

Information and algorithmic support of a multi-level integrated system for the investment strategies formation

D A Gercekovich¹, O Yu Basharina^{1,2}, I S Shilnikova¹, E Yu Gorbachevskaya³,
and S A Gorsky²

¹Irkutsk State University, Karl Marx St. 1, Irkutsk, Russia, 664003

²Matrosov Institute for System Dynamics and Control Theory of SB RAS,
Lermontov St. 134, Irkutsk, Russia, 664033

³Irkutsk National Research Technical University, Lermontov St. 83, Irkutsk,
Russia, 664074

basharinaolga@mail.ru

Abstract. The article summarizes the accumulated practical experience of the authors in the development of algorithms for the formation of investment strategies. For this purpose, the optimization of the studied parameters, information support of investment activities, verification, monitoring and adjustment in the testing mode and the subsequent practical application of the described tools are considered. The system is based on the main provisions of the Markowitz portfolio theory. The analytical block of the Information System Portfolio Investor includes Profitability-Risk model; empirical models of optimal complexity; hybrid predictive model systems; the principle of combining (integrating) both models and forecasts, as well as decision rules; optimization of the training sample length (modified Markowitz model); optimization of the frequency of monitoring and adjusting the composition of the investment portfolio. The principles of design and development of the information block of the system, its replenishment and functioning are described in detail. All the above listed components of the algorithmic content of the investment decision making system are described sequentially. The system modules have been successfully tested on a wide class of financial instruments: ordinary shares, preferred shares, government and corporate bonds, exchange commodities, stock, commodity, industry and bond indices, exchange-traded investment funds and real estate funds. The implemented Markowitz model with a dynamic database of historical data can significantly increase the efficiency of investment decisions, which is facilitated by taking into account the characteristics of both the markets under study and the corresponding financial instruments.

1. Introduction

The development of the world economy, trade, the processes of integration of national economies into the world economic system, continuous communications and telecommunications improvement, information technology allow you to make investment transactions around the world "in a matter of minutes". This makes the investment environment more dynamic, more accessible, more vulnerable and, as a result, increases the risks of financial losses. The above has led to a sharp increase in the relevance and importance of developing methods, models, and tools (including software) for making rational decisions that consider the current state of development of financial markets.

The paper has the following structure. In Section 2, the authors briefly review the related works and software products on the problem under study. Section 3 provides a detailed description of the

architecture of the investment decision information system. The results of using this system are presented in Section 4. Section 5 concludes the paper.

2. Related work

Today it is impossible to imagine the work of an individual or an institutional investor without the use of computer technology. The use of computers has a rich history and, as a result, a wide range of currently popular software systems focused on solving a wide class of investment analysis problems. A distinctive feature of foreign software, developed at the end of the 20th century, was the emphasis on providing first services. First of all, this is investment accounting, tracking quotes, minimizing tax liabilities, tracking commission payments and receiving dividends, generating reports, facilitating accounting, maintaining descriptive statistics, managing expenses by checkbook, communication service, text editors Dow Jones PLUS, Dow Jones PLUS 2.0, Dow Jones Spreadbeat Link, Disclosure Database, Dow Jones & Company Programs, Investment Record, Managing Your Money, Modern Jazz, Symphony, etc. The most advanced software packages contain a mathematical toolkit of portfolio theory designed to form an investment portfolio – Value / Screen, CompuServe, Market Analyzer Plus, Compustock, Fundamental Investor, Technical Investor, Stock Portfolio Allocator, Stock Portfolio System, and many others.

A computer equipped with modern software allows you to quickly sort financial instruments among several thousand actively traded stocks, bonds, commodities, derivatives, currencies, and many other [1-5].

Among the developments of software systems devoted to investment analysis, the following can be noted. The world's leading providers of financial information such as Bloomberg, Thomson Reuters offer their own software and hardware tools for professional participants in financial markets. The main products of these companies (Bloomberg Terminal [6], Thomson Reuters Eikon [7]) are terminals for accessing electronic trading platforms and terminals for receiving and analyzing financial information. Information terminals allow access to current and historical prices on virtually all world exchanges and many OTC markets.

Bloomberg and Reuters terminals are quite expensive and intended for large brokers, so alternative products began to appear on the market. For example, aggregators of financial data Yahoo! Finance and Google Finance provide news and background information on topics related to business, finance, and economics. Reference information includes quotes and ratings of securities, press releases and financial statements of most multinational companies. Yahoo! Finance also offers a service for managing an investment portfolio [8]. It allows you to analyze the financial market and track selected investments in real time.

Among the Russian broker companies that provide news of the financial market, and the world economy, various rates, quotes, and indicators should be noted Finam, Finmarket, RBC, AKM, Prime and others.

Investment management programs are becoming more and more popular among individuals who understand that investing can become an additional source of income or a way to accumulate capital, considering inflation. The client can buy foreign exchange, national and foreign stocks, corporate and government bonds (including Eurobonds), ETFs, BPIFs, and structured notes. Also, with several conditions, access to the IPO (initial public offering) is provided.

There are a lot of mobile applications for investment: Webull, Investing, Smart-lab, Intelinvest, Interactive Brokers, Quik and others. The largest banks offer their own investment management services: Tinkoff Investments, VTB My Investments, Sberbank Investor, Otkritie Broker, Alfa Direct, BCS My Broker, etc. The standard set of functions for these services contains the following features: opening an investment account; tracking income and expenses; analytical charts; stock catalogue, exchange rates; news feed; suggestions and advice on the purchase of packages; expert advice.

Some applications use the artificial intelligence algorithms. Its use in Tinkoff Investments [9] allows you to analyze the data of thousands of securities and promptly inform investors about the most important changes in companies' performance to make correct decisions based on real performance.

The application VTB My Investments [10] based on artificial intelligence helps banks create personalized trading offers for traders and makes it possible to track actions on previous transactions, identifies buy or sell transactions that led to financial losses in the past. For each participant in the process, artificial intelligence develops a Profile-of-One, taking into account their trading preferences, internal bank data and information from social networks.

Despite its advantages, the above software, does not make full use of analytical tools of investment analysis. These systems have limited functionality. They are geared towards follower users who prefer to follow the advice of others.

A new Information System (IS) Portfolio Investor was developed in this regard. It includes a wide range of software that generates standardized files compatible with MS Excel. It allows the investor (in real time) not only to perform a wide range of statistical and technical analysis operations, but also to apply for these purposes, the *Search for a solution* and *Data analysis* add-ins built into the spreadsheet processor.

3. Architecture of Information System

IS Portfolio Investor being a multifunctional system for making financial decisions is based on the analysis of historical data. It performs the following functions:

- Formation of an initial set of financial instruments, financial and market indicators, stock, commodity, industry, and other indices; tracking the dynamics of their market quotations, accumulating in the database such a volume of information that is necessary and sufficient for the functioning of the analytical unit of the system.
- Evaluating the investment attractiveness of financial instruments; sorting, and ranking the lists of financial instruments according to the specified quality criteria; calculating expected returns and risk levels; forecasting quotes, forecasting returns..
- Conducting fundamental and technical analysis.
- Construction of functions graphs, including visualization of tendencies and trends (with an assessment of their strength or weakness).
- Making a list of leading instruments for which it is recommended to open transactions.
- Solution of a complex of portfolio tasks: selection of financial instruments; portfolio formation; conducting trading, monitoring the portfolio and adjusting that segment of the portfolio of financial instruments that has ceased to meet the criteria formulated by the investor, for example, due to changes in the market. Tracking the level of risk, both for individual financial instruments and for the portfolio.
- Development, testing and direct application of investment strategies.

To facilitate further presentation, we will divide the conditionally functional content of the IS Portfolio Investor into four blocks: information, analytical, a block for verifying models, forecasts, strategies, portfolios, and an optimization block. Let us consider each of them in more detail.

3.1. Information block

This block is a storage of various kinds of data and operations with them.

The sources of information on the state of the financial markets are the real-time systems. They supply data on quotations of currencies, securities, exchange commodities, etc., as well as the financial and economic news from the international agencies: Reuters, Dow Jones, Telerate, CQG, by Bloomberg, Tenfore, the FT Extel and others. From the Russian information sources Finam.ru, Investing.com, Bhom.ru, InvestFunds.ru, Finanz.ru, etc. are used.

Also in this module, the *initial filling of the database, verification and, if necessary, correction of the incoming information and automatic data replenishment* are carried out. The database is filled (and its subsequent replenishment) with tick data, based on which information is aggregated into time intervals (bars) of a specified value: minute, hour, day, week, month, etc. [11].

The values of time intervals are determined not only by the tasks set by the investor, the specifics of the market, the peculiarities of the dynamics of some financial instruments, but also by the technical

capabilities and established traditions in this area. For the securities market, as a rule, time intervals of one month are formed [12, 13], since the majority of databases on prices and dividends of financial instruments are replenished with monthly data, brokerage firms send monthly reports. When analyzing the feasibility of adjusting the structure of the investment portfolio and the weights of the instruments included in it, time intervals of a smaller dimension can be used [11]. Software trading (trading robots) operates on ultra-short time intervals.

Filling in the gaps in the incoming information is also a necessary tool for the subsequent investment analysis. The lack of data in this time interval may be due to the absence of tenders or data rejection. Missing data are filled in either by "pulling" the data of the nearest previous time interval (this approach is based on the following statement: there was no bidding, which means that the price has not changed), or using interpolation methods.

The information module contains a *news section*. This section, seems to be secondary at first glance, is one of the most important and difficult in investment analysis. It is the ability to carry out an adequate analysis of market changes, and, accordingly, to predict trends, which makes it possible to determine the possibility of making a profit or significant losses in the nearest time interval. Therefore, a competent assessment of quotes, certain events, rumors, expectations, forecasts, and their proximity to real values, etc. is a necessary component of the investor's work. The main purpose of such an analysis is to determine the "moment" of a change in the trend of the financial instrument under study and to answer the question: "When and to what extent should this financial instrument be included (excluded) in the investment portfolio?".

As part of the news analysis, various reports on financial events in the world, political and economic phenomena that may affect the development of the financial market are studied. Here the information about the work of stock exchanges and large companies, discount rates of central banks, the economic course of the government, possible changes in the political life of the country, forecasts and assessments of the state of important economic indicators, as well as all kinds of rumors and expectations are important.

After the message appears, the investor must answer the following questions:

- Is this information expected or not?
- How will it affect (positively or negatively) the state of the financial instrument or even the economic sector under study?

The task is solved by analyzing the situation by publications, while studying market reviews in the information systems Reuters, Dow Jones Telerate, CQG, Bloomberg, Tenfore, etc. The information obtained is gradually developing into a model of various scenarios for further developments in the markets.

3.2. Analytical block

The basis of a multi-level integrated system for the formation of investment strategies is the following components:

- Risk-return model,
- Empirical models of optimal complexity,
- Systems of hybrid empirical models,
- The principle of combining (integrating) models, forecasts, decision rules,
- Optimization of the length of the training sample.

The Return-Risk model, in turn, is based on three pillars: portfolio analysis, winner portfolio, and the concept of optimal training length.

The proposed model is based on the basic principles of classical *portfolio analysis* which are based on the following basic assumptions. An investor, based on statistical data, can:

- Calculate the yield based on historical data. Obtain estimates of the expected (average) values of returns, their pairwise covariances and possible risk.
- Compare selected portfolios based on the ratio of three criteria: average return, risk, and return-to-risk ratio.

- The investor is not inclined to risk, that is, from two portfolios (assets) with the same return, he will choose a portfolio (asset) with a lower risk, or from two portfolios (assets) with the same risk, he will choose a portfolio (asset) with a higher return.

The portfolio of winner is formed from the fixed and predetermined number of shares so that its cost is mainly grown for much of the previous year. The authors of [14-18] showed experimentally that the formed portfolio "by inertia" continues to increase its value. The formation of the winner portfolio procedure similar to the method of O'Shaughnessy [19], the choice of betting on sports scores [20], a method for selecting shares from Graham [20]. Sharpe [21] calls this model momentum investors-impulse investors (inertial investors).

The concept of optimal instruction length was firstly formulated by Gershengorn [22]. Later Fama et al. [23] quantified the optimal training sample size for stock markets.

According to Markowitz [24], the risk of investing in a certain type of financial instruments is determined by the probability of the yield deviation from the expected value. Suppose that during a certain nearest time interval the trends existing in the market will continue, then the predicted value of profitability can be determined based on the processing of historical data on the dynamics of quotations of these assets in the past, and the risk as the standard deviation from the expected profit. In what follows, the term "investor" will mean both an institutional (portfolio manager) and an individual investor.

The algorithm for finding the optimal training length includes the following stages:

Stage 1. The profitability R_t of a financial instrument in every temporarily interval (bare) is calculated by the formula, proposed by Sharpe [20]:

$$R_t = \frac{P_{close} - P_{open}}{P_{open}} \cdot 100(\%),$$

where R_t is the expected profitability of a financial instrument, calculated for the period under consideration, P_{close} is the opening quote at time t , P_{open} is the closing quote.

Step 2. With the help of MS Excel means the sport diagram on the abscissa axis is build and it indicates a risk (standard deviation), and the ordinate axis indicates the expected profit (average profit). Then those instruments are excluded from consideration, which in the studied time interval showed an absolute return close to zero. Further, financial instruments are removed, which, with the same risk, showed a lower profitability, and with the same income, a greater risk. This makes it possible to form a group of leading instruments according to the Profitability-Risk criterion optimum. The remaining tools are used to build a regression equation using the Data Analysis add-in.

Step 3. Fundamental and technical analysis to determine the final tool group leaders is carried out.

Stage 4. Correlation analysis is also carried out. The results of which the assets of the original group leaders forming several subgroups of financial instruments, so as to minimize (to the extent possible within the framework of the problem being solved) the dynamics of the relationship between the analyzed asset returns within a subgroup [25].

Step 5. The efficiency of subgroups is being identified by quantitative comparison of the results with the results of the dynamics of respective currency, stock, raw materials, industry and many other indices (see the next paragraphs). Verification of the results is carried out outside the training sample (Out-Of-Sample method).

Empirical models of optimal complexity are constructed by the elimination method [26-28] according to the principle from complex to simple. The quality of the synthesized model is estimated by the value of the determination coefficient. The model is built using the MS Excel add-in "Data Analysis". A requirement suitability synthetic model is the condition that the standard errors are empirical coefficients, calculated with 95% - were smaller units of its factors:

$$\left| \frac{\Delta x_i}{x_i} \right| < 1, \quad (1)$$

here Δx_i is the error coefficient is calculated with 95%, the confidence level, x_i is coefficient.

At the initial step, a model is built, in the right side of which all the original variables are included. Further, among the input variables, there is a subgroup of variables for which inequality (1) does not hold. Further, the variable for which this ratio is maximal is removed from this subgroup. The model is

rebuilt without this variable. The process of elimination, i.e. reduction of the number of input variables continues until the computed standard errors for all variables are less than the absolute values of the corresponding coefficients.

For all output variables, testing of which on the test sequence showed their limited suitability, *systems of hybrid models* are constructed. For this purpose, in the right parts of all hybrid models under construction, without exception, those basic variables with the previously synthesized predictive models demonstrated acceptable quality on the test sequence are included as essential. That is, the initial list of autoregressive terms is supplemented with previously predicted input variables without lag. As a result, the hybrid model system operates on a forecast-on-forecast basis.

The principle of combining (integrating) models, forecasts, decision rules allows you to increase the efficiency of unitized models and, above all, forecasts developed on their basis. The combined use of other algorithms (systems of hybrid models, indicators, oscillators, etc.), parallel analysis for the presence of trends on time intervals of a different size, etc., is also effective [11, 29].

By combining algorithms is meant the summation of solutions with certain weights. To solve the problem, it is necessary to rank the algorithms, i.e. determine the contribution of each algorithm to the final solution. The weight for each algorithm and their optimal number is estimated empirically basing on the test results from historical data, based on the quality of the forecasts generated. In the papers Conroy et al. [30], Brown et al. [31], and Guerard [32], it is noted that the advantages associated with the combination of forecasts synthesized on the basis of time series models and forecasts of analysts. And this is quite justified, since this combination implies a combination of the results obtained on the basis of fundamental analysis and forecasts developed according to the models described above.

Optimization of the training sample length (modified Markowitz model). In the classical setting, the set of financial instruments under study is given by the corresponding time series of the *same length*, which represent the realizations of their quotes. In other words, the data that the investor has is a two-dimensional array, where the number of rows is equal to the number of financial instruments under consideration, and the number of columns is the amount of historical data, i.e. number of time intervals (bars). Let us introduce the following designations: $D_R(i, j)$ is realized profitability of the i th financial instrument, j is several the time interval (bar), $D_A(i, j)$ is expected (calculated by the investor) profitability of the i th financial instrument.

Then the difference between the investor's expectations and the actual profitability of the analyzed time interval for each of the assets under consideration will fully reflect the quality of the forecast.

In other words, if $D_R(i, j) - D_A(i, j) = 0$, then the investor's expectations are met exactly, but if $D_R(i, j) - D_A(i, j) > 0$, then the investor's expectations are exceeded. And vice versa, if $D_R(i, j) - D_A(i, j) < 0$, then the investor's expectations turned out to be overestimated, and the real profitability was less than the expected one. Let us take this quantitative expression as a basis for the formation of an optimality criterion.

In turn, the expected return is the average, which can be calculated for the last one, two, three, etc. the previous values of the profitability of the studied financial instruments. Continuing to increase the number of historical data to calculate the expected return from the current point in time to the past, the investor, according to the a priori selected quality criterion, will receive the optimal length of the training sample (i.e., the most acceptable number of historical data that will be used to calculate the expected return).

To implement the algorithm for optimizing the size of the training sample, we will divide the entire series of returns in chronological order into two non-intersecting sequences: training q_1 and test q_2 .

According to preceding returns from subsystem q_1 for each asset i , chosen optimum interval of length S , where at

$$D_R(i, t_{q_1-s}), D_R(i, t_{q_1-s+1}), \dots, D_R(i, t_{q_1}) \quad (2)$$

the most accurate forecasts for

$$D_R(i, t_{q_1+1}), D_R(i, t_{q_1+2}), \dots, D_R(i, t_n). \quad (3)$$

The forecast based on the test sequence data is carried out on the basis of the sliding verification principle.

The forecast error is $E_R = D_R(i, q_1 + 1) - D_A(i, q_1 + 1)$ remembered. Next, the first time point from the training sequence is discarded and the first point from the test sequence (on which the adaptive algorithm was tested) is added. The expected profitability is restated according to the updated data. The model with a given number of observations on training is sequentially tested at all time points q_2 from the test sequence.

Sequentially increasing the length of the training sequence from 1 to q_1 , the optimal length of the training sequence for each i -asset is found according to the results of the values of the following criteria:

1) Minimum search (criterion K_1)

$$K_1 = \min(S(n_{ed})) = \sum_{j=q_1+1}^n (D_R(i, j) - D_A(i, j))^2.$$

The purpose of this criterion is to find such a length of the training sequence at which the average error in the profitability forecast is minimized (without taking into account the sign of the error).

2) Maximum search (criterion K_2):

$$S_1 = \sum_{j=q_1+1}^n (D_R(i, j) - D_A(i, j)), \text{ for the cases when } E_R(i, q_1 + 1) \geq 0.$$

$$S_2 = \sum_{j=q_1+1}^n (D_R(i, j) - D_A(i, j)), \text{ for the cases when } E_R(i, q_1 + 1) < 0.$$

Then the optimality criterion K_2 can be written as: $K_2 = S_1 + S_2$.

The purpose of the criterion is to find such (optimal) length of the training sample, at which the sum of the results of the predictions favorable for the investor, minus the negative results, reaches the maximum. Let us call it the criterion of optimism.

A consistent comparative analysis of the effectiveness of the model is carried out with restrictions on the maximum permissible level of risk:

$$\sqrt{\frac{1}{q_2} \sum_{j=q_1+1}^n (D_R(i, j) - D_A(i, j))^2} \leq V, \quad (4)$$

where (4) is the standard deviation of the results of testing the model with a fixed number of historical data on its training, V is the maximum risk level set by the investor. The essence of the last inequality is to exclude from consideration those options from those considered that do not meet the set task of synthesizing the optimal investment portfolio. Parallel synthesis of the optimal training sample lengths according to two criteria is due to two reasons:

- The criterion for minimizing the sum of the square of errors is not effective when studying time series where the random component (commodities) prevails.
- Analysis by two criteria gives a "panoramic" effect, because it allows you to teach to study the object in more detail.

After identifying the optimal training length sequentially for each financial instrument, out of the total number m we obtain a one-dimensional array containing m elements (m unique training lengths for each financial asset under study) $L(1: m)$.

Lack of financial resources for carrying out the intended investments is one of the initial problems that a novice investor has to solve. He (the investor) can invest in any of the pre-planned financial instruments, it is obvious that such a decision increases the level of possible risk due to the lack of diversification. These include [11]:

- Development of forecasts of quotations, incomes based on historical data, for example, using empirical models of optimal complexity.
- The Lens algorithm allows you to build both investment and hedging strategies on mirror pairs.

Testing the above approaches showed quite acceptable results in terms of profitability with a high level of risk. In the absence of funds, Exchange-Traded Funds (ETF) or Real Estate Investment Trust (REIT) are more suitable [33].

3.3. Verification block for models, forecasts, strategies, portfolios

A portfolio manager pursuing an active investment portfolio management strategy should regularly reassess the effectiveness of investments and timely reinvestment of funds in the event of unfavorable

circumstances. Therefore, the analytical module of the multi-level integrated system for the formation of investment strategies includes activities related to the current adjustment of its structure, up to the exit from some investment areas of investment.

In order to objectively assess the current situation, a financial instrument or investment portfolio as a whole in real time with a given time step is verified on a test sequence using a sliding verification method. For this purpose, the portfolio is tested according to the first (in chronological order) observation, the results of the yield forecast for the tested portfolio are remembered, and this observation is included in the training sequence, while the very first observation is discarded. Taking into account the newly received information in the training sequence, the group of leading assets included in the portfolio is recalculated and the weights characterizing the contribution of each asset are specified. After that, the portfolio is verified by the next observation, and so on for all observations from the verification sequence. Thus, in the process of verification, self-tuning of both the group of leading financial instruments included in the portfolio and the weight is carried out.

The described method is called sliding, since the training sample after each test is shifted one position (one observation) to the right forward, so the training sample seems to slide along the test sequence following the prices. Sliding verification is as close as possible to the tasks of practice and provides more effective investment decisions through monitoring and timely correction of the structure of the investment portfolio.

3.4. Optimization block

IS Portfolio Investor is a set of rules that are designed to answer a number of questions. Here are some of them:

- Which financial instruments should you choose?
- When to enter the market with long positions?
- How often should the portfolio be monitored to assess the need to adjust its composition?

Any information system for making financial decisions has several numerical parameters. Varying of which can change the rules of the system and the time for making certain decisions. For example, in a system of two moving averages, there are two parameters: the lengths of the short-term and long-term moving averages. In Markowitz model this is the size of the training sample of historical data, which are used to calculate the expected profitability and risk level, the value of the time interval (bar) where the opening, closing quotes, etc. are determined.

To carry out computational experiments aimed both at testing the algorithmic content of the system and at selecting the most appropriate values for the required parameters, it is necessary to create a software simulator that allows simulating the behavior of a trading account controlled by user-specified rules on historical data. The simulator automates the process of evaluating the effectiveness of the IS Portfolio Investor based on a sample of historical data. To perform the listed tasks, the software simulator is equipped with a number of optimization algorithms. To date, a large number of types and methods of optimization have been developed, each with its own advantages and disadvantages, strengths and weaknesses [34-38]. Let us consider those of them that are presented in the IS Portfolio Investor.

Optimization "head-on" [39] determines the optimal solution by systematic enumeration of potential options, that is, combinations of rules, parameters, or both together. For head-on optimization, no long programs are required, usually simple nested loops are used. The parameters are changed from the start to the end value with a certain step using the loop operators (brute force or loop optimization).

Optimization methods proper. This includes, first of all, multiple regression analysis, the purpose of which is to select such regression coefficients that minimize the sum of the squared errors.

Distinguish between linear and nonlinear programming, one-dimensional and multidimensional methods for finding an extremum, with and without a derivative, etc.

Non-linear programming. If the surface of the target function nonlinear and has a highly rugged shape with recesses (gullies) and projections (ridges) of irregular shape, in this case, so-called gully methods (Method rotating coordinate – Rosenbroke method [35]) or a modified method of steepest

descent. Idea of which consists of memorizing the starting point, then perform a certain number of optimization steps using the steepest descent method, then take an optimization step in the direction from the initial point to the final one. In the multidimensional case, this method, as a rule, leads to the vicinity of the minimum in fewer iterations than the classical the method of steepest descent, which can significantly reduce the amount of computational costs for numerical optimization [11].

Linear programming. Developed for optimization problems involving linear target functions with linear constraints on parameters or input variables [40, 41]. Linear programming tasks set within the IS Portfolio Investor are solved with the help of the "Search for a Solution" add-in MS Excel.

4. Practical use

The above modules of the IS Portfolio Investor were verified for a wide class of practical problems and in most cases showed quite acceptable results:

- Several empirical models have been synthesized for the optimal complexity of forecasting the yield of such exchange commodities as fattening cattle, sugar, rapeseed, coarse rice, etc., as well as for forecasting the yield on treasury bills and changes in the consumer price index for the US stock market, etc [42].
- The constructed systems of hybrid models for forecasting the profitability of financial instruments of the US stock market and forecasting the yield of the main agricultural crops of the Russian Federation contribute to a significant expansion of the number of predicted financial instruments and, as a result, to an increase in the efficiency of investment decisions [43].
- Independent material has shown that the combination of forecasts and models reduces the forecast error to 20% [11].
- The efficiency acceptable for the needs of investment practice was demonstrated by the simplest portfolios synthesized using the Return-Risk model. The approbation was carried out on an independent material for common shares of the US and RF stock markets, analysis of the US industries, the FOREX market, the commodity market, ETF, REIT, stock, industry and commodity indices [33, 44].
- Improving the efficiency of the investment strategies adopted is facilitated by the optimization of the learning length of empirical models, the use of sliding verification and optimization of the monitoring frequency objectifies the process of updating portfolios and increases their practical significance [11, 43, 44].
- Approbation of the modified Markowitz model with a variable length of the training sample on long-term data on the profitability of the main financial instruments of the US stock market and data on the profitability of major agricultural crops in the Russian Federation made it possible to objectify the process of forming optimal portfolios based on the above information [45].

5. Conclusions

The article describes the information and algorithmic support of the Portfolio Investor IS. The practical use has shown its effectiveness, versatility, ease of implementation and a sufficient number of analytical tools to solve problems at all stages of the formation of investment strategies. In the future, as part of a cycle of work on the development of a multi-level integrated system for the formation of investment strategies, it is planned to conduct a comprehensive testing of the system modules. In addition, we plan to implement the support of continuous integration for the applied software of IS Portfolio Investor using the Orlando Tools framework [46, 47].

Acknowledgements

The study was supported by the Ministry of Science and Higher Education of the Russian Federation, project "Technologies for the development and analysis of subject-oriented intelligent group control systems in non-deterministic distributed environments", as well as the Russian Foundation of Basic Research and Government of Irkutsk Region, project no. 20 47 380002.

References

- [1] Gitman L J and Joehnk M D 1993 *Basis of investments* (New York: HarperCollins Publishers) p 629
- [2] Gullen R 1987 Patterns for Profit *Personal Investor* **11** 52–53
- [3] Morgenson G A 1987 Checklist for Stock Market Prognosticators *Forbers* **May 4** 110–114
- [4] Paulson M C 1988 You Know All About the Dow. But. Did You Know Ir`s One of the Least Useful Indexes? *Changing Times* **12** p 19
- [5] Smith P A 1987 Gearing Up: More Small Investors Turn to Computers for Assistance *The Wall Street Journal* **June 25** p 27
- [6] Bloomberg Professional Services. Available at: <https://www.bloomberg.com/professional/solution/bloomberg-terminal/> (accessed: 04/30/2021)
- [7] Refinitiv Eikon. Available at: <https://eikon.thomsonreuters.com/index.html> (accessed: 04/30/2021)
- [8] Yahoo! Finance. Available at: <https://finance.yahoo.com/> (accessed: 04/30/2021)
- [9] Tinkoff Investments. Available at: <https://www.tinkoff.ru/invest/> (accessed: 04/30/2021)
- [10] VTB My Investments. Available at: <https://broker.vtb.ru/trade/vtbinvest/> (accessed: 04/30/2021)
- [11] Gertsekovich D A 2008 *Quantitative methods of analysis of financial markets* (Irkutsk: Irkutsk State University Press) p 335
- [12] Brealey R and Myers S 2008 *Principles of corporate finance* (New York: McGraw-Hill) p 996
- [13] Ferri R 2012 *All about asset allocation. The easy way to get started* (New York: McGraw-Hill) p 993
- [14] Damodaran A 2007 *Investment valuation: Tools and techniques for determining the value of any asset* (New Jersey: John Wiley & Sons) p 993
- [15] DeBondt W F and Thaler R 1985 Does the stock market overreact? *Journal of Finance* **40** 793–805
- [16] DeBondt W F and Thaler R 1987 Further evidence on investor overreaction and stock market seasonality *Journal of Finance* **42** 557–581
- [17] Jegadeesh N 1990 Evidence of Predictable Behavior of Security Returns *Journal of Finance* **45(3)** 890–891
- [18] Jegadeesh N and Titman S 1993 Returns to buying winners and selling losers: Implications for stock market efficiency *Journal of Finance* **48(1)** 65–91
- [19] O`Shaughnessy J 2005 *What Works on Wall Street* (New York: McGraw-Hill) p 433
- [20] Graham B 2018 *The Intelligent Investor* (New York: HarperCollins Publishers) p 568
- [21] Sharpe W, Alexander G and Bailey J 1999 *Investments* (New Jersey: Prentice-Hall) p 962
- [22] Gershengorn G I 1977 Software package for building empirical differential equations *Long-term predictions of natural phenomena* pp 133–137
- [23] Fama E F and French K R 1988 Dividend Yields and Expected Stock Returns *Journal of Financial Economic* **1(22)** 3–25
- [24] Markovitz H M 1952 Portfolio selection *Journal of Finance* **7(1)** 77–91
- [25] Gertsekovich D A 2017 Formation of an optimal investment portfolio on a set of efficient portfolios *Vestnik Moscow University. Ser.: Economics* **5** 86–101
- [26] Draper N Smith H 1998 *Applied Regression Analysis* (New Jersey: John Wiley & Sons) p 736
- [27] Seber G and Lee A 2003 *Linear regression analysis* (New Jersey: John Wiley & Sons) p 562
- [28] Efroimson M A 1960 Multiple regression analysis *Mathematical Methods for Digital Computers* **1** 191–203
- [29] Nelson C R 1972 The prediction performance of the ERB-MITTPENN model of the US economy *American Economic Review* **62(5)** 902–917
- [30] Conroy R and Harris R 1987 Consensus Forecast of Corporate Earnings: Analysts' Forecasts and Time Series Methods *Management Science* **33** pp 725–738

- [31] Brown L D, Hagerman R L, Griffin P A and Zmijewski M E 1987 An Evaluation of Alternative Proxies for the Market's Assessment of Unexpected Earnings *Journal of Accounting and Economics* **2(9)** 159–193
- [32] Guerard J D 1989 Combining Time-Series Model Forecasts and Analysts' Forecasts for Superior Forecasts of Annual Earnings *Financial Analysts Journal* **1(45)** 69–71
- [33] Gertsekovich D A, Gorbachevskaya L I and Gorbachevcrfya E Yu 2020 Real Estate Funds. Quantification of investment efficiency *Problems of social and economic development of Siberia* **1 (39)** 13–17
- [34] Vasiliev F P 1967 *Numerical methods for solving problems of extreme* (Moscow: Science) p 549
- [35] Wilde D J 1964 *Optimum seeking methods* (New York: Prentice-Hall) p 267
- [36] Hadley G 1964 *Nonlinear and dynamic programming* (Boston: Addison-Wesley) p 496
- [37] Himmelblau D 1972 *Applied nonlinear programming* (New York: McGraw-Hill) p 498
- [38] Fletcher R 1987 *Practical methods of optimization* (New York: Wiley) p 436
- [39] Katz J O and McCormick D L 2000 *The encyclopedia of trading strategies* (New York: McGraw-Hill) p 387
- [40] Dantzig G B 1963 *Linear programming and extensions* (New Jersey: Princeton University Press) p 625
- [41] Kantorovich L V 1959 *Economic calculation of the best use of resources* (Moscow: Academy of Sciences of the USSR Press) p 344
- [42] Gercekovich D A, Podlinyaev O L, Tonkikh A V and Kaverzina O V 2019 Formation of a strategy for the development of investment policy of economic entities based on empirical models *Problems of socio-economic development of Siberia* **1(35)** 29–35
- [43] Gercekovich D A, Podlinyaev O L and Tonkikh A V 2021 Systems of hybrid models for forecasting crop yields as a basis for the synthesis of investment strategies *Problems of socio-economic development of Siberia* **1(43)** 19–25
- [44] Gercekovich D A, Caetano J S and Zmanovskaya O S 2020 Comparative analysis of the potential preference of various investment areas *Vestnik Moscow University. Ser.: Economics* **2** 62–76
- [45] Gercekovich D A, Gorbachevskaya E Yu and Shilnikova I S 2021 Identification of basic criteria of portfolio analysis based on the rolling verification principle *Proc. of the 1st International Workshop on Advanced Information and Computation Technologies and Systems (AICTS) (CEUR-WS Proceedings)* **2858** pp 57–63
- [46] Feoktistov A, Gorsky S, Sidorov I, Bychkov I, Tchernykh A and Edelev A 2020 Collaborative Development and Use of Scientific Applications in Orlando Tools: Integration, Delivery, and Deployment *Comm. Com. Inf. Sc.* **1087** 18–32
- [47] Gorsky S, Kostromin R, Feoktistov A and Bychkov I 2020 Orlando Tools: Supporting High-performance Computing in Distributed Environments *Proc. of the 6th International Conference on Information Technology and Nanotechnology (ITNT 2020)* (IEEE Press) pp 1–6