Bouffard, F., Galiana, F.D., 2008. Stochastic security for operations planning with significant wind power generation. IEEE Trans. Power Syst. 23 (2), 306–316.
Box, G.E., Jenkins, G.M., Reinsel, G.C., Ljung, G.M., 2015. Time Series Analysis: Forecasting and Control. John Wiley & Sons, pp. 1–2.
Brooks, C., 2002. Introductory Econometrics for Finance. Cambridge university press, pp. 206–207.
Chatfield, C., 2000. Time-series Forecasting. CRC Press.
Chiu, C.-H., Hou, S.-H., Li, X., Liu, W., 2017. Real options approach for fashionable and perishable products using stock loan with regime switching. Ann. Oper. Res. 257 (1–2), 357–377.
Datar, V.T., Mathews, S.H., 2004. European real options: an intuitive algorithm for the black-scholes formula. J. Appl. Finance 14 (1).
Friedman, J., Hastie, T., Tibshirani, R., 2009. The elements of statistical learning: data mining, inference, and prediction, second ed. In: Springer Series in Statistics, Springer.
Gardiner, C., 1986. Handbook of stochastic methods for physics, chemistry and the natural sciences. Appl. Opt. 25, 3145.
Hahn, W.J., Dyer, J.S., 2008. Discrete time modeling of mean-reverting stochastic processes for real option valuation. European J. Oper. Res. 184 (2), 534–548.
Hahn, H., Meyer-Nieberg, S., Pickl, S., 2009. Electric load forecasting methods: tools for decision making. European J. Oper. Res. 199 (3), 902–907.
Kaido, H., White, H., 2009. Inference on risk-neutral measures for incomplete markets. J. Financ. Econ. 7 (3), 199–246.
Kantz, H., Schreiber, T., 2004. Nonlinear Time Series Analysis, vol. 7. Cambridge university press.
Ledolter, J., 1981. Recursive estimation and adaptive forecasting in ARIMA models with time varying coefficients. In: Applied Time Series Analysis II. Elsevier, pp. 449–471.
Lu, X.F., Lai, K.K., Liang, L., 2014. Portfolio value-at-risk estimation in energy futures markets with time-varying copula-GARCH model. Ann. Oper. Res. 219 (1), 333–357.
Merton, R.C., et al., 1973. Theory of rational option pricing. In: Theory of Valuation. pp. 229–288.
Nembhard, H.B., Shi, L., Aktan, M., 2002. A real options design for quality control charts. Eng. Econ. 47 (1), 28–59.
Pourhabib, A., Huang, J.Z., Ding, Y., 2016. Short-term wind speed forecast using measurements from multiple turbines in a wind farm. Technometrics 58 (1), 138–147.
Ruppert, D., 2015. Statistics and Data Analysis for Financial Engineering. Springer Texts in Statistics.
Segal, G., Shaliastovich, I., Yaron, A., 2015. Good and bad uncertainty: macroeconomic and financial market implications. J. Financ. Econ. 117 (2), 369–397.
Shapiro, S.S., Wilk, M.B., 1965. An analysis of variance test for normality (complete samples). Biometrika 52 (3/4), 591–611.
Shreve, S.E., 2004. Stochastic Calculus for Finance II: Continuous-time Models, vol. 11. Springer Science & Business Media.
Sohn, S.Y., Lim, M., 2007. Hierarchical forecasting based on AR-GARCH model in a coherent structure. European J. Oper. Res. 176 (2), 1033–1040.
Thorsen, B.J., 1999. Afforestation as a real option: some policy implications. For. Sci. 45 (2), 171–178.
Tran, N., Reed, D.A., 2004. Automatic ARIMA time series modeling for adaptive I/O prefetching. IEEE Trans. Parallel Distrib. Syst. 15 (4), 362–377.
Trigeorgis, L., et al., 1996. Real Options: Managerial Flexibility and Strategy in Resource Allocation. MIT press.
Wang, J., Alshelahi, A., You, M., Byon, E., Saigal, R., 2018. Integrative probabilistic short-term prediction and uncertainty quantification of wind power generation. arXiv preprint arXiv:1808.07347.
Whitt, W., 1981. The stationary distribution of a stochastic clearing process. Oper. Res. 29 (2), 294–308.
Wu, L.-L.B., Wu, D., 2015. Dynamic pricing and risk analytics under competition and stochastic reference price effects. IEEE Trans. Ind. Inf. 12 (3), 1282–1293.
Xiao, S., Ma, S.-H., Li, G., Mukhopadhyay, S.K., 2015. European option pricing with a fast fourier transform algorithm for big data analysis. IEEE Trans. Ind. Inf. 12 (3), 1219–1231.
Zhai, Q., Ye, Z.-S., 2017. Rul prediction of deteriorating products using an adaptive Wiener process model. IEEE Trans. Ind. Inf. 13 (6), 2911–2921.
Zhang, G.P., 2003. Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing 50, 159–175.