mathematics

Article

Validation of Stock Price Prediction Models in the Conditions of Financial Crisis

Vesela Mihova [1,*], Ivan Georgiev [1,2], Elitsa Raeva [1], Slavi Georgiev [1,2] and Velizar Pavlov [1]

1 Department of Applied Mathematics and Statistics, University of Ruse, 8 Studentska Str., 7004 Ruse, Bulgaria; irgeorgiev@uni-ruse.bg or irgeorgiev@math.bas.bg (I.G.); eraeva@uni-ruse.bg (E.R.); sggeorgiev@uni-ruse.bg or sggeorgiev@math.bas.bg (S.G.); vpavlov@uni-ruse.bg (V.P.)

2 Department of Informational Modeling, Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, 8 Acad. Georgi Bonchev Str., 1113 Sofia, Bulgaria

* Correspondence: vmicheva@uni-ruse.bg

Abstract: The distribution laws of various natural and anthropogenic processes in the world around us are stochastic in nature. The development of mathematics and, in particular, of stochastic modeling allows us to study regularities in such processes. In practice, stochastic modeling finds a huge number of applications in various fields, including finance and economics. In this work, some particular applications of stochastic processes in finance are examined in the conditions of financial crisis, aiming to provide a solid approach for stock price forecasting. More specifically, autoregressive integrated moving average (ARIMA) models and modified ordinary differential equation (ODE) models, previously developed by some of the authors to predict the asset prices of four Bulgarian companies, are validated against a time period during the crisis. Estimated rates of return are calculated from the models for one period ahead. The errors are estimated and the models are compared. The return values predicted with each of the two approaches are used to derive optimal risk portfolios based on the Markowitz model, which is the second major aim of this study. The third aim is to compare the resulting portfolios in terms of distribution (i.e., weights of the stocks), risk, and rate of return.

Keywords: stochastic processes; time series; ARIMA; ODE; price forecasting

MSC: 60G35; 91B84; 34A34; 34A55; 91B28; 90C90

Citation: Mihova, V.; Georgiev, I.; Raeva, E.; Georgiev, S.; Pavlov, V. Validation of Stock Price Prediction Models in the Conditions of Financial Crisis. *Mathematics* **2024**, *12*, 33. https://doi.org/10.3390/math12010033

Academic Editors: Anatoliy Swishchuk and Antonella Basso

Received: 26 October 2023
Revised: 16 November 2023
Accepted: 18 December 2023
Published: 22 December 2023

1. Introduction

Financial markets have always attracted considerable attention due to their critical influence on developed countries' economic infrastructure, and also because of their facilitation of huge volumes of capital flow. Functioning as complex systems for the trading of various financial instruments—both domestically and internationally—these markets serve as a mainstay for resource allocation and liquidity provision in the broader economic landscape. In contemporary economic theory, financial markets are accredited with six fundamental roles: price determination, liquidity provision, transactional efficiency, facilitation of credit and lending mechanisms, real-time information dissemination on cash flows, and risk management strategies.

To cater to the diverse investment preferences of market participants, various types of financial markets exist. These range from foreign exchange markets, which are arguably the largest of their kind, to capital markets trading in shares, bonds, and ETFs. Additionally, money markets largely concern short-term financial instruments and are typically dominated by banking institutions. Other categories include derivatives markets, commodities markets dealing in raw materials, the nascent cryptocurrency markets, mortgage markets focused on real-estate investments, and insurance markets that allow for risk transference through premiums. Each of these markets has its own temporal focus—some being more suitable for short-term investments, while others are geared toward long-term holdings.

In Bulgaria, where the capital market is yet to reach full maturity, stock price prediction has increasingly become an imperative subject of financial analysis, not just for professional investors but also for the general audience. Within stock markets, the necessity for robust investor education is emphasized due to the market's inherent volatility and sensitivity to a multitude of variables. Indeed, the dynamics of the market, coupled with various external and internal factors, contribute to rapid fluctuations in the price of traded assets. Publicly available historical data often exhibit predictive correlations with future stock returns; the challenge lies in accurately identifying and utilizing these data for predictive modeling. This challenge is even greater in the conditions of a financial crisis, when there is the worsening dynamics of macroeconomic indicators and increased volatility of financial markets.

In the area of portfolio management theory [1], the principle of separation postulates that given identical input variables, all investors would arrive at the same optimal risk portfolio. This separates the challenge of portfolio selection into two problems: The first is essentially a technical problem that involves the construction of an optimal risk portfolio, which is designed to maximize the reward-to-variability ratio based on input variables like rates of return, standard deviations, and correlation matrices. The second problem necessitates the proportional allocation of risk-free treasury bonds and the aforementioned risk portfolio, contingent upon the individual risk tolerances and preferences of each investor.

In the present research, we delve into specific applications of stochastic processes within the scope of finance, particularly during periods of financial turmoil. Our focus is directed towards the validation of two distinct models: the autoregressive integrated moving average (ARIMA), and the modified ordinary differential equation (ODE) model. These models, previously devised by some of the authors (see [2]) to forecast asset prices for four prominent Bulgarian companies, were applied over a timeline encompassing the financial crisis.

To establish the models' predictive accuracy, we computed the estimated rates of return for one period ahead. Subsequent to this, the error's magnitudes [3] were quantified and a comparative analysis between the models was conducted. The forecasted return values procured from both methodologies were instrumental in formulating optimal risk portfolios, grounded in the principles of the Markowitz model. A meticulous examination of these portfolios was undertaken, considering variables such as asset distribution, inherent risk, and estimated rate of return.

2. Materials and Methods

The investment portfolio constructed in the current study comprises a combination of shares from four Bulgarian companies operating in distinct sectors, namely, technology, real estate, courier/transport services, and finance. These companies include the technology holding Allterco JSC, the joint-stock company Elana AgroCredit JSC, the courier company Speedy JSC, and the holding Chimimport JSC. Data on a daily basis for the period 1 June 2020–29 October 2020 were used for the development of the models, and daily data for 1 June 2022–28 October 2022 were used for their validation.

Allterco JSC—Sofia Allterco Robotics is a well-established firm in the telecommunications and smart technology domain, having commenced operations in 2010 in Sofia, Bulgaria. It has gained recognition as an innovative enterprise focusing on the development and trade of IoT (Internet of Things) devices. Notable products include the widely popular MyKi smartwatches for children and the Shelly technology tailored for household applications. The company's shares were officially listed on the Bulgarian Stock Exchange (BSE) on 1 December 2016

Elana AgroCredit JSC—Sofia REIT (Real Estate Investment Trusts) provides agricultural loans to farmers for the acquisition of farming land, facilities, and advanced technologies. It stands as the pioneering entity specialized in financial leasing. The company's shares have been publicly traded on the BSE since the latter part of 2013.

Speedy JSC, established in 1998, has emerged as a prominent courier enterprise in Bulgaria. Its strategic partnership with DPD, Europe's largest ground delivery network, has significantly contributed to its successful global deliveries. Currently holding a substantial market share of 25.5%, Speedy became the first company in its industry to go public when it was listed on the BSE in November 2012.

Chimimport JSC, established in 1947 as an external holding, focuses on the chemical products market. The conglomerate includes over 68 successful companies renowned in various sectors, such as banking services, insurance, pension fund management, and mutual fund management, as well as receivables and real-estate securitization. Chimimport JSC is dedicated to providing comprehensive financial services and a diverse range of operations to cater to its clientele.

Table 1 shows the financial results of the four companies at the end of 2022.

Table 1. Financial results of the four companies at the end of 2022.

Company	Market Capital (Million BGN)	Turnover for the Previous Year (BGN)	Total Assets (Pieces)	Price per Share (BGN)
Allterco JSC	743,698,934	20,757,003	17,999,999	41.20
Elana AgroCredit JSC	37,362,524	193,821	36,629,925	0.95
Speedy JSC	548,517,138	52,109	5,377,619	97.18
Chimimport JSC	189,320,551	540,312	239,646,267	0.79

In this study, we address the challenge of constructing an investment portfolio comprising stocks from four companies and treasury bills. The planning horizon for this portfolio is set for the next period. The projections made are specifically related to the returns expected over one ownership period.

We considered the asset prices on 29 October 2020 (and similarly, on 28 October 2022, the second-last business day of October) as the initial timepoint (time 0) prices. The anticipated rates of return, denoted as $E(r_i)$, were determined following the methodology outlined in Equation (1):

$$E(r_i) = \frac{E\left(P_i^1\right) - P_i^0}{P_i^0} \quad (1)$$

where P_i^0 represents the prices at time 0, while $E\left(P_i^1\right)$ represents the estimated prices from the obtained models at time 1 (30 October 2020 and 31 October 2022—the last working day of October).

2.1. ARIMA Models

In contemporary financial analytics, autoregressive integrated moving average (ARIMA) models have been extensively harnessed to extrapolate potential trajectories for stock prices, anchored primarily on their historical performance, or to project a company's future earnings by scrutinizing previous fiscal intervals. These analytical structures are anchored in the broader spectrum of regression studies, proficiently elucidating the relative potency of a chosen dependent variable in contrast to other evolving determinants. Fundamentally, the ARIMA methodology aspires to forecast impending shifts in securities or broader financial market trajectories. Intriguingly, this is accomplished not by direct examination of absolute values but, rather, by analyzing the variances between consecutive data points in a series.

The construction of ARIMA models pivots around three parameters, colloquially denoted as p, d, and q [4]. The autoregressive component, symbolized by "p", encapsulates the influence exerted by data from "p" antecedent intervals within the analytical framework. Conversely, the integrated component, represented by "d", captures the overarching trend manifest in the dataset. The moving-average segment, denoted by "q", delineates the number of sequential data points that are leveraged to temper minor oscillations using a moving-average methodology.

To encapsulate the theoretical construct, an ARIMA model with the specified parameters p, d, and q, as cited in references [4,5], can be mathematically represented as per Equation (2):

$$Y_t = C + \varphi_1 \Delta^d Y_{t-1} + \cdots + \varphi_p \Delta^d Y_{t-p} - \theta_1 \varepsilon_{t-1} - \cdots \theta_q \varepsilon_{t-q} + \varepsilon_t. \tag{2}$$

where C, φ_i, $i = \overline{1,p}$, and θ_j, $j = \overline{1,q}$ are the parameters sought, while ε_j is a randomly distributed variable with mathematical expectation of zero and dispersion σ^2. If we do not have any information about the distribution of this variable, then by default it is assumed that the distribution is normal [5]. Δ is the difference operator, defined as (3):

$$\Delta^0 Y_t = Y_t, \Delta^1 Y_t = Y_t - Y_{t-1}, \ldots, \Delta^k Y_t = \Delta^{k-1} Y_t - \Delta^{k-1} Y_{t-1}. \tag{3}$$

In this study, we experimented with diverse permutations of the parameters (p, d, q), with the intent of meticulously characterizing the underlying time-series dynamics. The graphical representations of autocorrelation functions (ACFs) and partial autocorrelation functions (PACFs) for each unique parameter combination were reviewed. These functions are predicated upon a predetermined number of lags and are systematically computed for every temporal moment "t", barring certain terminal instances where their derivation proves infeasible.

Critical to our assessment is the scrutiny of discontinuities or "jumps" manifested within both the ACFs and PACFs. These fluctuations serve as invaluable markers, guiding the optimal selection of parameters for each model. It is noteworthy that if for a given Y_t both the ACF and PACF portray either an absence of jumps or a singular, marginal deviation beyond the confines of the 95% confidence intervals, such a model is adjudged to be adequately robust and congruent with the research objectives delineated in this study.

The formula for the autocorrelation function (ACF) in the current moment "t" for a "k" lag can be articulated according to Equation (4):

$$r_k = \frac{\frac{1}{n-k}\sum_{t=1}^{n-k} \left(Y_t - \overline{Y}\right)\left(Y_{t-k} - \overline{Y}\right)}{\frac{1}{n-1}\sum_{t=1}^{n} \left(Y_t - \overline{Y}\right)^2}. \tag{4}$$

In the given context, n denotes the total count of observations present within the time series, while k signifies the lag, quantified in terms of the number of time delays. $\overline{Y}$ stands as the arithmetic mean of the entire time series. The denominator encapsulates the variance inherent to the time series. The canonical error associated with autocorrelation is derived from the squared magnitude of the autocorrelation spanning all preceding autocorrelations.

The mathematical representations for ascertaining partial correlations are intrinsically intricate, necessitating the application of recursive methodologies [6].

2.2. Modified ODE Models

In those cases where extensive datasets are utilized, the application of more complex forecasting techniques based on numerical solutions for ordinary (ODE), partial (PDE), and stochastic (SDE) differential equations becomes imperative [7–9].

This article uses a modified ODE methodology for simulating the price fluctuations of a specific asset over a defined period. The benefits of this adaptation are illustrated through numerical examinations in [2] and references therein. The numerical price prediction models in this study are grounded on solving the Cauchy initial value problem for first-order ordinary differential equations (ODEs). The performance characteristics of the adapted ODE were examined through MATLAB 2020a [10]. These computational experiments encompass various data-fitting models to ascertain the best forecasted values. These are calculated as a weighted mean of all forecasts for the relevant financial instruments, where the weights are inversely proportional to the "final errors".

Consider the temporal moments t_i $\left(i = \overline{0, n}\right)$, sorted in ascending order, along with the observed asset prices y_i $\left(i = \overline{0, n}\right)$, represented as a time series, as indicated in (5):

$$t_0,\ t_1, \ldots,\ t_n \textit{ and } y_0,\ y_1,\ \ldots,\ y_n. \tag{5}$$

As demonstrated in prior studies [7,8], it is feasible to calibrate an ordinary differential equation to this particular time series.

$$y'(t) = g(t, y),\ \ y(t_0) = y_0, \tag{6}$$

describing the discrete timepoints of the time-series values provided. The function g(t, y) can be significantly diverse.

In the scenario where $g(t, y) = ay$, Equation (6) can be addressed either through numerical integration or by acquiring an analytic solution, provided that the value of a is known. Furthermore, it is plausible to calibrate the variable a at various temporal intervals. One method for resolving Equation (6), with $g(t, y) = ay$, is delineated by Lascsáková in [8]. Another approach, presented by Xue and Lai [7], suggests several alternatives in the structure of the first derivative in (6). They suggest the subsequent form for $g(t,\ y)$:

$$g(t, y) = a(t)y + s(y), \tag{7}$$

where $a(t)$ and $s(y)$ are representable by elementary functions such as polynomial functions, exponential functions, logarithmic functions, etc. They could be expanded to include series of degrees subject to specific conditions. Furthermore, within the dataset (5), discernible periodic trends can be identified. Consequently, the function $g(t,\ y)$ in the structure of (7) could potentially comprise a polynomial aspect, along with a trigonometric component.

Let us consider the scenario where the first derivative is in the following format:

$$g(t, y) = \left(\sum_{i=0}^{M} a_i t^i\right) y + b_0 + \sum_{j=1}^{N} b_j \sin\left(\frac{2\pi j}{\theta} y + c_j\right) \tag{8}$$

i.e.,

$$y'(t) = \left(\sum_{i=0}^{M} a_i t^i\right) y + b_0 + \sum_{j=1}^{N} b_j \sin\left(\frac{2\pi j}{\theta} y + c_j\right). \tag{9}$$

The parameters under consideration are as follows:

$$a_0,\ a_1, \ldots a_M,\ b_0,\ b_1, \ldots, b_N,\ c_1, c_2, \ldots, c_N, \theta. \tag{10}$$

The number of these parameters is $2N + M + 3$. If the condition $n \geq 2N + M + 4$ (which will be assumed to always hold) is fulfilled, they could be determined by resolving an inverse problem employing a numerical method involving a one-step explicit or implicit approach (or a combination of the two):

$$\frac{(y_{k+1} - y_k)}{h} = Z(g) \textit{ or } y_{k+1} = y_k + hZ(g), \tag{11}$$

where $h = t_{k+1} - t_k$, while g denotes the right-hand side of (6), and $Z(g)$ signifies a specific one-step numerical method, such as explicit or implicit Euler, Runge–Kutta, etc. As demonstrated in [7], one could utilize the last $2N + M + 4$ values of the time series (5). Employing (11) leads to the establishment of a system of nonlinear equations concerning the coefficients (10). Once this system is solved, the considered coefficients (10) can be determined. These coefficients are then employed to define the forthcoming unknown value, y_{n+1}, at timepoint t_{n+1}. In Equation (6), all coefficients are already known, enabling the computation of the next y_{n+1} value through the numerical method Z.

One limitation of this methodology is its failure to utilize the complete information embedded within the time series (5). As per this approach, only as many values are extracted from series (5) as necessary to close the system of nonlinear equations, determined by the specific selection of M and N. However, this shortcoming was effectively circumvented in this research, wherein an arbitrary number of k values could be selected (when $n \geq k > 2N + M + 4$) from line (5). Subsequently, the overdetermined system of nonlinear algebraic equations was solved through the application of the least squares method (LSM). The LSM could also be applied in a weighted format, allowing for the incorporation of a weight function. It stands to reason that the weight function should assign higher weights to the more recent values in the time series (5), thus demonstrating an increasing trend or, at the very least, remaining constant.

Within the context of this paper, the weight of the errors in the least squares method (LSM) is determined by the ascending function:

$$w(t) = t^{\alpha}, \;\; t > 0, \, \alpha > 0. \tag{12}$$

Various simulations were conducted at diverse weight values of α (at $\alpha > 1$ the function (12) is convex; at $\alpha < 1$ it is concave). To be precise, the weights (13) were employed in lieu of the weight function (12).

$$w_i = w(t_i) = \left(\frac{t_i - t_0}{t_k}\right)^{\alpha}, \;\; t_i > 0, \, \alpha > 0, \, i = \overline{0, \ldots k}. \tag{13}$$

The utilization of (13) is motivated by the assignment of a weight to each error within the range of [0, 1]. A weight of 0 is assigned to the data point furthest in time, while the closest (most recent) data point is assigned a weight of 1. This selection process is primarily subjective, allowing for the potential use of varying weights, provided they follow an increasing pattern.

A predicament arises when employing the LSM (10) to define the coefficients, resulting in the resolution of a nonlinear system of algebraic equations. It is widely acknowledged that nonlinear systems could yield multiple solutions, and different selections among these solutions can lead to diverse predictions [11].

Due to the inherent nonlinearity of the system, it is not feasible to ascertain beforehand the exact number of solutions that it may offer. When employing a specific numerical method to solve the nonlinear system regarding the coefficients (10), initial approximations for these coefficients are provided. Different choices of initial approximations typically result in varied solutions. In this study, the approach involved the selection of L different initial approximations:

$$a_0^{0_r}, \, a_1^{0_r}, \ldots a_M^{0_r}, \, b_0^{0_r}, \, b_1^{0_r}, \ldots, b_N^{0_r}, \, c_1^{0_r}, c_2^{0_r}, \ldots, c_N^{0_r}, \theta^{0_r}, \, r = \overline{1, \ldots L}. \tag{14}$$

Solutions in this context are generated through a pseudo-random mechanism, with each one allocated within a predetermined range. These individual solutions stand distinct from one another. To pick a particular solution from this array, the most recent l values in the time series are extracted and used to calculate an error, thus validating a specific solution from Equation (10). These values are referred to as “test values”. The following value is then forecasted and compared with its actual counterpart, after which the real value is reintegrated into the time series and its associated error is noted. This procedure is methodically applied to these l values in the series. The “final error” is then determined as the weighted mean of these l errors, where an ascending function can be utilized for the weighting. The solution from Equation (11) that results in the least “final error” is chosen to define Equation (11) for a given set of M and N values. Thereafter, the upcoming value y_{n+1} at time t_{n+1} is forecasted.

Furthermore, solving the overdetermined system of nonlinear equations can be accomplished via the minimax method, which aims to minimize the maximum discrepancy between the estimated and real values. In addition, the implementation of a weighted minimax approach is also feasible in this scenario.

The approach for predicting the subsequent value in the time series at given intervals and given time series values (5) can be summarized as follows:

1. Fit an ordinary differential equation with initial condition (6)—a Cauchy problem to the time series.
2. Choose the form (8) for the function $g(t,\ y)$ by fixing the parameters M and N.
3. Select a specific numerical method—such as explicit, implicit, or a combination of both, including one-step methods like Euler, Milne, Runge–Kutta, etc.—for the numerical solution of (6).
4. Select the last k values of the time series (5) (where $n \geq k > 2N + M + 4$).
5. Solve an inverse problem with respect to the unknown coefficients (10), which reduces to solving an overdetermined system of nonlinear equations.
    - Choose a method to minimize the given error for solving the overdetermined system—whether the LSM, the minimax method, or a similar approach.
    - Select a weight function to apply this method, ensuring that the weight function is non-decreasing (in this specific case, of type (13)).
    - Apply the chosen method with the selected weight function to reduce the overdetermined system to a nonlinear system of equations.
    - Due to the nonlinearity of the obtained system, generate L different pseudo-random initial approximations (14) for the coefficients (10) and obtain different solutions for (10).
6. Select the last l values of (5) and segregate them to measure an error and validate the chosen solution of (10). Compute the "final error" as a weighted average of the l errors obtained, where the weights are increasing.
7. Choose the solution of (10) with the least "final error".
8. Substitute the chosen solution of (10) into (11), and predict the future y_{n+1} value at time t_{n+1} for $k = n$.
9. Repeat Steps 1–8 for different choices of the numerical method used to solve the initial Cauchy problem, the parameters M, N, k, l, and L, the numerical method for solving the overdetermined system of nonlinear equations, the weight function for solving the overdetermined system of nonlinear equations, and the weight function for estimating the "final error".
10. After obtaining various forecasts and their "final errors", compute their weighted average value, with the weights being inversely proportional to their "final errors", and consider it as the final forecast.

2.3. Risk Portfolio Optimization

The authors utilized a previously established MATLAB programming code [12] to generate an efficient risk portfolio comprising a set of n volatile securities. The primary objective was to maximize the capital allocation line (CAL) for every eligible risk portfolio labeled as p. The code effectively resolved the optimization problem (15) while adhering to the specified condition (16):

$$\min F = -max\ S_P = -\frac{E(r_p) - r_f}{\sigma_P} \tag{15}$$

$$\sum_{i=1}^{n} w_i = 1, \tag{16}$$

where:

- w_i—weight of the i-th asset;
- $E(r_p)$—the anticipated return rate of the risk portfolio refers to the average value of the expected return rates of the volatile assets, weighted by their respective proportions

within the risk portfolio. This value was computed in accordance with the equation provided in (17).

$$E\left(r_p\right) = \sum_{i=1}^{n} w_i E(r_i) \tag{17}$$

- σ_P—the standard deviation, calculated as shown in (18).

$$\sigma_P = \sqrt{\sum_{i=1}^{n} w_i^2 \sigma_i^2 + \sum_{i=1,\ i\neq j}^{n} \sum_{j=1}^{n} w_i w_j \sigma_i \sigma_j \rho\left(r_i, r_j\right)} \tag{18}$$

- $\rho\left(r_i, r_j\right)$—the correlation coefficient between the rates of return of the i-th and the j-th asset.

Solving the optimization challenge involved utilizing a revised version of the Markowitz model [13], which is built on certain key presumptions: the existence of risk-free investment options, the ability to lend at a risk-free interest rate, and the allowance for short-selling of fluctuating assets.

The specific MATLAB code used is detailed in Appendix A. The data fed into this code included forecasts of returns, their deviations from the average, the relevant correlation matrix, and the yield of risk-free investments. The code's output delivers an optimized risk portfolio for n assets across a single period, and it also calculates the anticipated return for this specified risk portfolio. Within the framework of this program, the calculations initially overlook the signs of the return rates. Yet, when analyzing the results, these signs are crucial to decide whether to opt for a short (negative) or long (positive) stance on a particular asset.

3. Results

The data for the four companies used in the current work were taken from the Bulgarian Stock Exchange [14].

3.1. Model Development

3.1.1. ARIMA Model Development

Using the daily closing prices of the shares of the companies Allterco (A4L), Chimimport (CHIM), Elana AgroCredit (EAC), and Speedy (SPDY) from the interval 1 June 2020–29 October 2020 (151 data points—from 0 to 150, including weekends and missing data), diverse forecasts were generated using the ARIMA approach for the closing price value on the following day, i.e., 30 October 2020. A subset of the last known six periods ($l = 6$) was reserved for validation purposes.

The development sample for each instrument encompassed data excluding the validation periods specific to the corresponding instrument. Successful models identified in the development sample underwent incremental testing by integrating the validation periods step by step. Calculations were made for the errors derived from the predictions at each step, which were then employed to compute the weighted average error, adhering to the principle that recent observations bear greater significance.

A careful examination was conducted to ascertain whether the models, when applied to the development and validation data, exhibited any significant deviations beyond the 95% confidence intervals in the ACF (autocorrelation function) and PACF (partial autocorrelation function) residual graphs. If no substantial deviations or minor discrepancies were observed, the model was applied to the complete dataset for the relevant instrument, thus producing the forecasted value for one period ahead.

For each period during which the respective models undergo validation, forecast errors are computed. The outcomes are systematically presented in [2].

The aim was to generate a forecast for $l = 151$. The computation of the weighted average error was conducted as follows: the assigned weights for computing the weighted average error were $\frac{1}{7.5}; \frac{1.1}{7.5}; \frac{1.2}{7.5}; \frac{1.3}{7.5}; \frac{1.4}{7.5}; \frac{1.5}{7.5}$. These weights correspond to the errors in the retained validation cases. The weight allocation was determined subjectively, based on a linear function established through empirical observation. The latest error was given 1.5 times the weight of the first. Table 2 provides the projected values and weighted average errors for the selected models for each financial instrument.

Table 2. Predicted value and weighted average error of the chosen models for 2020.

Instrument	Model	Predicted Value 30 October 2020	Weighted Average Error
A4L	ARIMA (2, 1, 6)	5.1525	0.0965
	ARIMA (1, 0, 10)	5.0102	0.0993
CHIM	ARIMA (4, 1, 7)	0.8826	0.0152
	ARIMA (3, 1, 10)	0.8758	0.0135
EAC	ARIMA (2, 1, 6)	1.0477	0.0101
	ARIMA (1, 0, 6)	1.0394	0.0031
SPDY	ARIMA (2, 1, 9)	56.3522	0.5224
	ARIMA (3, 1, 10)	56.0732	0.6015

A linear combination of the predictions for the two models was employed for each instrument to derive the ultimate projected value. The coefficients in this linear combination were inversely correlated with the weighted average errors obtained from the specified models. The findings are outlined in [2].

For example, the projected price for the instrument A4L on 30 October 2020 was derived as illustrated in (19):

$$\frac{\frac{1}{0.0965} * 5.1525 + \frac{1}{0.0993} * 5.0102}{\frac{1}{0.0965} + \frac{1}{0.0993}} = 5.0824. \tag{19}$$

3.1.2. Modified ODE Model Development

In a manner akin to the ARIMA methodology, to finalize a solution from among various possibilities, the most recent l values (here, $l = 6$, accounting for periods of non-availability) are extracted from the time series. These values are segmented to generate an error, which, in turn, validates a specific solution as per Equation (6). These extracted values are termed “test values”. The value that follows is then forecasted and compared with its actual counterpart, after which this actual value is reintegrated into the time series, with the ensuing error recorded. This procedure is methodically repeated for the $l = 6$ values in the series. The “final error” is then calculated as the weighted mean of these six errors, where an ascending function might be used for determining the weights. The solution from Equation (6) that yields the smallest “final error” is then utilized to formulate Equation (7), taking into account chosen values of M and N. Following this, the next value y_{n+1} at time t_{n+1} is projected.

Utilizing the same dataset as in the ARIMA approach, multiple forecasts for each instrument were produced using varied M and N values in the modified ODE approach. In addressing the nonlinear system of algebraic equations via the least squares method, a linearly increasing weight function was selected. These linearly ascending weights were similarly employed in calculating the “final error”.

The results of the predictions and the “final errors” for each set of parameters M and N are given in [15].

The ultimate forecasts were computed as a weighted mean of all predictions for the individual instruments. The weights were reciprocally correlated with the “final errors”. The ultimate forecasts (estimated values), the real values, and the percentage relative error between the predicted and actual closing price values for the various instruments as of 30 October 2020 are given in [15]. Notably, the forecasts demonstrate exceptional precision (with a maximum relative error of less than 0.6%).

3.2. Model Validation

Data from 2022 were used to validate the models obtained on the data from 2020. The purpose was to check the behavior and the robustness of the models in the conditions of financial crisis. The methodology used on the data from 2020 was applied to the recent data from 2022. This included building the same models, using the daily closing prices of the shares for A4L, CHIM, EAC, and SPDY from the time window 1 June–28 October 2022 (150 periods—from 0 to 149, including weekends and missing data). The last known $l = 6$ periods were kept for validation in order to calculate the corresponding errors. Forecasts for 31 October 2022 were obtained (30 October 2022 was a Sunday; therefore, this time the prediction was for 31 October, which was a Monday—the last working day of October).

3.2.1. ARIMA Model Validation

Drawing a parallel to the ARIMA methodology employed in 2020, the present study adopted a sequential validation approach for the period of 1 June–20 October 2022. When a model demonstrated satisfactory performance on this development sample, it underwent further testing by progressively including the validation periods of 21–30 October 2022, which comprise six distinct periods when weekends are excluded. Following this, prediction errors were computed at each incremental step. These errors subsequently informed the computation of a weighted average error, which was then applied to the predictions for 31 October 2022.

Diagnostic checks carried out on both the development and validation data for 2022 affirmed the robustness of the existing models. Notably, these models exhibited no discernible anomalies or jumps that exceeded the boundaries of the 95% confidence intervals in their ACF and PACF residual plots. Given this affirmation of the models' stability, they were then deployed on the comprehensive datasets spanning 1 June–28 October 2022 for the pertinent financial instrument. The consequent predicted value for 31 October 2022 was subsequently extracted. The outcomes of this exercise are elucidated in Table 3. It is imperative to note that, in instances where data were absent on account of weekends, such omissions were categorized as "no deals", implying that the price remained unchanged from the preceding day.

Table 3. Absolute errors of the predictions from the ARIMA models, applied to the data for 2022.

Instrument	**ARIMA (*p*, *d*, *q*)**	**Absolute Errors**					
		***l* = 142**	***l* = 145**	***l* = 146**	***l* = 147**	***l* = 148**	***l* = 149**
A4L	(2, 1, 6)	0.1048	0.1420	0.0532	0.6884	0.4907	0.1009
	(1, 0, 10)	0.0397	0.0483	0.2304	0.6647	0.5802	0.1506
CHIM	(4, 1, 7)	0.0186	0.0089	0.0302	0.0100	0.0128	0.0050
	(3, 1, 10)	0.0031	0.0014	0.0354	0.0086	0.0269	0.0122
EAC	(2, 1, 6)	0.0196	0.0013	0.0059	0.0191	0.0114	0.0024
	(1, 0, 6)	0.0229	0.0153	0.0219	0.0198	0.0198	0.0224
SPDY	(2, 1, 9)	3.5699	1.1275	0.2297	0.6225	0.4253	0.6345
	(3, 1, 10)	4.3457	0.7079	0.2853	0.2296	1.2233	0.4009

The weighted average error was derived as described in Section 3.1.1. Table 4 shows the predicted values and weighted average errors of the chosen models, applied to the data for 2022.

In a comparable approach to the one applied to the 2020 data, this study employed a linear combination of the 31 October 2022 forecasts derived from both models for each financial instrument to ascertain the ultimate forecasted value for each respective instrument. The coefficients used in this linear combination were inversely related to the weighted average errors produced by the respective models. These findings are documented in Table 5.

Table 4. Predicted value and weighted average error of the chosen models for 2022.

Instrument	Model	Predicted Value for 31 October 2022	Weighted Average Error
A4L	ARIMA (2, 1, 6) ARIMA (1, 0, 10)	17.6247 17.8545	0.2744 0.3029
CHIM	ARIMA (4, 1, 7) ARIMA (3, 1, 10)	0.7786 0.7804	0.0137 0.0152
EAC	ARIMA (2, 1, 6) ARIMA (1, 0, 6)	0.9789 1.0144	0.0097 0.0204
SPDY	ARIMA (2, 1, 9) ARIMA (3, 1, 10)	107.6251 107.5661	0.9923 1.0772

Table 5. Predicted price values, actual price values, and relative % error between them, using ARIMA models.

Financial Instrument	Estimated Value for 31 October 2022	Actual Value for 31 October 2022	Relative Error in %
A4L	17.7339	18.65	4.9119
CHIM	0.7795	0.78	0.0701
EAC	0.9903	0.995	0.4714
SPDY	107.5968	107.00	0.5578

3.2.2. Modified ODE Model Validation

Using a similar approach to that described in Section 3.1.2, and with the same values for l $(l = 6)$, M, and N, estimates for the prices and final errors were made for 31 October 2022. The results are given in Tables 6–13.

The final predictions were derived as a weighted average of forecasts across all individual financial instruments. Intriguingly, these weights were inversely related to the "final errors". Table 14 showcases these definitive forecasts (anticipated values), juxtaposed against the actual values, alongside the percentage discrepancy between the forecasted and realized closing prices of the instruments as of 31 October 2022. It is noteworthy that the forecasts exhibit remarkable accuracy, manifesting a maximal relative deviation of under 2.6%.

Table 6. Estimated values of the closing prices of the A4L instrument for 31 October 2022 under different model selections.

Values of M/N	$N = 2$	$N = 3$	$N = 4$	$N = 5$
$M = 0$	18.28096355	18.09513597	18.08016462	17.89179719
$M = 1$	18.17829696	18.00695886	18.01112314	18.05661753
$M = 2$	19.66848655	19.15769258	18.72543461	16.97272997
$M = 3$	16.91553048	18.70972867	16.12132696	21.36109005

Table 7. "Final errors" in the "test values" for the closing price of the A4L instrument for 31 October 2022 under different model choices.

Values of M/N	$N = 2$	$N = 3$	$N = 4$	$N = 5$
$M = 0$	0.10055900	0.07681535	0.08236390	0.04989364
$M = 1$	0.15965715	0.08840078	0.07980459	0.08133888
$M = 2$	0.01886979	0.05639571	0.05568973	0.00743097
$M = 3$	0.07079675	0.03217803	0.03329889	0.00650684

Table 8. Estimated values of the closing prices of the CHIM instrument for 31 October 2022 under different model selections.

Values of *M*/*N*	*N* = 2	*N* = 3	*N* = 4	*N* = 5
M = 0	0.80500336	0.78203687	0.79825092	0.78912960
M = 1	0.79096175	0.79034851	0.79107247	0.79014385
M = 2	0.77761242	0.77231375	0.76321097	0.81030482
M = 3	0.79378890	0.77494780	0.82276553	0.78763064

Table 9. "Final errors" in the "test values" for the closing price of the CHIM instrument for 31 October 2022 under different model choices.

Values of *M*/*N*	*N* = 2	*N* = 3	*N* = 4	*N* = 5
M = 0	0.03933323	0.05206715	0.03723752	0.07488105
M = 1	0.10632697	0.10693371	0.10709779	0.10159614
M = 2	0.03494464	0.05875035	0.03669832	0.03072809
M = 3	0.05611642	0.06096068	0.01475785	0.08157008

Table 10. Estimated values of the closing prices of the EAC instrument for 31 October 2022 under different model selections.

Values of *M*/*N*	*N* = 2	*N* = 3	*N* = 4	*N* = 5
M = 0	0.97349605	1.05003867	0.98470937	0.99199108
M = 1	0.99967003	0.99980277	0.99943251	1.00033667
M = 2	1.00902475	0.98998030	0.97504687	1.0045637
M = 3	1.03042746	0.99630033	0.97227905	0.9887555

Table 11. "Final errors" in the "test values" for the closing price of the EAC instrument for 31 October 2022 under different model choices.

Values of *M*/*N*	*N* = 2	*N* = 3	*N* = 4	*N* = 5
M = 0	0.02386445	0.01023558	0.10329024	0.01240245
M = 1	0.10902632	0.13969440	0.13510329	0.12584670
M = 2	0.02729210	0.05297982	0.06652198	0.06267512
M = 3	0.03251327	0.02605377	0.03249936	0.04000114

Table 12. Estimated values of the closing prices of the SPDY instrument for 31 October 2022 under different model selections.

Values of *M*/*N*	*N* = 2	*N* = 3	*N* = 4	*N* = 5
M = 0	106.82119691	110.10191152	107.37708304	106.83075939
M = 1	106.89208241	106.82065168	107.02550172	106.91839134
M = 2	100.89401340	97.422710508	112.29010460	102.64214538
M = 3	95.362169097	89.974116039	119.80544693	107.98492701

Table 13. "Final errors" in the "test values" for the closing price of the SPDY instrument for 31 October 2022 under different model choices.

Values of *M*/*N*	*N* = 2	*N* = 3	*N* = 4	*N* = 5
M = 0	0.09308016	0.09909670	0.09640184	0.08576109
M = 1	0.09448638	0.10674863	0.09032125	0.08246354
M = 2	0.03791003	0.02761275	0.04940139	0.06708805
M = 3	0.00618474	0.00493533	0.01082146	0.04768667

Table 14. Predicted price values, actual price values, and relative % error between them, using modified ODE models.

Financial Instrument	Estimated Value for 31 October 2022	Actual Value for 31 October 2022	Relative Error in %
A4L	18.1758	18.65	2.5426
CHIM	0.7899	0.78	1.2692
EAC	0.9998	0.995	0.4824
SPDY	106.6469	107.00	0.3300

3.3. Risk Portfolio Optimization

The programming code from Appendix A was applied, using the input data from 2020 and 2022. These data include the following: the estimates of the expected rates of return (Table 15), derived as a relative difference between the estimated value for 30 October 2020 (31 October 2022) and the real value for 29 October 2020 (28 October 2022, respectively); the standard deviations, calculated using historical data (Table 15); the correlation matrix, based on historical data (see [15] for 2020 and Table 16 for 2022); the return of the risk-free asset.

Table 15. Expected rates of return and standard deviations.

Financial Instrument	Expected RoR–ARIMA 2020	Expected RoR–Modified ODE 2020	Std. Deviation 2020	Expected RoR–ARIMA 2022	Expected RoR–Modified ODE 2022	Std. Deviation 2022
A4L	1.6480%	−0.31%	3.3276%	0.1917%	2.6881%	2.1724%
CHIM	−0.1136%	−0.54%	1.7765%	−1.0846%	0.2411%	1.9702%
EAC	0.1250%	0.04%	1.2198%	−0.4714%	0.4824%	7.1714%
SPDY	−1.3640%	0.19%	1.5851%	0.5578%	−0.3300%	1.9739%

Table 16. Correlation matrix for 2022.

Financial Instrument	A4L	CHIM	EAC	SPDY
A4L	1.0000	−0.0523	−0.0444	−0.0255
CHIM	−0.0523	1.0000	−0.0246	0.0880
EAC	−0.0444	−0.0246	1.0000	−0.0476
SPDY	−0.0255	0.0880	−0.0476	1.0000

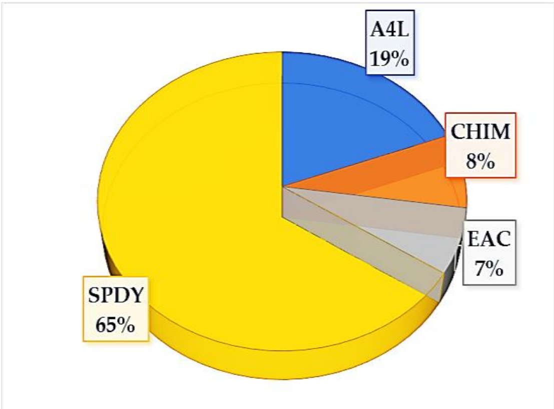
For 2020, a government security with an annual yield of 3%, which equals 0.0083% daily yield (3%/360), was employed. For 2022, the annual yield of the government securities was 4.36%, or 0.0121% daily yield.

It can be seen from Table 15 that the expected rates of return differ not only by year (2020 vs. 2022), but also by the type of model. The ARIMA predictions for 2022 (Table 15, Column 5) for A4L, CHIM, and SPDY confirmed the movement (increase/decrease) observed in the ARIMA predictions from 2020 (Table 15, Column 2), while that for EAC was not confirmed—it underwent an increase of 0.125% in 2020 but a decrease of 0.4714% in 2022. The modified ODE predictions gave us confirmation of the movement only for EAC. From the second and the third columns of Table 15, it can also be noted that the ARIMA and modified ODE approaches predicted the same trend (increase) in 2020 for CHIM and EAC and showed discrepancies for A4L and SPDY, while in 2022 we obtained the same trend predictions only for A4L. The standard deviations for the two periods (Table 15, Columns 4 and 7) show that EAC is a much riskier instrument for 2022 (7.1734% standard deviation) than for 2020 (1.2198% standard deviation).

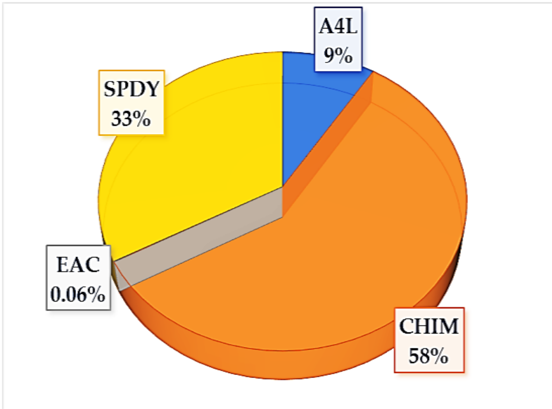
The correlation matrices for the two periods also indicate different movements and connections between the companies in the time window in 2022 compared to that in 2020.

The outcomes from the optimization algorithm were as follows: For 30 October 2020, the optimal risk portfolio shaped by ARIMA forecasts (as shown in Figure 1a) includes

19.23% A4L stock in a long position, 8.37% CHIM stock in a short position, 7.20% EAC stock in a long position, and 65.21% SPDY stock in a short position. Conversely, the portfolio for the same period crafted based on the modified ODE predictions (illustrated in Figure 1b) comprises 9.18% A4L stock in a short position, 57.77% CHIM stock in a short position, a mere 0.06% EAC stock in a long position, and 32.99% SPDY stock in a long position.



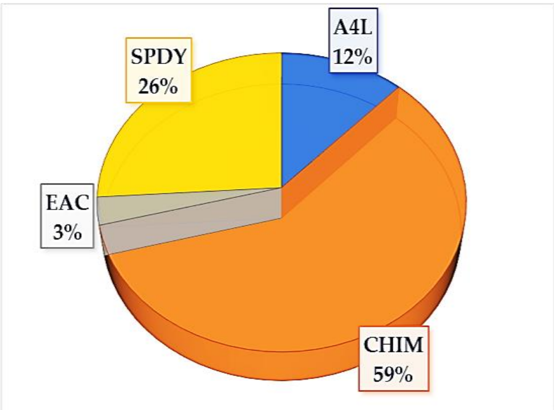
(**a**)



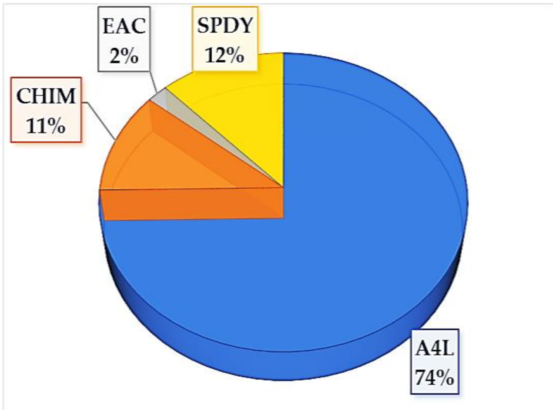
(**b**)

Figure 1. Assets' shares in the optimal risk portfolio for 30 October 2020: (**a**) Portfolio constructed using the ARIMA predictions. (**b**) Portfolio constructed using the modified ODE predictions.

Similarly, for 31 October 2022, the optimal risk portfolio, as per the ARIMA forecasts (depicted in Figure 2a), includes 12.10% A4L stock in a long position, 58.86% CHIM stock in a short position, 2.85% EAC stock in a short position, and 26.18% SPDY stock in a long position. For the same date, the portfolio structured using the modified ODE predictions (shown in Figure 2b) contains 74.41% A4L stock in a long position, 11.12% CHIM stock in a long position, 2.42% EAC stock in a long position, and 12.05% SPDY stock in a short position.



(**a**)



(**b**)

Figure 2. Assets' shares in the optimal risk portfolio for 31 October 2022: (**a**) Portfolio constructed using the ARIMA predictions. (**b**) Portfolio constructed using the modified ODE predictions.

Table 17 delineates the assets' positions within the structured risk portfolio, with long positions indicated by a (+) and short positions by a (−). The last column, bearing the designations (+), (−), and (0), reflects the authentic trajectory of the closing prices, wherein a "0" signifies a stagnant price with no observed alteration. For each financial instrument, the deviation between the last known price and the actual price was computed.

Table 17. Predictions, actual values, and relative % error between predicted and actual closing price values for different instruments for 30 October 2020 and 31 October 2022.

Financial Instrument	Date	Estimated Value ARIMA	Portfolio Weights ARIMA %	Estimated Value ODE	Portfolio Weights ODE %	Last Known Price	Actual Price
A4L	30 October 2020 31 October 2022	5.0824 17.7339	19.23 (+) 12.10 (+)	4.9846 18.1758	9.18 (−) 74.41 (+)	5.00 17.70	4.98 (−) 18.65 (+)
CHIM	30 October 2020 31 October 2022	0.8790 0.7795	8.37 (−) 58.86 (−)	0.8752 0.7899	57.77 (−) 11.12 (+)	0.88 0.79	0.87 (−) 0.78 (−)
EAC	30 October 2020 31 October 2022	1.0413 0.9903	7.20 (+) 2.85 (−)	1.0404 0.9998	0.06 (+) 2.42 (+)	1.04 1.00	1.04 (0) 0.995 (−)
SPDY	30 October 2020 31 October 2022	56.2225 107.5968	65.21 (−) 26.18 (+)	57.1085 106.6469	32.99 (+) 12.05 (−)	57.00 107.00	57.00 (0) 107.00 (0)

Given the asset weights in the portfolio, and considering the positions—both long and short—as indicated in Table 17, we can compute the actual percentage return for each methodology on both 30 October 2020 and 31 October 2022, factoring in the known direction of the trend. For instance, employing the ODE approach for 2022 yields

$$\left(\frac{|18.65 - 17.70|}{18.65}\right) * 74.41 + \left((-1) * \frac{|0.78 - 0.79|}{0.78}\right) * 11.12 + \left((-1) * \frac{|0.995 - 1|}{0.995}\right) * 2.42 + 0 = 3.6356\% \quad (20)$$

Table 18 shows a summary of the expected rates of return and estimates of the standard deviations of the optimal risk portfolios, constructed using the predictions from ARIMA and ODE, for both 30 October 2020 and 31 October 2022. It also shows the real percentage return, calculated as shown in (20).

Table 18. Expected rates of return, standard deviations, and real percentage return of the optimal risk portfolios.

Method Used for the Predictions	Date	Expected RoR of the Risk Portfolio	Standard Deviation Estimate of the Risk Portfolio	Real Percentage Return
ARIMA	30 October 2020 31 October 2022	1.2248% 0.8211%	1.1622% 1.3278%	0.0190% 1.3853%
ODE	30 October 2020 31 October 2022	0.4031% 2.0784%	1.1371% 1.6338%	0.7009% 3.6356%

4. Discussion

First, we must comment on the relative errors in prediction. It is obvious that the maximal errors generated by the modified ODE approach are strictly less than their ARIMA counterparts. Furthermore, the maximal relative errors for both methods are larger for the newer period in 2022 than for 2020. Interestingly, in three of the four cases, the largest errors were obtained for A4L. This is the instrument with the largest standard deviation for the 2020 period. However, its standard deviation decreased in 2022, while it increased for all other instruments. In general, the larger variance for the more recent period could be explained by the unstable global situation in economic and political terms. This is also the reason for the increased risk in the portfolio optimization for 2022 compared to 2020.

As discussed earlier, in the case of modified-ODE forecasting, the maximal relative errors were less than 0.6% for 2020 and less than 2.6% for 2022. These are remarkably lower than those reported in the literature. For example, in [8], the commodity prices traded on the London Stock Exchange were predicted. In half of the cases, the error in the monthly forecasts was lower than 10%, while in the others it was greater than 10%. In comparison, the authors of [7] forecasted the prices of petroleum and bank stocks. In most cases, their prediction error was slightly larger than 2%, while in two cases it was around 1%. All of this unequivocally means that the forecasting results reported in this study significantly outperform the known ones in the discussed literature.

It is noteworthy that all of the real returns observed in Table 4 are positive. This is true in spite of the fact that the absolute prices of some instruments (i.e., A4L and SPDY) are much larger than their levels two years ago. The real percentage returns obtained using the modified ODE forecasts for both periods were greater than those obtained by means of the ARIMA forecasts. Furthermore, regardless of the prediction method, the real percentage returns for 31 October 2022 were significantly higher than the ones for 30 October 2020. Of course, this comes with the price of the increased risk of the portfolio, in terms of standard deviation, for the new period in view.

With respect to the results obtained for 31 October 2022, the optimal risk portfolio constructed using the ARIMA forecasts correctly predicted the trend (i.e., increase/decrease) for three of the four instruments, constituting 73.82% of the total exposure. Conversely, for the risk portfolio obtained by means of the modified ODE forecasts, only one instrument was predicted correctly. However, this instrument was EAC, which represents 74.41% of the risk portfolio's shares, compensating for the mispredicted trends for the other three instruments.

5. Conclusions

In the intricate realm of financial market predictions, especially during economically volatile times, accuracy and precision in forecasting models play a paramount role. This study's validation of two forecasting methodologies, ARIMA and a modified ODE approach, the latter of which was developed by the authors, has shed light on their respective efficacies in predicting stock prices, especially within the context of a financial crisis. It was discerned that the modified ODE method consistently rendered predictions with smaller relative errors compared to its ARIMA counterpart. Furthermore, the observation that the discrepancies in these predictions were more pronounced for 2022, vis-à-vis 2020, suggests that external economic and political factors introduced greater market uncertainty in the more recent period.

One key implication of this research pertains to the practicality of portfolio optimization. By leveraging more accurate predictions, professionals in investment and hedge funds can refine their strategies to achieve optimal returns. This is epitomized by the results for 31 October 2022, wherein the modified-ODE-based portfolio, despite having only one instrument predicted with the correct trend, managed to yield significant positive returns due to the dominance of that correctly predicted instrument. This underscores the importance of not just the number of accurately predicted instruments, but their relative weights within a portfolio, which is the foundation of a sound diversification.

In summary, this research accentuates the paramount importance of continuous model validation and refinement in the world of finance. In an environment fraught with uncertainty, especially during financial downturns, harnessing better prediction tools is not merely a means of achieving better returns, but also of ensuring the stability and resilience of investments. The methodologies explored and the insights garnered from this study can serve as valuable tools for professionals aiming to navigate the turbulent waters of financial markets during crises.

There are many ways in which the current research could be further developed. One possible way would be to perform portfolio hedging for a number of consecutive periods—e.g., weeks or days—in order to gain more confidence in the resulting relative profit. Furthermore, we could combine the ARIMA, modified ODE, and even LSTM methodologies in a single hybrid approach in order to obtain more precise forecasts. Eventually, we could upgrade the classical Markowitz idea of portfolio optimization, as the risk could be evaluated in a more advanced way than by simply observing the standard deviation. This combination could lead to higher and more reliable returns.

Author Contributions: Conceptualization, V.M. and I.G.; methodology, V.M. and I.G.; software, I.G. and S.G.; validation, V.M., I.G., E.R., S.G. and V.P.; formal analysis, I.G.; investigation, V.M. and I.G.; resources, V.M. and I.G.; data curation, V.M., I.G. and S.G.; writing—original draft preparation, V.M., I.G. and S.G.; writing—review and editing, V.M., I.G., E.R., S.G. and V.P.; visualization, V.M. and I.G.; supervision, V.P.; project administration, V.P.; funding acquisition, I.G., S.G. and V.P. All authors have read and agreed to the published version of the manuscript.

Funding: This study is supported by the project BG05M2OP001-1.001-0004 UNITe, funded by the operational program "Science and Education for Smart Growth", co-funded by the European Union through the European Structural and Investment Funds and by the Bulgarian National Science Fund under project KP-06-M62/1 "Numerical deterministic, stochastic, machine and deep learning methods with applications in computational, quantitative, algorithmic finance, biomathematics, ecology and algebra" from 2022.

Data Availability Statement: The data are freely available online (see [14]).

Acknowledgments: The authors would like to thank the anonymous reviewers for their constructive comments, which have significantly improved the quality of the paper.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, Analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Appendix A MATLAB Programming Code for Risk Portfolio Optimization

```
function [S,W_p,Er_p, StD_p]=fopt_portfolio(ERoR,StD,CM,RoRf)
%%% INPUT DATA %%%
% ERoR—Expected Rate of Return (RoR)—n-dimensional vector
% StD—Standard Deviation of RoR—n-dimensional vector
% CM—Correlation Matrix—nxn matrix
% RoRf—RoR of the risk-free asset
n=length(ERoR);
ERoR = ERoR(:);
StD = StD (:);
% CovM—Covariance Matrix
CovM = ( StD*StD').*CM;
% F = -S—optimization function
F = @(w)-(w(:)'*ERoR-RoRf)./sqrt(w(:)'*CovM*w(:));
% w—vector of weights
% Search that w, where min(F)
[W_p,S] = fmincon(F,ones(n,1)/2,[],[],ones(1,n),1,zeros(n,1),ones(n,1));
%%% OUTPUT DATA %%%
Er_p = W_p'*ERoR; S = −S;
% W_p—weights in the optimal risk portfolio
% Er_p—Estimated RoR of the optimal risk portfolio
StD_p=sqrt(W_p'* CovM*W_p);
% StD_p—Standard Deviation of RoR of the optimal risk portfolio
```

References

1. Cheng, L.; Shadabfar, M.; Sioofy Khoojine, A. A State-of-the-Art Review of Probabilistic Portfolio Management for Future Stock Markets. *Mathematics* **2023**, *11*, 1148. [CrossRef]
2. Mihova, V.; Centeno, V.; Georgiev, I.; Pavlov, V. Comparative Analysis of ARIMA and Modified Differential Equation Approaches in Stock Price Prediction and Portfolio Formation. In *New Trends in the Applications of Differential Equations in Sciences*; Springer: Berlin/Heidelberg, Germany, 2023; pp. 341–351.
3. Dimov, I.; Georgieva, R.; Todorov, V. Balancing of systematic and stochastic errors in Monte Carlo algorithms for integral equations. In *Lecture Notes in Computer Science*; Springer: Cham, Switzerland, 2015; Volume 8962, pp. 44–51.
4. Ngo, H.D.; Bros, W. The Box-Jenkins methodology for time series models. In Proceedings of the SAS Global Forum 2013 Conference, San Francisco, CA, USA, 28 April–1 May 2013; Volume 6, pp. 1–11.

5. Matlab's ARIMA Methodology. Available online: https://www.mathworks.com/help/econ/arimaclass.html?searchHighlight=arima&s_tid=doc_srchtitle (accessed on 26 October 2023).
6. Tabachnick, B.; Fidell, L.; Ullman, J. *Using Multivariate Statistics*; Pearson: Boston, MA, USA, 2007; Volume 5.
7. Xue, M.; Lai, C.H. From time series analysis to a modified ordinary differential equation. *J. Algorithms Comput. Technol.* **2018**, *12*, 85–90. [CrossRef]
8. Lascsáková, M. The analysis of the numerical price forecasting success considering the modification of the initial condition value by the commodity stock exchanges. *Acta Mech. Slovaca* **2018**, *22*, 12–19. [CrossRef]
9. Dimitrov, Y.; Dimov, I.; Todorov, V. Numerical solutions of ordinary fractional differential equations with singularities. *Stud. Comput. Intell.* **2019**, *793*, 77–91.
10. Matlab. Available online: https://www.mathworks.com/products/matlab.html (accessed on 26 October 2023).
11. de Melo, M.K.; Cardoso, R.T.N.; Jesus, T.A.; Raffo, G.V. Investment portfolio tracking using model predictive control. *Optim. Control. Appl. Methods* **2023**, *44*, 259–274. [CrossRef]
12. Centeno, V.; Georgiev, I.; Mihova, V.; Pavlov, V. Price Forecasting and Risk Portfolio Optimization. In *Application of Mathematics in Technical and Natural Sciences*; AIP Publishing: Melville, NY, USA, 2019; Volume 2164, pp. 060006-1–060006-15.
13. Bodie, Z.; Kane, A.; Marcus, A. *Investments*, 10th ed.; McGraw-Hill Education: Berkshire, UK, 2014.
14. Bulgarian Stock Exchange. Available online: https://www.bse-sofia.bg/en/ (accessed on 26 October 2023).
15. Mihova, V.; Centeno, V.; Georgiev, I.; Pavlov, V. An application of modified ordinary differential equation approach for successful trading on the Bulgarian stock exchange. In Proceedings of the Eighth International Conference New Trends in the Applications of Differential Equations in Sciences (Ntades2021), St. Constantin and Helena, Bulgaria, 6–10 September 2021; AIP Publishing: Melville, NY, USA, 2022; Volume 2459, pp. 030025-1–030025-9.