Adv Syst Sci Appl 2022; **04**; 126-143
Published online at http://ijassa.ipu.ru.

Raw Material Price Forecasting on Commodity Markets: Application of Expert and Quantitative Information

Zinaida K. Avdeeva[1*], Elena A. Grebenuyk[1], Svetlana V. Kovriga[1]

[1)] *Trapeznikov V.A. Institute of Control Sciences of RAS, Moscow, Russia*

E-mail: zinaida.avdeeva@gmail.com

E-mail: lngrebenuk12@yandex.ru

E-mail: kovriga@ipu.ru

Abstract: The article considers the problem of forecasting prices on the commodity market for the year ahead, broken down by months. The uncertainty level of the forecast increases with the increase of its horizon, due to changes caused by events of the external environment on the forecasting horizon. To reduce this uncertainty, we have proposed a hybrid model for the formation and correction of a monthly price forecast for the year ahead. This model uses for forecasting expert-analytical information processed using the FCM model, and data from time series of commodity market prices and macro indicators. Model of forecasting has based on the use of an ensemble of VECM and VAR models built at various time scales. For correcting the target indicator forecast on the forecasting horizon we have developed an algorithm based on signals of situation monitoring conducted on the FCM-model of commodity market situation. We demonstrate the efficiency of the proposed technique for monthly forecasting of prices for black scrap a year ahead for 2019. The accuracy of the obtained forecast we compared with naive and ARIMA forecasts.

Keywords: commodity markets; nonstationary processes; forecasting; fuzzy cognitive map, situation and digital monitoring, time series.

1. INTRODUCTION

One of the most important activities of enterprises is to reduce their expenses without losing the value of the products created for the consumer. Forecasting commodity prices for the year ahead, broken down by months, is a necessary element of enterprise procurement planning and inventory management and is crucial for reducing enterprise costs.

The purpose of our study is to build a monthly forecast of price of a certain type of for raw material or product (target indicator) in the commodity market for the year ahead.

Commodity markets, as objects of forecasting, have the following features:

- prices for various commodities follow common trends, such price trends are called co-movement [7,22];
- the product value chain determines the dynamics of price interaction between goods [21];
- the structure of the interrelations depends on the time scale in which we consider it and may change with changes in scale;
- there is a dependence of prices on macro indicators, such as the real interest rate, exchange rates, oil prices, world demand for goods, etc. [8];
- the processes of price changes are nonstationary, events occurring in the external environment, such as natural and economic disasters, political events, changes in participants' strategies in commodity markets, etc., can lead to the emergence of structural breaks in prices [13,33].

[*] Corresponding author: zinaida.avdeeva@gmail.com

Forecasting a time series several steps ahead for an extended period is a difficult problem, because the level of forecast uncertainty increases with the growth of the horizon [20]. To predict nonstationary processes, researchers use various methods: building e parametric models, singular spectral analysis, empirical mode decomposition, support vector machines, neural networks, etc. [5,10]. However, the high level of uncertainty in the dynamics of processes on commodity markets creates problems for the application of traditional forecasting models.

To reduce this uncertainty some of authors [29,31] explored different multistep forecasting strategies and shown that the most efficient for multistep forecasts are strategies with several outputs. The application of these strategies implies the preservation of stochastic dependencies between the members of the predicted sequence. Regular processes satisfy these conditions, because their dynamics can be described with sufficient accuracy by some dynamical system.

However, for processes whose dynamics change is often associated with external events, the approach based on extracting and accounting in the forecast additional information about the events of the forecast period that affect the value of the forecasted indicator is more effective.

According to the efficient market hypothesis [14], changes in the external environment are reflected in the quantitative data used in forecasting and thus taken into account in the forecasts. However, complex interactions of supply and demand, and extreme and unexpected events in the external environment are not predictable from commodity market data and are one of the sources of forecasting errors [11].

To take into account environmental events in the forecast, the authors of [32] proposed a methodology of multiscale forecasting, using search systems data in different time scales. To assess the impact of extreme events on the volatility of crude oil prices, the authors of [34] developed an estimation methodology based on the decomposition of the series by empirical modes at different time scales. The experiments have shown the promise of the method for analyzing possible changes in the price series under the influence of extreme events.

Modern approaches to forecasting point to the need to use expert information. When constructing medium-term (and long-term) forecasts, information about future events that change the dynamics of the target indicator in the interval between forecasts is not available in the data at the time of making the forecast.

If the price dynamics on the forecasting horizon changes under the influence of external environment events, the forecasting model becomes unsuitable.

For timely replacement and adjustment of the forecasting model, monitoring is necessary, the purposes of which are: detection of changes in time series (structural shifts), revealing of the reasons, which caused or may cause the occurrence of structural shifts, formation of scenarios of possible development of the situation, evaluation of the power and duration of the forthcoming changes.

The time series monitoring algorithms detect the structural shifts of the forecast indicators and determine the type of these changes (change in the trend, volatility, the type of non-stationarity of the process). However, the results of digital monitoring do not allow solving such problems as:

– identifying the causes of structural shifts;
– forecasting further development of the situation which led to the occurrence of a structural shift;
– forecasting the situation which may lead to a structural shift.

Situation monitoring based on processing and analysis of expert information is needed to support the solution of the above tasks.

Therefore, improving the quality of process forecasts in a changing external environment is impossible without the account of expert and analyst judgments [2,9,15,28]. The task of decision support systems for forecasting is to develop tools for extracting, integrating, and processing expert knowledge and generating signals to incorporate it into forecasts obtained by processing quantitative data.

In this study, we use methods for analyzing and modelling expert information based on fuzzy cognitive maps. Fuzzy Cognitive Maps (FCM) refer to a class of models representing expert knowledge and judgments. They are used to analyze and model Ill-structured situations. These are situations involving many elements of various nature, and the relationships between the elements can be both quantitative and qualitative [1,12,26].

The use of FCM in production sphere is associated with the support of managerial decision-making in situations of uncertainty or lack of quantitative data to solve the problems posed. Typical FCM applications are the following examples:

– analysis of factors supporting collaborative planning, forecasting and replenishment to ensure effective supply chain management [6];

– risk analysis of ERP (Enterprise Resource Planning) maintenance in the system lifecycle [23];

– risk analysis in a production system, for which the possibility of loss of profit is considered a key risk factor [33];

– factors' analysis of influence on planning of production in manufacturing plants using the FCM of multi-criteria decision making [30];

– selection and justification of technological solutions for creating final products with specified goal characteristics, where the technology solutions are products and technologies-components of the final product [4], etc.

In accordance with the purpose outlined in the article, the contribution of this study is as follows:

1. We propose a hybrid model for the formation and correction of the monthly price forecast for the year ahead, combining the information obtained from the processing of expert judgments through FCM, and the data of time series of commodity market prices and macroeconomic indicators. The developed hybrid model performs the following functions:

– processes expert judgments and analytical information and simulates scenarios of possible consequences of the impact of significant events of environment on the change of the target indicator by means of FCM – a formalized model of causal influences between the system-forming factors of the situation and the target indicator;

– forms a monthly price forecast for the year ahead on the basis of the information obtained from the processing of expert judgments and time series data as a result of applying the ensemble of multidimensional time series models and taking into account the cause-and-effect effects between the factors of the situation and the target indicator.

2. We have developed an algorithm for correcting the target indicator forecast on the forecast horizon based on signals

– situation monitoring implemented on the basis of analysis and modelling methods on FCM [3,4];

– digital monitoring carried out by methods of sequential analysis of nonstationary processes [19].

3. We tested the effectiveness of the hybrid model and algorithms for generating and correcting on monitoring signals forecast when generating a monthly forecast of prices for ferrous scrap for the year ahead. The constructed forecast showed a significant increase in forecasting accuracy compared to the plan-fact ("naive") forecast and the forecast on the ARIMA model.

2. PROBLEM STATEMENT, ITS FEATURES AND GENERAL SOLUTION PRINCIPLE

We consider the problem of forming and adjusting the monthly forecast of commodity prices for the year ahead. The forecast indicator (target) $Y = \{y_1, y_2, ..., y_t, ...\}$ is a non-stationary time series, whose properties may change under the influence of:

- factors affecting the domestic market of the forecasted indicator;
- factors affecting changes in world prices for raw materials and products;
- factors that characterize the global economy;
- factors, characterizing the state policy of commodity markets regulation.

We call these factors system-forming, because in their unity they form a system that reflects a holistic view of the situation in the context of the problem under consideration. The forecasted target at time t may depend on its previous values; previous values of other processes; significant events in external environment, uncontrolled and unpredictable factors described by a random sequence.

We call significant events (*Inf*) events that can lead to changes in the characteristics of the process: disruption of relationships, gradual or abrupt changes in trends, increase or decrease in volatility.

Required to build:

a) an algorithm for monthly forecasting of the target indicator $Y = \{y_1, y_2, ..., y_t, ...\}$ for the year ahead;

b) a monitoring system that includes digital monitoring of price time series and situation monitoring to detect events and system-forming factors that change the dynamics of processes;

c) an algorithm for correcting the forecast of the indicator $Y = \{y_1, y_2, ..., y_t, ...\}$ on the forecasting horizon based on monitoring signals.

d) assessing the quality of the forecast obtained, built according to algorithms (a)-(c).

.2.1. General principle of the solution

We divide the problem of constructing a monthly forecast of the values of the target indicator for the year ahead into two stages. Stage 1 is a building a forecast for the year ahead (performed before the start of the forecast period); stage 2 is a correction of the forecast on the forecasting horizon based on situation and digital monitoring signals.

To assess the state of the external environment and the strength of the influence of the factors that may change the values of the forecasted target indicator we build a formalized description of the current situation in the form of the FCM. To receive information about the effects of changes in the external environment that are not present in the time-series at this moment we use situation monitoring, according to the results of which the FCM generates signals about possible changes in the dynamics of the forecasting target indicator. To detect changes (structural shifts) in the processes and determine the type of these changes (trend change, volatility, type of process non-stationarity) we perform digital monitoring.

Figure 1 shows a general scheme of the forecasting process. Stage 1 consists of three steps.

Step 1.1. Based on the constructed description of the situation in the form of FCM we form groups of system-forming factors (SF-groups) that have a direct or indirect impact on the forecasted target indicator. The basis of FCM is a graph, the vertices of which represent the factors of the situation under study; the edges are the direct causal effects of some factors on others; any factor with all the factors of influence on them forms a node of FCM.

The constructed FCM allows you to formalize the process of transferring expert information to the prediction algorithm (Figure 1).

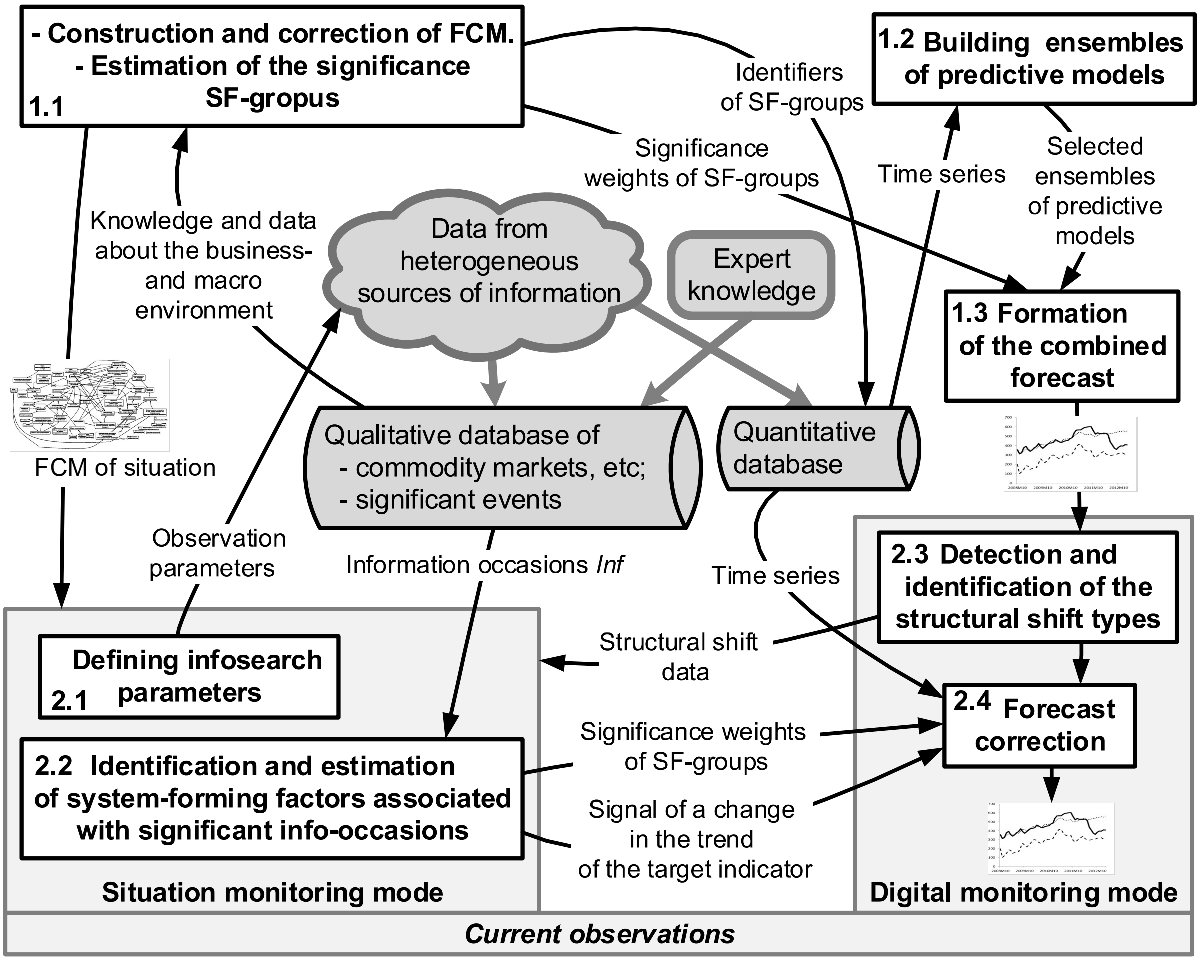

Fig. 1. General scheme of the forecasting process

Each of the SF-groups describes one aspect of the situation under study (production factors related to the predicted indicator, market factors of the business environment related to supply and demand in commodity markets, macro-environment factors related to financial and economic indicators, factors of influence of state regulators, etc.).

Separate SF-groups correspond to time series of the data, changes in which arise under the influence of system-forming factors of the given SF-group. The formation of SF-groups is aimed at expanding the information space of the search for quantitative data to replenish the quantitative database and its structuring by distributing the series by the formed SF-groups. We calculate the weights of SF-groups according to the significance of their influence on the target indicator and pass them to step 1.3 to calculate the final forecast.

Step 1.2. To build ensembles of forecast models we use monthly data. Each model includes a target indicator and time series of a certain SF-group. For each SF-group for which the corresponding time series are in the database, we compute a forecast on the ensemble of models. We selected only those time series that are Granger-causes Y [17]. Since the series are non-stationary, integrated of the first order, we build a model describing the target indicator using the differences of the series. If there is cointegration between the series, we build VEC models to trace the long-run dependencies.

Step 1.3. We form a combined forecast based on the forecasts obtained in step 1.2, and SF-group weights calculated in step 1.1. The proposed strategy allows us to obtain more accurate forecasts [16,22,32].

Stage 2. To account for the influence on the built forecast of environmental events on the forecasting horizon, we adjust it according to the signals received from the results of situation and digital monitoring. The signals of actualization of the forecast correction process in the mode of current observations are of different nature:

– signals triggered by significant events in the external environment, which may coincide with digital monitoring signals or may serve as a warning of possible changes that will occur in the lagged observations;

– signals about the presence of changes in the predicted indicator and the related price series (different for each type of change).

3. ANALYSIS OF COMMODITY MARKETS BASED ON HISTORICAL DATA AND EXPERT KNOWLEDGE AND FORMATION OF A TARGET INDICATOR FORECAST

3.1. FCM construction and estimation of the SF-groups' significance

We conduct the structuring of the subject knowledge domain to form a holistic representation of the situation associated with the impact of the external environment on the dynamics of the predicted process.

According to the results of the structuring, we obtain a conceptual model of the situation, which includes the main subject-oriented concepts and related subject knowledge, characterizing the production cycle and the value chain of the target-related products, as well as its business- and macro environment in the context of the forecasting problem set by us.

Taking into account, the heterogeneity of environmental factors affecting the predicted processes, we used expert knowledge (in the mode of direct interaction with experts in the analysis of the commodity market) and analytical information from heterogeneous sources as input data for structuring.

Based on the conceptual model of the situation, we identify, together of experts, the system-forming factors influencing the predicted process and establish the relationship between them and the target indicator, i.e., formalize the selected subject knowledge in the form of the FCM.

The constructed FCM allows formalizing the process of transferring expert information into the forecasting algorithm (Figure 1).

In Figure 2, we show a fragment of the FCM associated with the commodity market (grey arrows with ellipsis show the fragment boundaries in the full FCM). This map includes 59 system-forming factors, including the $\tilde{y}$ factor associated with the target forecasted indicator Y (purchase prices for black scrap metal (BSM)), which are related by cause-effect relationships.

We represent factors as variables $\tilde{X} = \{\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_{N-1}, \tilde{x}_N = \tilde{y}\}$, described by certain sets of values (Figure 2).

Let $K_f(\tilde{X}, W, f\}$ be an FCM of the situation (S) in commodity markets related to a target indicator Y, where $\tilde{X} = \{\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_{N-1}, \tilde{x}_N = \tilde{y}\}$, $\tilde{y}$ is a factor characterizing Y; $W = [w_{ij}]$ is the $N \times N$-matrix of factors mutual influence, where the weight $w_{ij} \in [-1;1]$ determine the degree of expert confidence that activity of factor $\tilde{x}_i$ causes on the factor $\tilde{x}_j \in [-1;1]$; f is a function of factors' activation at any discrete time t

$$f : a_i(t+1) = a_i(t) + \sum_{j \in I_i} w_{ji} \times \Delta a_j(t) + g_i(t), \tag{1}$$

where $a_i(t), a_j(t)$ are factor activity values $\tilde{x}_i$ and $\tilde{x}_j$ at the moment t, $\Delta a_j(t) = a_j(t) - a_j(t-1)$; I_i is a set of factor codes $\{\tilde{x}_j\}$ that directly influence factor $\{\tilde{x}_i\}$; $g_i(t) = g_i(0)$ for $t = 0$ and $g_i(t) = 0$ for $t \neq 0$, $g_i(0) \in [-1;1]$.

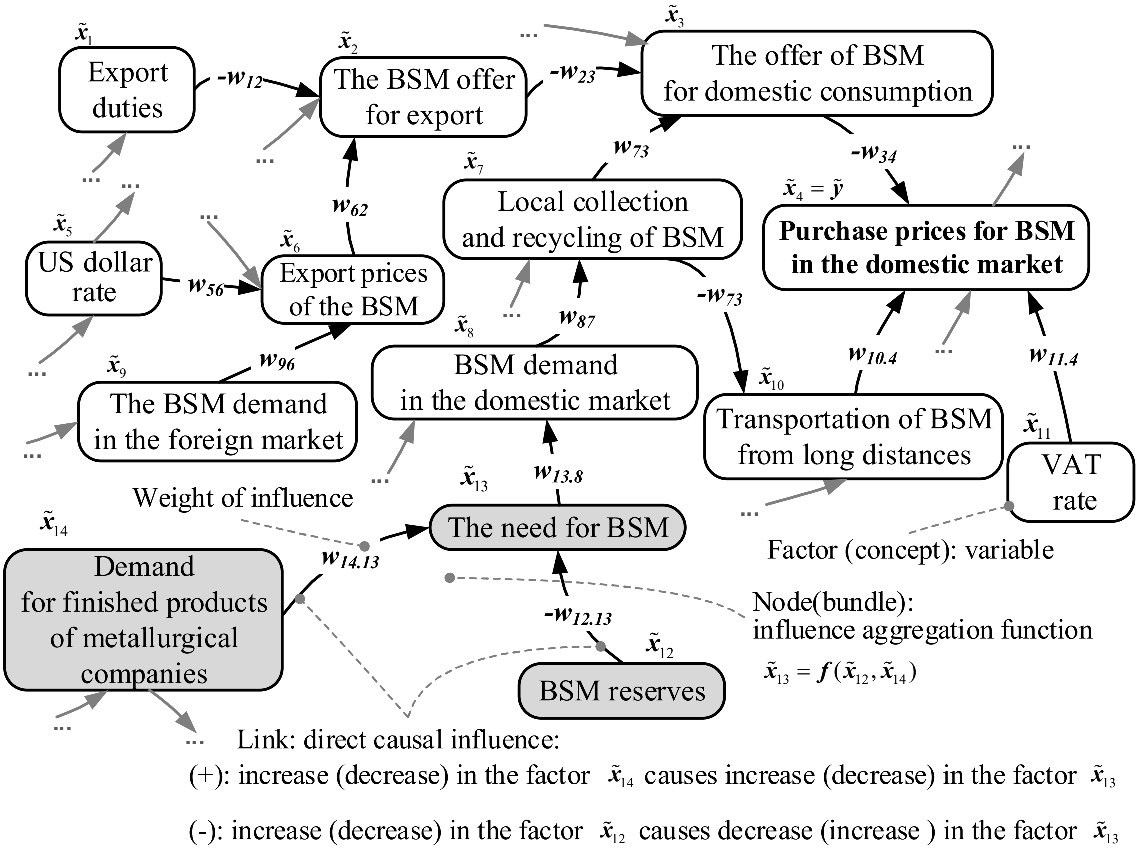

Fig. 2. The fragment of an FCM associated with the commodity market

We define the state $A(t+1)$ of the situation S at time $t+1$ as the result of the influence of all factors $\tilde{X}$ on the interval $[0,t]$:

$$A(t+1) = Q(t)(A(0) + G(0)) \tag{2}$$

where $Q(t) = E_N + W + W^2 + ... + W^t$, $A(0)$ is an initial state of factors $\tilde{X}$, $G(0)$ is a vector of external impacts on factors at $t=0$, W^k is a matrix of influences to k-th degree, $k = 1,...,t$, E_N is the identity matrix.

The property $Q = \lim_{t\to\infty} Q^t = (E_N - W)^{-1}$ allows us to identify the steady states by factors. Then Q is the matrix of integral influences of FCM factors, the elements of which characterize all the direct and indirect influences between factors.

Q-matrix analysis is based on the method of [3], which allows

– to determine the cumulative (direct and indirect) influence of factors on the target;
– determine the nature (favorable or unfavorable) and weight of the cumulative influence of one or another factor (taking into account the current dynamics of this factor) on the favorable dynamics of the target indicator;
– to estimate efficiency of influence of positive and negative factors on a target indicator and their ranking (taking into account their inclusion in different SF-groups);
– to assess the activity of factors with a high rank of influence on the target indicator.

Based on this data, we calculate for each SF-group the integral positive q^+ and negative q^- influence on the target indicator of all factors included in the SF-group. Then the total influence of the SF-group is $q = q^+ + q^-$.

According to estimates of type $\{q_i\}$, we calculate the normalized weights of the influence of SF-groups in the form

$$C_i = \frac{q_i}{q_1 + q_2 + \ldots + q_m}, \tag{3}$$

where q_i is the total integral influence of SF-group, m is number of SF-groups.

We transfer the weights of SF-groups to the numerical forecasting block 1.3 to account for the influence of systemic factors of the situation on the target.

3.2. Construction of model assembles, and formation of a target indicator forecast

From the structured time series database (step 1.1 on Figure 1), we select the sets of series $\langle \breve{F}_i, \bar{Y} \rangle, i = 1, 2, \ldots, m,$ where $\bar{Y}$ is the series of target indicator values, $\breve{F}_i = (\bar{X}_1, \bar{X}_2, \ldots, \bar{X}_{ni})$ are the series of SF-group with number $i, i = 1,2,\ldots,m$. We break down the available set of time series $\breve{F}$ according to the structure of the available SF-groups: $\breve{F} = (\breve{F}_1, \breve{F}_2, \ldots, \breve{F}_m),$. We obtain the sets of series $\langle \breve{F}_m, \bar{Y} \rangle$, each of which describe the dynamics of the target indicator as a function of its past values and the system-forming factors of its corresponding SF-group.

The process of building a forecast includes the steps below, which we do for each set of $\langle \breve{F}_i, \bar{Y} \rangle$, and for each time scale (month, quarter, half-year, year). The models we apply vary for different time scales, because of changes in the relationships between successive observations and decreases in the size of the training sample as the scale increases.

Step 1. We apply structural shift detection algorithms to series $\bar{Y}$ [19], to determine the presence of trends, their direction, and the moments of their changes.

Step 2. We check if the statistical properties of the series of the set $\breve{F}_i$ correspond to the properties of the target $\bar{Y}$. We include in the forecast model only those series from $\breve{F}_i$, which have the order of integration and seasonality index equal to the corresponding characteristics of the series $\bar{Y}$.

Step 3. Granger causality check [24] of the series from $\breve{F}_i$ relative to the target indicator $\bar{Y}$. For each group, we form variants of the VAR models that include the target indicator and the series from $\breve{F}_i$.We exclude a series $\bar{X}_i \in \breve{F}_i$ from further analysis if it is not Granger causal for the target indicator.

Step 4. Building models. For series with a unit root, regression models linking the past values of the series $\bar{Y}$ with its current values and with the values of the series from $\breve{F}_i$, describe a spurious regression if there is no cointegration between the series included in the model [17]. For these series we build VAR-models, taking the necessary number of differences in advance, but such VAR-models reflect only short-term dependence. If there is cointegration between the series, we build a VEC model to build a forecast that allows us to take into account the long-term dynamics of the series.

The construction of VAR and VEC models is associated with the estimation of a large volume of parameters: variables, lags of variables, constants and coefficients of trend, coefficients of error correction equations - correction term (in the case of VEC models).

The volume of estimated parameters increases with the number of variables included in the model, and large samples are required to obtain consistent estimates of model parameters. We consider series in which the length of intervals without structural shifts rarely exceeds 4-5 years, which even for monthly data is about 50-60 points. With increasing data scale (transition to 2-month, quarterly, semi-annual data, etc.), the sample size decreases significantly, therefore, in order to obtain statistically significant results, it is necessary to reduce the number of regressors, included in the model.

According to the results of our experiments on real data, we chose models with two and three variables for monthly data, and models with two variables for quarterly data. Analysis of the dynamics of annual and semi-annual data we performed according to the ARIMA models built for the series $\bar{Y}$.

In order to take into account, the possible effects on the values of the target indicator of all the series of the group that have an impact on it, we use an ensemble of models, because as the results of research [25] show, estimates by ensembles of models allow us to obtain more accurate forecasts.

Step 5. For each of the SF-groups we construct forecasts for ensembles of models at different time scales. Forecast construction for individual SF-groups on a certain time scale includes:

– evaluating the quality of the built models by Schwartz 's criterion, $r_1, r_2, ..., r_{k_j}$, where r_i is the value of the Schwartz criterion of the model number i, k_j is the quantity of models in SF-group j, $j = 1, 2, ..., m$; calculation of the average value of the Schwartz criterion for each SF-group: $\breve{r}_j = \frac{1}{k_j} \sum_{i=1}^{k_j} r_i$;

– calculation of forecast weights $w_1, w_2, ..., w_{k_j}$ for each of the models in the SF-group by the formula: $w_k = \frac{p_j}{\sum_{i=1}^{k_j} p_i}, k = 1, 2, ..., k_j$, where $p_i = \frac{1}{r_i}$, j is a number of SF-group;

– formation of the forecast by the ensemble of models, built using time series from a separate SF-group, by the formula:

$$\breve{Y}_j = \sum_{i=1}^{k_j} w_i \breve{Y}_j^i \tag{4}$$

where $\breve{Y}_j^i$ is the vector of forecasts for the model with number i of SF-group j, $j = 1, 2, ..., m$. The sizes of vectors are 12, 4, 2, 1, respectively, depending on the scale. For each SF-group, we build models using monthly, quarterly, semi-annual, and annual data. As a result, we get 12 monthly forecasts, four quarterly forecasts constructed using ensembles of models, two semi-annual and one annual forecasts constructed using a time series of target indicator.

Step 6. Correction of forecasts of SF-groups with respect to time scales. We adjust the monthly forecasts $\breve{Y}_j$ (equation (4)) for each SF-group j using forecasts made at other time scales.

We adjust the monthly forecasts for March and September by quarterly forecasts, the forecast for June with quarterly and semi-annual forecasts, and the forecast for December with quarterly, semi-annual, and annual forecasts. We calculate the adjusted forecast as the weighted average of the forecasts included in it, the weights of which are inversely proportional to the errors of the forecasting models calculated at the interval of model building.

Step 7. Final forecast with expert information. We construct the final forecast as a weighted average of the SF-group weights calculated using equation (3) (Section 3.1): $Y_F = \sum_{i=1}^{m} C_i \breve{Y}_j$, where $\breve{Y}_j$ we calculate by equation (4).

4. FORECAST CORRECTION IN THE FORECAST INTERVAL

4.1. Application of situation monitoring to identify significant environmental events

Situation monitoring signals the effects of changes in the external environment that are not currently reflected in the time series data.

Situation monitoring implements the following functions

– regular (or at the request of digital monitoring) tracking and preliminary filtering of information from diverse sources (specialized resources that publish weekly analytical reports, reviews and summaries of the dynamics of commodity markets; news portals and media that provide information about various events);

– processing of filtered information about significance events (information occasions) by forming scenarios of the situation development to assess the significance of the consequences of these events on the predicted process (target indicator);

– recalculation of SF-groups weights and/or generation of a signal associated with the identification of a change in the trend of the target indicator based on the results of scenario modelling.

The basis for tracking and filtering information is the FCM of the situation, which we consider as a semantic model of observation in information sources on significant events and topics. This approach allows us to organize a directed search for information (step 2.1 in Figure 1) on the selected system-forming factors of the FCM, taking into account their relationships with each other and the target indicator.

The processing of filtered information is as follows. Upon detection of an event *Inf* (information occasion), which, according to a preliminary expert assessment, is significant for changing the target indicator and/or other system-forming factors of FCM, we model the "what if" scenario.

To form the scenario $S^{\inf}$, we determine which system-forming factors $\{\tilde{x}^{\inf}\}$ can be used to reflect info-occasion *Inf* in $S^{\inf}$.

Next, we, together with experts, evaluate the activity of change factors $\{\tilde{x}^{\inf}\}$ (their increasing or decreasing and how much) at the time of detection of an info-occasion. These values are the initial data for modelling (according to Equation (1)) the net effect of factors $\{\tilde{x}^{\inf}\}$ on the target indicator Y.

We get the resulting state, which is a qualitative forecast for the $S^{\inf}$ scenario. This state determines at a qualitative level the change in the activity of the $\tilde{y}$ (and other factors of the FCM) in the $S^{\inf}$ scenario. The obtained value of the factor $\tilde{y}$ allows us to assess whether *Inf* affects the change in the target indicator Y.

After evaluating $\tilde{y}^0$ – value (activity) of the factor $\tilde{y}$ at the initial moment t, we compare it with the resulting $\tilde{y}^*$ – value in the $S^{\inf}$ scenario.

If $\tilde{y}^* \neq \tilde{y}^0$, then the target indicator Y has changed.

If $sign(\tilde{y}^0) = sign(\tilde{y}^0) \times sign(\tilde{y}^* - \tilde{y}^0)$ then the trend of Y does not change (the trend change indicator $In(Y) = 0$).

Otherwise $sign(\tilde{y}^0) \neq sign(\tilde{y}^0) \times sign(\tilde{y}^* - \tilde{y}^0)$, we fix a trend change Y. Moreover, if $sign(\tilde{y}^0) = +1$ and $sign(\tilde{y}^0) \times sign(\tilde{y}^* - \tilde{y}^0) = -1$ then the increase of Y is replaced by a decrease (trend change indicator $In = -1$); in the case of $sign(\tilde{y}^0) = +1$ and $sign(\tilde{y}^0) \times sign(\tilde{y}^* - \tilde{y}^0) = +1$, the decrease of Y is replaced by an increase ($In(Y) = +1$).

For recalculation the weights $\{C_i\}$ of SF-groups based on the results of the $S^{\inf}$ scenario, we use the formula

$$C_i = \frac{c_i^{\inf}}{c_1^{\inf} + c_2^{\inf} + \ldots + c_p^{\inf}},$$

where $c_i^{\inf}$ is an estimate of the significance of the SF-group i influence on Y, taking into account the activity of system-forming factors in the $S^{\inf}$ scenario, p is number of SF-groups in which there are active factors associated with the info-occasion *Inf*.

Each $c_i^{\inf}$ estimate is calculated as the following sum $\sum_{j \in I_i} q_j^{\inf} \times a_j^{\inf}$, where $q_j^{\inf}$ is cumulative influence of the active factor from the SF-group i on $\tilde{y}$ and $a_j^{\inf}$ is its degree of activity in the $S^{\inf}$ scenario, I_i is a set of active factor codes for the SF-group in the $S^{\inf}$ scenario.

The weights $\{C_i\}$ of SF-groups characterize the contribution of each group to the change in the target indicator, taking into account the influence significance of active factors in these groups on Y and the activity degree of manifestation of these factors in the $S^{\inf}$ scenario. In other words, only those groups participate in the formation of weights, in which the system-forming factors associated with the info-occasion *Inf* were activated.

The obtained weights $\{C_i\}$ of SF-groups and the trend change indicator $In(Y)$ are transferred to step 2.4 of forecast correction (Figure 1).

4.2. Correction of the forecast based on the results of digital monitoring

If the prediction model matches the process, then we add new observations to the prediction formula when they arrive, to reduce uncertainty and increase the accuracy of the prediction. However, the forecast interval is large enough to assume that events may occur within it, leading to drastic changes in the process and requiring correction of the forecasting model. Therefore, after making a forecast, it is necessary to monitor the situation and the forecasted process in the macro - and business environment. We consider a generalized algorithm for monitoring the situation and the signals it generates in Section 4.1.

Based on the results of monitoring performed during the forecasting period, we adjust the forecast on the forecasting horizon in the following cases:

- when we receive new data at each forecast time horizon: month, quarter, six months;
- when we get signals from the monitoring algorithms of situation using FCM;
- when we receive the signals from time series monitoring algorithms.

To successfully adjust the forecast based on the results of detected changes, it is necessary not only to detect the change, but also to determine its type and magnitude. To improve the quality of the numerical forecast, we apply algorithms for detecting changes in series level, volatility, trend direction and relationships between time series: Granger causality, and cointegration [18,19].

Rules for adjusting the forecast.

If the drift has changed in the monitored target indicator, we adjust the deterministic trend coefficient, if the process volatility has changed or we detected deviations in the model coefficients, we recalculate the model.

If the relationship between the predicted indicator and the regressors in any of the models has changed (Granger causality has been violated), then we exclude this model from the forecast formula. If cointegration is broken for the rows included in the VECM model, we replace the VEC model with a VAR model based on differences.

If we find that the cointegration between the series in the VECM model is broken, we replace the VEC model with a difference-based VAR model.

5. RESULTS OF EXPERIMENTAL STUDY

The results of the implementation of the proposed technique we have demonstrated on the example of building a monthly forecast for 2019 prices for ferrous scrap in the commodity market.

Stage 1. Because of the structuring and formalization of the subject area, we developed a FCM, the core of which are the key parameters of the scrap value chain: the price of primary and secondary raw materials, production cost and demand for the final product and system factors affecting changes in the price of scrap. We identified four groups of series corresponding to four groups of system factors: primary (SF1-group) and secondary (SF2-group) raw materials for the creation of metal products, finished metal products (SF3-group), macroeconomic indicators (SF4-group).

The 2019 forecast was constructed using monthly data (January 2016 - December 2018), quarterly, semi-annual, and annual data (2010-2018). For each SF-group, we built a forecast by ensembles of models, taking into account the dynamics of the target indicator and related series at different time scales. We combined the constructed forecasts in the final forecast, taking into account the influence of SF-group factors on the dynamics of the current situation (section 3.2).

Figure 3 shows for comparison the actual prices for scrap metal for 2019, the monthly forecast for the year ahead obtained at stage 1, the "naive" forecast (for which the forecasts for January 2019, February 2019, ..., etc. are equal to the actual prices in January 2018, February 2019, ..., etc.).

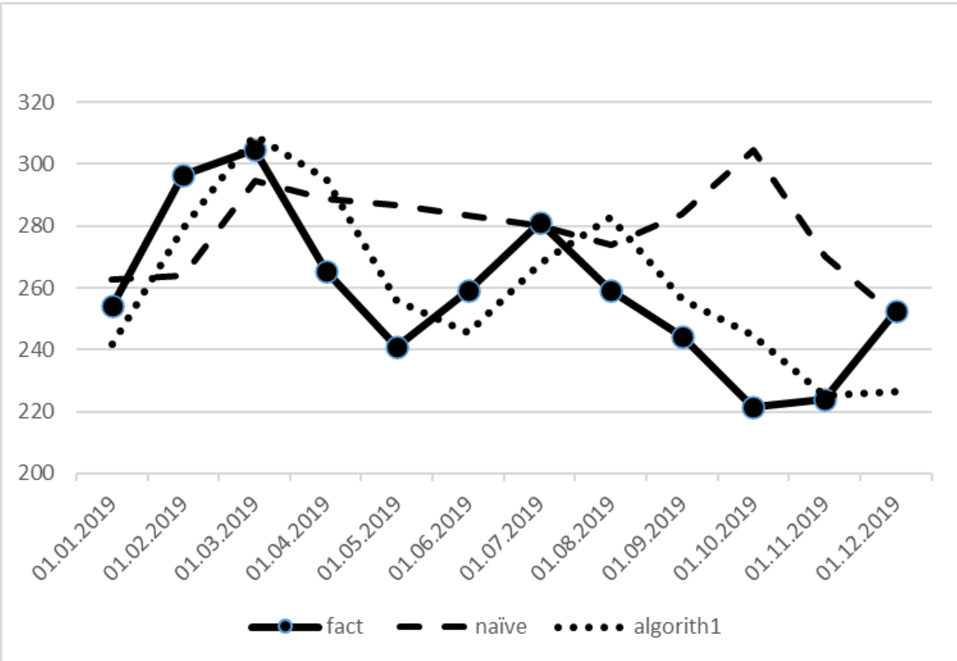

Fig. 3. The result of the forecasting on Stage 1 and naïve.

Stage 2. The algorithm adjusts the forecast during 2019 based on signals of situation and digital monitoring signals. As part of the situation monitoring information on commodity markets and other events of the external environment in 2019, we conducted a month-by-month analysis.

When detecting events (information occasions), which according to preliminary expert evaluation are potentially significant for the change in the forecasting indicator, we modeled $S^{\inf}$ scenarios "what happens if" to assess the significance of the consequences of these information occasions on the forecasting process now of their occurrence. In forming the scenario, we correlated each information occasions with the system-forming factors of the FCM. A fragment of the situation monitoring log for 7 months is presented in Table 1, which contains signals to be transmitted to the digital forecasting correction block.

Let us explain the formation of the trend change signal and the formation of weights of significance of groups by the examples of January 2019, March 2019 and August 2019 events.

Figure 4 demonstrates the initial data and the results of the scenario modeling for the formation of the signals by monitoring the events of January 2019.

For example, in January 2019, we identified a potentially significant information factor - a fire at a large steel plant in Brazil and its shutdown, which caused a sharp increase in steel prices in the U.S. market. The baseline scenario factor activity for this information trigger is given on a linguistic scale (0.2 - weak increase, -0.2 - weak decrease, 0.5 - average increase, etc., see the legend to Figure 4).

Table 1. Situation monitoring results: Signals generated from the results of scenario modelling for transmission to the forecast correction block

Observation period		*01'19*	*02'19*	*03'19*	*04'19*	*05'19*	*06'19*	*07'19*	...
1		2	3	4	5	6	7	8	
$In(Y)$		+1 Δ = 0,3	0	-1 Δ = -0,1	0	+1 Δ = 0,1	-1 Δ = -0,1	+1 Δ = 0,3	
$sign(\tilde{y}^0)$			+1	+1		-1	+1	-1	
M_i	C_i^0	C_i^1	C_i^2	C_i^3	C_i^4	C_i^5	C_i^6	C_i^7	
SF1	0,4	0,7	0,6	0,3		0,3	0,7	0,4	
SF2	0,3		0,4	0,2		0,2		0,6	
SF3	0,2			0,5		0,5	0,3		
SF4	0,1	0,3							

Based on the results of the simulation, we formed a trend change signal for the scrap price $In(Y)$ in January 2019, and adjusted the significance weights for the influence of SF-groups (see column 2 in Table 1 and the histogram on the right in Figure 4).

In March 2019, the situation monitoring algorithm identified changes in demand for steel products in Europe and the introduction of US duties on imported steel products. Based on these data, the qualitative forecast for the FCM ($S^{\inf}$ scenario) showed that in March the significance of the quantitative forecast for the group of factors reflecting demand for finished products (SF3-group) is the highest (see column 2 in Table 1).

In addition, in August 2019, the numerical monitoring algorithm revealed a lack of relationships between causal factors and target indicator in some forecast models and, as a consequence, a violation of cointegration in them.

Based on changes identified (March 2019, August 2019), the algorithm assigns the maximum weight (March 2019) to the group of factor models reflecting demand for finished goods, and replaces models with co-integration violations with other models from the same SF group (August 2019).

Thus, given the signals generated by situation monitoring (Table 1), the numerical prediction algorithm recalculates for each SF-group the prediction for ensembles of models.

Figure 5 demonstrates the results of forecasting in two variants: the forecast at Stage 1 (the broken line "Stage 1 (algorithm 1)") and with its correction by signals from monitoring carried out on the forecast horizon (the broken line "Stage 2 (algorithm 2)").

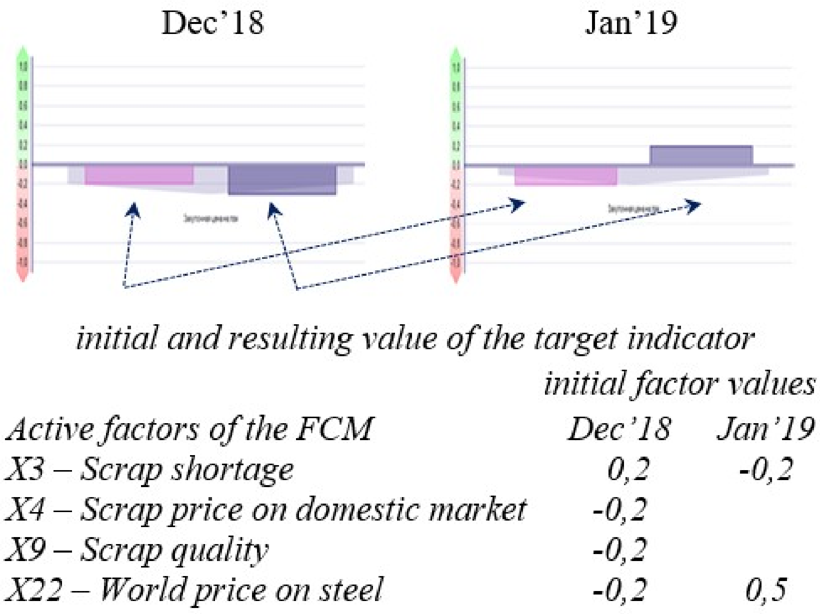

Fig 4. The examples of scenarios: results for Dec'18 and Jan'19

We compared the results of forecasting with the following algorithms: "naive" forecast, forecasting using ARIMA autoregressive model, forecast generated at Stage 1 without correction at forecast horizon (Stage 1) and algorithm with correction at forecast horizon (Stage 2). The following metrics were used to assess the quality of the prediction:

mean absolute error in percent (MAPE) and root mean square error (RMSE) to assess the accuracy of the prediction:

$$MAPE = \frac{100}{n} \times \sum_{i=1}^{n} \left| \frac{y_t - \hat{Y}_{t|t-h}}{y_t} \right|, \ RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_t - \hat{Y}_{t|t-h})^2} \ ;$$

mean percentage error (MPE) to assess the systematic error

$$MPE = \frac{100}{n} \times \sum_{i=1}^{n} \left(\frac{y_t - \hat{Y}_{t|t-h}}{y_t} \right),$$

where y_t is the value of the forecast indicator at time t (monthly), calculated with the horizon h (2019 year) at time $t-h$ (2018 year).

Comparison of forecasting errors was carried out according to the following algorithms: naive forecasting (Naive), forecasting by ARIMA model. The results are presented in Table 2.

The experimental results presented here demonstrate the advantage of our technique. At step 1, the forecast accuracy increases almost twice as much as the "naive" forecast and ARIMA, and the bias, 10.5 and 3 times, respectively. Applying correction on the forecast horizon reduces the absolute percentage error by 20% and the bias by a factor of 5 compared to the forecast constructed from data up to December 2018. The results shown in Table 2 demonstrate the advantage of the proposed forecasting technique relative to naive and ARIMA forecasts by the criteria of accuracy and bias of forecast.

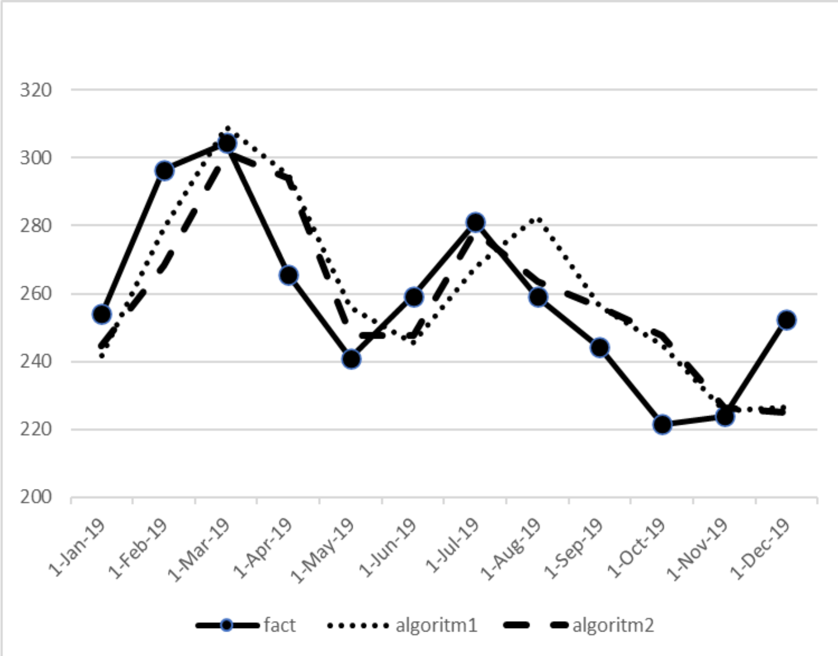

Fig. 5. The result of the forecasting on Stage 1 and Stage 2

Table 2. Average error of the annual forecast with a monthly split built by different algorithms

Algorithm	*MAPE* 2019	*RMSE* 2019($)	*MPE* 2019
Algorithm 2	5,29	17,09	-0,2
Algorithm 1	6,26	17,9	-1,09
ARIMA	12,5	38,57	-3,5
Naive	11,2	35,14	-9,67

6. CONCLUSION

This article considered the problem of forecasting commodity prices for the year ahead, broken down by months. The purpose of this study was to form a monthly forecast of the price of a certain type of raw material (target indicator) on the commodity market for the year ahead.

When constructing medium- and long-term forecasts, information about future events that change the dynamics of the forecasted value in the interval between forecasts is not available in the data at the time of the forecast. For such situations of uncertainty, we proposed a forecasting technique based on monitoring the situation and regularly including expert information in the generated digital forecast over the forecasting interval.

This technique includes

– structuring and processing of expert judgments and information from heterogeneous sources through the use of following instruments:

a) the FCM, which reflects the structure of relationships (causal influences) between system-forming factors that characterize the main processes in the studied situation on commodity markets;

b) scenario modelling on FCM of possible consequences the influence of significant events in the external environment on the change in the target indicator;

– formation of a monthly price forecast for the year ahead, based on information obtained as a result of processing expert judgments and time series data by constructing ensembles of VECM, VAR models in different time scales.

This technique allows us to automate the process of systematic processing of expert and analytical information and its transmission in digitized form for inclusion in a forecast based on quantitative data. At the same time, we consider modelling on the FCM of situation as a driver (driving force) of the forecasting process, which ensures the directed and regular inclusion of expert information in the formation of the forecast and its correction in the forecasting interval. It is important to note that the structure of representing the mutual influence of factors in the FCM is quite flexible, which allows us to quickly make changes to it (adding factors, links) to reflect new aspects of the situation caused by external environment events.

The results of the experimental study demonstrate the performance of the proposed technique with respect to the most important criterion for the practical applicability of the forecasting method - the accuracy of the resulting forecasts. The experiment showed that the proposed method the forecast error reduced the forecast error in comparison with other methods due to the inclusion of the results of the analysis and modelling of the situation on the FCM in the forecasting process in the current observation mode.

The experiment showed that the proposed method significantly reduces the forecast error compared to other methods due to the inclusion in the process of forecast the results of analysis and simulation of the situation on FCM.

We see further development of the proposed forecasting methodology in the following directions:

– to expand the composition of signal types about a possible change in the development of the situation according to situation and digital monitoring data and their formalization – for the interaction of blocks of qualitative and quantitative forecasting;

– to develop a formalized model of the situation monitoring system, combining the FCM of situation, considered as the basis for identifying and filtering of information-search parameters, and modern approaches for setting up search and intellectual analysis of text content.

– to develop a formalized model of the situation monitoring system, combining the FCM of the situation, considered as the basis for the identification and filtering of information-search parameters, and modern methods of organization of search and intellectual analysis of texts.

The practical significance of the developed forecasting methodology is to increase the effectiveness of expert-analytical and predictive activities in situations of uncertainty and instability.

REFERENCES

1. Andreou, A., Neophytou, H. & Stylianou, C. (2014). A review of the application of fuzzy cognitive maps in the policy decision-making life cycle. In *Handbook of research on advanced ICT integration for governance and policy modelling*, https://doi.org/10.4018/978-1-4666-6236-0.ch008.
2. Armstrong, J.S. (2005). The forecasting canon: nine generalizations to improve forecast accuracy, *International Journal of Applied Forecasting*, **1**, 29–35.
3. Avdeeva, Z., Kovriga, S. & Makarenko, D. (2016). On the statement of a system development control problem with use of SWOT-analysis on the cognitive model of a situation, *IFAC-Papers Online*, **49**(12), 1838–1843, https://doi.org/10.1016/j.ifacol.2016.07.850.
4. Avdeeva, Z.K., Grebenyuk E.A. & Kovriga, S.V. (2020). The Technology of the Strategic Goal-Setting and Monitoring of a Manufacturing System Development on the

Basis of Cognitive Mapping. In M.A. Mellal (Ed.), *Manufacturing Systems: Recent Progress and Future Directions* (chapter 3), New York: Nova Science Publishers.

5. Bayükşahina, Ü. & Ertekin, S. (2019). Improving forecasting accuracy of time series data using a new ARIMA-ANN hybrid method and empirical mode decomposition, *Neurocomputing*, **361**, 151–163, https://doi.org/10.1016/j.neucom.2019.05.099.
6. Büyüközkan, G., Feyzioglu, O. & Vardaloglu, Z. (2009). Analyzing CPFR Supporting Factors with Fuzzy Cognitive Map Approach. *Proc. of 2009 International Conference on Computers & Industrial Engineering*, https://doi.org/10.5281/zenodo.1085579.
7. Byrne, J.P., Fazio, G. & Fiessa, N. (2013). Primary commodity prices: co-movements, common factors and fundamentals, *Journal of Development Economics*, **101**, 16–26, https://doi.org/10.1016/j.jdeveco.2012.09.002.
8. Cai, X. J., Fang, Z., Chang, Y., Tian, S. & Hamori, S. (2020). Co-movements in commodity markets and implications in diversification benefits, *Empirical Economics*, 58, 393-425, https://doi.org/10.1007/s00181-018-1551-3.
9. Cheikhrouhou, N., Marmier, F., Ayadi, O. & Wieser, P. (2011). A collaborative demand forecasting process with event-based fuzzy judgements, *Computers & Industrial Engineering*, **61**(2), 409–421, https://doi.org/10.1016/j.cie.2011.07.002.
10. Cheng, C., Sa-Ngasoongsong, A., Beyca, O., Le, T., Yang, H., et. al. (2015). Time series forecasting for nonlinear and nonstationary processes: a review and comparative study, *IIE Transactions*, https://doi.org/10.1080/0740817X.2014.999180.
11. Dées, S., Karadeloglou, P., Kaufmann, R. K. & Sánchez, M. (2007). Modelling the world oil market: Assessment of a quarterly econometric model, *Energy Policy*, **35**(1), 178–191, https://doi.org/10.1016/j.enpol.2005.10.017.
12. Dickerson, J.A., and Kosko, B. (1994). Virtual worlds as fuzzy cognitive maps, *Teleoperators & Virtual Environments*, **3**(2), 173–189, https://doi.org/10.1162/pres.1994.3.2.173.
13. Dooley, G. & Lenihan, H. (2005). An Assessment of time series methods in metal price forecasting, *Resources Policy*, **30**(3), 208–217, https://doi.org/10.1016/j.resourpol.2005.08.007.
14. Fama, E.F. (1970). Efficient capital markets: A review of theory and empirical work, *The journal of Finance*, **25**(2), 383–417.
15. Fildes, R., Goodwin, P. & Lawrence, M. (2006). The Design features of forecasting support systems and their effectiveness, *Decision Support Systems*, **42**(1), 351–361, https://doi.org/10.1016/j.dss.2005.01.003.
16. *Forecasting: theory and practice*. (2020). Collective monograph submitted to International Journal of Forecasting, URL: https://arxiv.org/abs/2012.03854
17. Granger, C.W.J. & Newbold, P. (1974). Spurious regressions in econometrics, *Journal of Econometrics*, **2**, 111–120, https://doi.org/10.1016/0304-4076(74)90034-7.
18. Grebenyuk, E.A. (2003). Detection of changes in the properties of time-varying random processes, *Automation and Remote Control*, **64**(12), 1868–1881.
19. Grebenyuk, E. A. (2020). Monitoring and identification of structural shifts in processes with a unit root. *Proc. of 13th International Conference* "*Management of large-scale system development*" *(MLSD'2020)*, Moscow, Russia, https://doi.org/10.1109/MLSD49919.2020.9247829.
20. Hogarth, R.M. & Makridakis, S. (1981). Forecasting and planning: an evaluation, *Management science*, **27**(2), 115–138, https://doi.org/10.1287/mnsc.27.2.115.

21. Kannegiesser, M., Günther, H.-O., P. van Beek, Grunow, M. & Habla, C. (2009). Value chain management for commodities: a case study from the chemical industry, *OR Spectrum*, 31, 63–93, https://doi.org/10.1007/s00291-008-0124-9.
22. Liu, C., Sun., X., Wang, J., Li, J. & Chen, J. (2021). Multiscale information transmission between commodity markets: An EMD-based transfer entropy network, *Research in International Business and Finance*, **55**(1), 101318, https://doi.org/10.1016/j.ribaf.2020.101318.
23. Lopez, C. & Salmeron, J. (2014). Dynamic risks modelling in ERP maintenance projects with FCM, *Information Sciences*, **256**, 25–45, https://doi.org/10.1016/j.ins.2012.05.026.
24. Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Berlin: Springer-Verlag.
25. Montgomery, J.M., Hollenbach, F.M. & Ward, M.D. (2012). Improving predictions using ensemble Bayesian model averaging, *Political Analysis*, **20**(3), 271–291.
26. Papageorgiou, E. (2013). Review study on fuzzy cognitive maps and their applications during the last decade. In M. Glykas (Ed.), *Business process management* (pp. 281–298), https://doi.org/10.1007/978-3-642-28409-0_11.
27. Patra, K. & Mondal, S. K. (2017). Risk analysis in a production system using fuzzy cognitive map, *International Journal of Mathematics in Operational Research*, 11(1), 29-44, https://doi.org/10.1504/IJMOR.2017.10006266.
28. Perera, H.N., Hurley, J., Fahimnia, B. & Reisi, M. (2019). The human factor in supply chain forecasting: A systematic review, *European Journal of Operational Research*, **274**(2), 574–600, https://doi.org/10.1016/j.ejor.2018.10.028.
29. Sahoo, D., Sood, N., Rani, U., Abraham, G., Dutt, V., et. al. (2020). Comparative Analysis of Multi-Step Time-Series Forecasting for Network Load Dataset. *Proc. of 2020 11th International Conference on Computing, Communication and Networking Technologies*. https://doi.org/10.1109/ICCCNT49239.2020.9225449.
30. Sona, P., Johnson, T. & Vijayalakshmi, C. (2018). Analysing factors in production management using fuzzy cognitive mapping, *International Journal of Pure and Applied Mathematics*, **118**(23), 517–524, URL: https://www.acadpubl.eu/hub/2018-118-23/articles/23b/24.pdf.
31. Taieb, S., Bontempi, G., Atiya, A. & Sorjamaab, A. (2012). A review and comparison of strategies for multi-step ahead time series forecasting based on the NN5 forecasting competition, *Expert System with Applications*, **39**(8), 7067–7083. https://doi.org/10.1016/j.eswa.2012.01.039.
32. Tang, L., Zhang, C., Li, L. & Wang, S. (2020). A multi-scale method for forecasting oil price with multi-factor search engine data, *Applied Energy*, **257**, 114033, https://doi.org/10.1016/j.apenergy.2019.114033.
33. Taylor, S. J. (1988). Forecasting market prices, *International Journal of Forecasting*, **4**(3), 421–426, https://doi.org/10.1016/0169-2070(88)90109-4.
34. Zhang, X., Yu, L., Wang, S & Lai, K. K. (2009). Estimating the impact of extreme events on crude oil price: An EMD-based event analysis method, *Energy Economics*, **31**(5), 768–778, https://doi.org/10.1016/j.eneco.2009.04.003.