DSS for implementing systemic approach to forecasting

Peter Bidyuk [1], Alexander Gozhyj [2], Alexander Trofymchuk [3], Alexei Bidyuk [4]

Dr. of eng. sci., professor at the Institute for Applied System Analysis, NTUU "KPI",Kiev,Ukraine[1]

pbidyuke@gmail.com

PhD., associate professor, Department of Information Technology and Program Systems at the Black Sea State University named after Petro Mohyla ,Nikolaev, Ukraine[2]

alex_daos@mail.ru

Dr. of eng. sci., professor, Institute of Telecommunications and Global Information Sphere at NAS of Ukraine,Kiev,Ukraine[3]

itelua@kv.ukrtel.net

Master, Profigent Labs (NL), Complex industrial system control software: research and development, Amsterdam, NL [4]

Oleksii.Bidiuk@gmail.com

Abstract

A computer based decision support system is proposed the basic tasks of which are modeling and forecasting of financial processes and credit risk estimation. The system is developed on the basis of system analysis principles, i.e. it has hierarchic architecture, there exists the possibility for taking into consideration some stochastic and information uncertainties, forming alternatives for models and forecasts, and tracking the computing procedures during all stages of data processing. A modular architecture is implemented that provides a possibility for the further easy enhancement and modification of the system functional possibilities with new forecasting and parameter estimation techniques, and the risk models. High quality of final result is achieved thanks to appropriate tracking of the computing procedures at all stages of data processing: preliminary data processing, model constructing, and forecasts estimation with appropriate sets of statistical quality criteria. Example is given for modeling selected processes and credit risk estimation. The example presented shows that the system developed has good perspectives for practical use.

Indexing terms/Keywords

Introduction

Very often it is possible to reach acceptable quality of forecasting dynamic processes using available on market data processing instruments. For example, well known SAS contains sophisticated data processing procedures usually capable of producing high quality final results [1]. However, such instruments are usually very costly, require special training for solving specific problems and need highly developed computers for implementation. All this creates conditions of rather restricted access for their usage except for banking system and large enterprises. Besides, new modeling and forecasting techniques that appear from time to time in specialized publications need to be appropriately implemented and approbated. To solve this problem it is would be reasonable to develop simpler and much less costly modeling systems constructed with taking into consideration modern principles of system analysis. The systemic approach is based on the results of application of systems analysis methodology.

Processing of statistical data, represented by the time series, is usually accompanied by uncertainties of various kind and nature. More particularly, these are (at least) uncertainties of structural, statistical and parametric form. The structural uncertainties are usually encountered in cases when analysis of time series data doesn't exhibit clear structure for respective model describing it. Remind that the notion of model structure includes the following elements: dimensionality (number of equations comprising a model); model order (highest order of a model equation); input delay time (lag) for independent variables (regressors); nonlinearity and its type (nonlinearity in variables or in parameters); stochastic disturbance and its type (distribution and its parameters) [2]. For example, the model order or input delay time (lags for independent variables) cannot be determined exactly enough using correlation analysis or appropriate time lag estimation algorithms. It is also not always possible to determine in a unique way type of nonlinearity, especially in the cases of processing short samples. The statistical uncertainties exist due to impossibility to determine in a unique way type of probability distribution for stochastic disturbances and the collected data itself, and as a consequence parameters of the distributions, degree of nonstationarity for the stochastic processes etc. As a consequence for existence of the two previous types of uncertainties we can trace the parametric uncertainties: mathematical model parameter estimates, computed with statistical data, can be not consistent, contain bias, and be inefficient.

Thus, existence of process and data uncertainties, and necessity for hierarchical organization for the data processing system as well as necessity for the functional completeness of the whole processing require development and application of systemic approach [3, 4], that provides a possibility for solving many substantial problems encountered in statistical data analysis, model constructing, forecasting and generating decision alternatives. In this study we consider some possibilities for constructing data processing scheme based on the principles of systemic approach.

Problem Statement

The purpose of the study is as follows: – to consider some principles of the systems analysis that are appropriate for solving the problem of successful short-term forecasting; – to develop an efficient data processing scheme for implementation in decision support systems based on these principles; – to consider some uncertainties inherent in model building and forecasting process and to show the ways for their elimination; – to show advantages of the approach developed.

Some System Analysis Principles Used

For development and implementation of modeling and forecasting system it is proposed to use the following systems analysis principles [4]: the systemic coordination principle; the principle of procedural completeness; the functional orthogonality principle; the principle of mutual informational dependence; the principle of goal directed correspondence; the principle of functional rationality; the principle of multipurpose generalization; the principle of multifactor adaptation, and the principle of rational supplement. The tasks to be solved arising from the systems analysis approach application to forecasting illustrates Fig. 1.
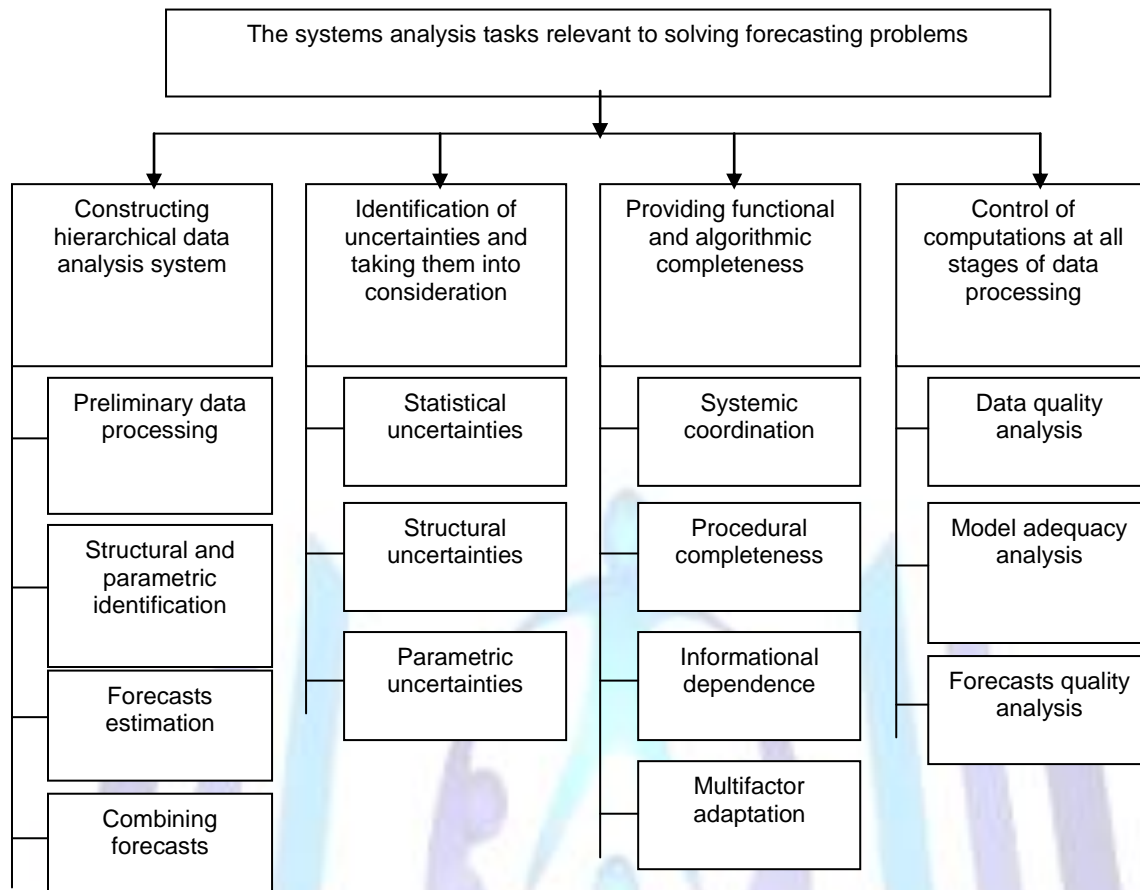
```
┌──────────────────────────────────────────────────────────────┐
│   The systems analysis tasks relevant to solving forecasting   │
│                         problems                               │
└──────────────────────────────────────────────────────────────┘
```

Constructing hierarchical data analysis system	Identification of uncertainties and taking them into consideration	Providing functional and algorithmic completeness	Control of computations at all stages of data processing
Preliminary data processing	Statistical uncertainties	Systemic coordination	Data quality analysis
Structural and parametric identification	Structural uncertainties	Procedural completeness	Model adequacy analysis
Forecasts estimation	Parametric uncertainties	Informational dependence	Forecasts quality analysis
Combining forecasts		Multifactor adaptation	

Fig. 1 Tasks arising from application of the systems analysis approach

to forecasting

According to the principle of systemic coordination all the techniques, approaches, and algorithms used in the system should be structurally and functionally coordinated, and should be mutually dependent (linked). This way it is possible to reach a unique systematic methodology for data analysis in the frames of modern decision support system (DSS) constructed, and to improve the quality of intermediate and final results.

The principle of procedural completeness guaranties that DSS developed will provide the possibility for timely and in place execution of all necessary functions directed towards data collection (editing and renewing), formalization of a problem statement, model constructing, computing forecast, and for estimating quality of the model and the forecasts based on it.

Implementation of all computational procedures in the DSS using mutually independent functions corresponds to the principle of functional orthogonality. Such approach to the DSS constructing favors to substantial enhancement of computational stability of the system and simplification of its possible further modifications and functional expansion. According to the principle of mutual informational dependence the computing results, generated by each procedure, should correspond to the formats and requirements of other procedures (or modules). This is provided with respective project development solutions for the system created.

Application of the principle of goal directed correspondence to computational procedures and functions provides a possibility for reaching of a unique goal – high quality of the final result in the form of forecast estimates for the process under study, as well as decisions based on the forecasts.

The principle of functional rationality prevents duplication of separate DSS functions. As a result of application of the principle it is possible to reach economy of respective computational resources.

According to the principle of multipurpose generalization all the functional modules developed possess necessary degree of generalization what provides a possibility for reaching high quality solutions for a set of distinctive problems that belong to the same class (say high quality forecasting for linear and nonlinear nonstationary processes). Among these problems are the following: accumulating and preliminary data processing; estimation of structure and parameters for the necessary mathematical models; constructing forecasting functions and computing of the forecasts; selecting the best results using appropriate sets of statistical quality criteria.

The principle of multifactor adaptation is directed towards the possibility of solving the problems of computing procedures adaption to the problems of modeling processes of different complexity depending on the completeness of available

---

information and user requirements. The adaptation is performed in the process of model structure and parameters estimation, i.e. the whole identification process of an object under study is represented by a set of adaptive procedures directed towards reaching the main goal: constructing adequate model and computing high quality forecasts.

The use of the rational supplement principle provides a possibility for enhancing the sphere of application of the DSS constructed by expanding it with new data types, computing procedures and criteria. The new procedures could be directed towards implementation of additional preliminary data processing procedures, model structure and parameter estimation, generating forecasting functions as well as selection of the best result for its further use aiming generation of appropriate decision.

A hierarchical structure of the DSS proposed for data analysis in the form of time series is given in Fig. 2. Implementation in the DSS of the system analysis principles mentioned above favors its functional flexibility, computational reliability, enhancement of quality for the final results, expansion of system life span in general, and simplification of drawback elimination and modification procedures.

Consider some functional elements of the general hierarchical structure of the DSS proposed that uses some principles of systemic approach (Fig. 1). After preliminary data processing directed towards data preparation for further model building, statistical analysis of data is performed aiming to determining type of the process under consideration (linear/nonlinear, and stationary/nonstationary) and model structure estimation. An important and useful feature of the system is that it uses separate sets of statistics for analyzing quality of data, adequacy of model constructed and quality of forecast estimates generated with the model. If the forecasts are used for generating alternative decisions then we include another set of criteria for testing decision quality. This way high quality forecasts are generated and the decisions based on these forecasts are usually acceptable.

Further statistical analysis of data is performed on the following purposes: testing for heteroskedasticity, for integration (availability of trend), nonlinearity, preliminary model structure estimation using correlation techniques, lag estimation and determining type of data distribution. After this step it is possible to determine class of the process under study and estimate the elements of model structure mentioned above in the introduction. The forecasting methods used in the system are as follows: regression analysis (linear and nonlinear models), the group method for data handling (GMDH), fuzzy GMDH, fuzzy logic, appropriate versions of Kalman filter, neural nets, support vector regression, nearest neighbor and probabilistic type techniques (Bayesian networks, Bayesian regression).

The set of techniques used covers linear and nonlinear nonstationary process. The nearest neighbor technique can be hired for generating long term forecasts in a case of availability long data samples. Practically all the techniques are implemented in adaptive versions what makes the system more flexible for newly coming data and to fight some types of uncertainties. On the step of model structure estimation an appropriate principal component analysis technique is applied when necessary.
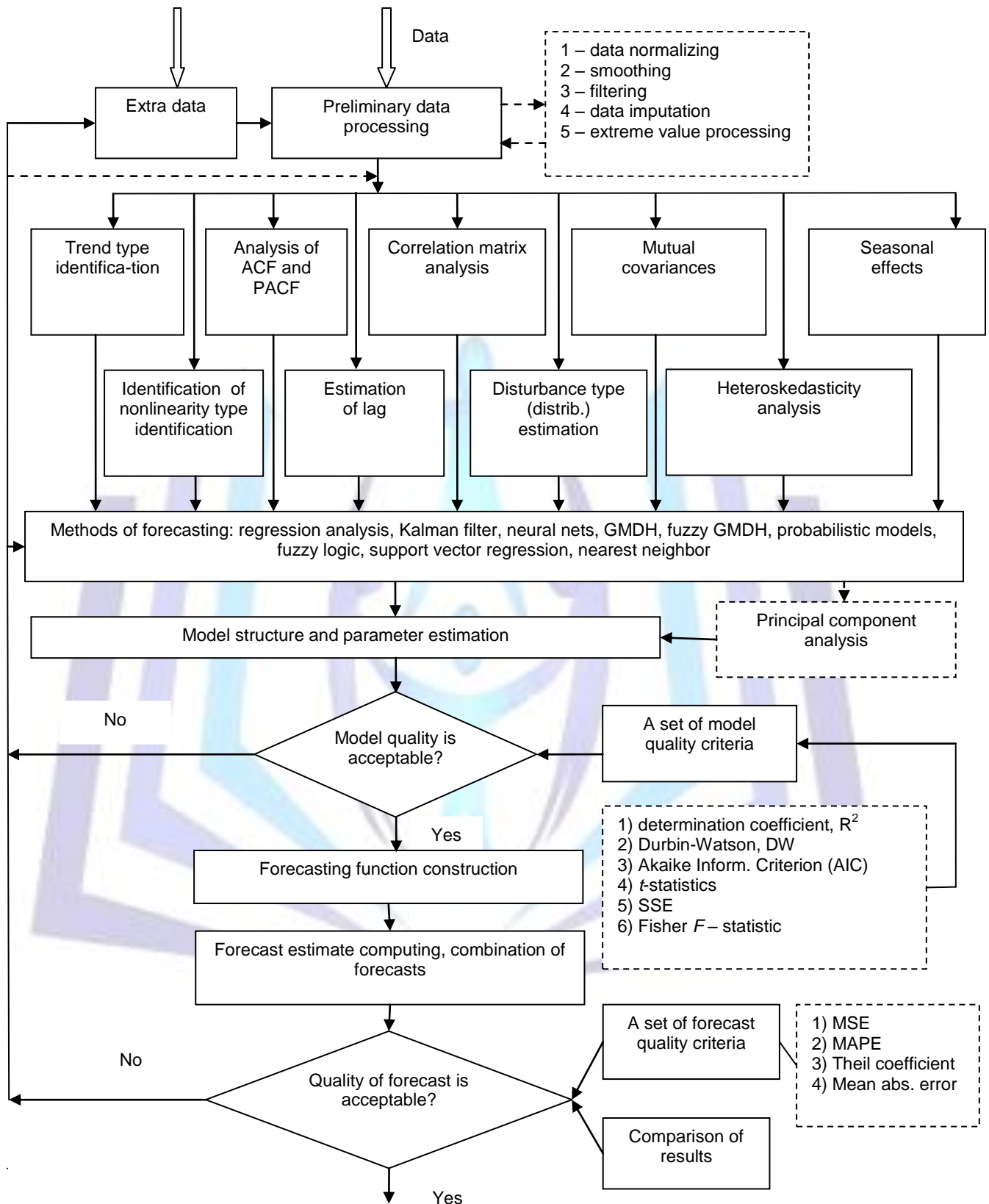
Fig. 2 Hierarchical structure of DSS for generating forecasts

The information of a model structure provides a possibility for selection of parameters estimation techniques for candidate models among which are: ordinary and nonlinear LS (NLS), recursive LS (RLS), maximum likelihood (ML), and some versions of Markov Chain Monte Carlo (MCMC). Some special optimization techniques are applied in a case of estimating fuzzy GMDH structures.

The candidates models estimated are tested with a set of statistical quality criteria some of which are shown in Fig. 1. One or more acceptable models are used for computing forecasts that are also tested for quality with another set of statistical criteria. Usually two or more forecasting techniques are used for the same dataset to get a possibility for combining the forecasts so that to further improve the final combined estimate.

Uncertainties Identification And Processing

Processing statistical uncertainty. The most often met statistical uncertainties related to model development and estimation of forecasts are provoked by the following factors:

- measurement errors (noise) that is available practically in all cases of data collection independently on the data origin (including economy and finances);

- stochastic external disturbances that usually negatively influence the process under study and shift the processes from desired mode (say, offshore capital transfer from some country, low quality of higher administration, unstable often changed laws);

- missed measurements (observations) and outliers;

- multicollinearity that requires special data processing techniques to reduce degree of mutual correlation between separate time series.

The most often means used to fight measurement noise and external stochastic disturbances are digital and optimal filters (among them Kalman filter is a widely used instrument) [5]. Digital filters (DF) help to select for subsequent processing the frequency band of interest by processing the time series data with linear structures that could be represented by autoregression (AR) or autoregression with moving average (ARMA) equations of the type:

$$y(k) = \sum_{i=1}^{p} a_i\, y(k-i) + \sum_{j=1}^{q} b_j\, x(k-j) \, ,$$

where $x(k)$ is input sequence of measurements; $y(k)$ is filtered output sequence; $\theta = [a_1 ... a_p\, b_1 ... b_q]^T$ is parameter vector that determines frequency band of a filter; p and q are orders of AR and MA parts, respectively. There are handy instruments for DF development based on optimization procedures, for example, in Matlab.

Appropriately designed adaptive Kalman filter provides a possibility for covariances estimation for stochastic disturbances and measurement noise as well as estimation of short-term forecasts [5, 6]. Optimal filter design requires model of the process (system) under study in the state space form:

$$\mathbf{x}(k) = \mathbf{A}(k, k-1)\mathbf{x}(k-1) + \mathbf{B}(k, k-d)\mathbf{u}(k-1) + \mathbf{w}(k) , \qquad (1)$$

$$\mathbf{z}(k) = \mathbf{H}\mathbf{x}(k) + \mathbf{v}(k) , \qquad (2)$$

where $\mathbf{x}(k)$ is system state vector, consisting of n elements; $\mathbf{u}(k)$ is m-dimensional vector of controls; $\mathbf{w}(k)$ is a vector of stochastic external disturbances; $\mathbf{A}(k, k-1)$ is $(n \times n)$-matrix of system dynamics; $\mathbf{B}(k)$ is $(n \times m)$-matrix of control coefficients; k is discrete time that is linked to continuous time t through sampling time $T_s : t = kT_s$; d is system delay time; $\mathbf{z}(k)$ is measurement vector with r elements; $\mathbf{H}(k)$ is matrix of measurement coefficients (very often it has diagonal form); $\mathbf{v}(k)$ is a vector of measurement noise (errors). Double time argument $(k, k-1)$ means that variable with this argument is used at time k, but its value is based on previous measurements up to $(k-1)$. Usually such double argument is not used in text to avoid symbols overload for mathematical expressions.

The main advantage of the model (1), (2) is that it includes explicitly two random process $\mathbf{w}(k)$ and $\mathbf{v}(k)$; consequently it is more adequate to reality than say linear regression. The main task of optimal filter is in computing state vector estimates with taking into consideration statistical characteristics (covariances) of the two stochastic processes mentioned. Such approach provides a possibility for improvement of state estimates and to compute estimates of nonmeasurable components of state vector $\mathbf{x}(k)$ if such are available. The main equation of the filter is as follows:

$$\hat{\mathbf{x}}(k) = \mathbf{A}(k)\hat{\mathbf{x}}(k-1) + \mathbf{B}(k)\mathbf{u}(k-1) + \mathbf{K}(k)[\mathbf{z}(k) - \mathbf{H}(k)\mathbf{A}(k)\hat{\mathbf{x}}(k-1)] , \qquad (3)$$

where $\mathbf{K}(k)$ is optimal coefficient of a filter in matrix form. The coefficient is computed by minimizing the functional:

$$J = \min_{\mathbf{K}} E\left\{[\hat{\mathbf{x}}(k) - \mathbf{x}(k)]^T [\hat{\mathbf{x}}(k) - \mathbf{x}(k)]\right\}.$$

where $\mathbf{x}(k)$ is exact value for state vector that could be found using deterministic part of model (1). In linear discrete case the coefficient is rather easily computed by solving Riccati equation.

Thus, optimal filter provides a possibility for reducing uncertainties in the form of influence of two stochastic processes $\mathbf{w}(k)$ and $\mathbf{v}(k)$, and estimation of nonmeasurable components of a state vector when respective components of covariance matrices are known (we mean covariance matrix $\mathbf{P}(k) = E\left\{[\hat{\mathbf{x}}(k) - \mathbf{x}(k)][\hat{\mathbf{x}}(k) - \mathbf{x}(k)]^T\right\}$ for system state vector estimation errors). Especially useful are adaptive versions of the filters that are most suitable for practical applications (in on-line and off-line modes of operation). They are suitable for repeated estimation of system (object) matrices $\mathbf{A}(k)$ and $\mathbf{B}(k)$ as well as covariances of the two stochastic processes mentioned [5, 7].

Coping with uncertainties appearing due to missing observations. For the data in the time series form the most suitable imputation techniques are: – simple averaging when it is possible (when only a few values are missing); – generation of forecast estimates with the model constructed using available measurements; – generation of missing (lost) estimates from distributions the form and parameters of which are again determined using available part of data; – the use of optimization techniques, say appropriate forms of EM-algorithms (expectation maximization); – exponential smoothing etc. The simplest model that could be hired for generating forecasts is AR(1):

$$y(k) = a_0 + a_1 y(k-1) + \varepsilon(k),$$

where a_0, a_1 are model parameters; random process $\varepsilon(k)$ takes into consideration model structure uncertainties (say lack of appropriate regressors, external random disturbances, errors of parameter computing etc.), and possible measurement errors. If parameters a_0, a_1 are known, we could compute one step-ahead prediction as conditional expectation of the form:

$$\hat{y}(k+1) = E_k[y(k+1)] = E_k[y(k+1) \mid y(k), y(k-1),..., \varepsilon(k), \varepsilon(k-1),...] =$$

$$= a_0 + a_1 E_k[y(k)] = a_0 + a_1 y(k),$$

as far as $y(k)$ at the moment k takes known value. Iteratively it is possible to derive the forecasting function for s steps-ahead [2]:

$$\hat{y}(k+s) = E_S[y(k+s)] = a_0\left(\sum_{i=0}^{S-1} a_1^i\right) + a_1^S y(k) = a_0 \sum_{i=0}^{S-1} a_1^i + a_1^S y(k).$$

The sequence of forecast estimates $\left\{\hat{y}(k+i)\right\}$, $i = 1,...,s$ is convergent if $\left|a_1\right| < 1$:

$$\lim_{s \to \infty} E_k[y(k+s)] = \frac{a_0}{1 - a_1}.$$

The last expression means that for stationary AR or ARMA processes the estimates of conditional forecasts asymptotically $(s \to \infty)$ converge to unconditional mean (long-term forecast). It should also be mentioned here that optimal filter can also be used for missing data imputation because it contains "internal" forecasting function that provides a possibility for generating quality short-term forecasts.

Further reduction of the uncertainty is possible thanks to application of several forecasting techniques to the same problem with subsequent combining of separate forecasts using appropriate weighting coefficients. The best results of combining the forecasts is achieved when variances of forecasting errors for different forecasting techniques do not differ substantially.

Coping with uncertainties of model parameters estimates. Usually uncertainties of model parameter estimates such as bias and inconsistency result from low informative data, or data do not correspond to normal distribution, what is required in a case of application LS for parameter estimation. This situation may also take place in a case of regressors multicollinearity and substantial influence of process nonlinearity that for some reason has not been taken into account when model was constructed. When power of data sample is not satisfactory for model construction it could be expanded by applying special techniques, or simulation is hired, or special model building techniques, such as GMDH, are applied. Very often GMDH produces results of acceptable quality with short samples. If data do not correspond to normal distribution, then ML technique could be used or appropriate Monte Carlo procedures for Markov Chains (MCMC) [8]. The last techniques could be applied with quite modest computational expenses when the number of parameters is not large.

Dealing with model structure uncertainties. When using DSS, model structure should practically always be estimated using data. It means that elements of model structure accept almost always only approximate values. When a model is constructed for forecasting we build several candidates and select the best of them with a set model quality statistics. Generally we could define the following techniques to fight structural uncertainties: gradual improvement of model order (AR(p) or ARMA(p, q)) applying adaptive approach to modeling and automatic search for the "best" structure using complex statistical quality criteria; adaptive estimation (improvement) of input delay time (lag) and data distribution type with its parameters; describing detected process nonlinearities with alternative analytical forms with subsequent estimation of model adequacy and forecast quality. An example of complex model and forecast criterion may look as follows:

$$J = \left|1 - R^2\right| + \alpha \ln\left[\sum_{k=1}^{N} e^2(k)\right] + \left|2 - DW\right| + \beta \ln(MAPE) + U \to \min_{\hat{\theta}_i},$$

where R^2 is determination coefficient; $\sum_{k=1}^{N} e^2(k) = \sum_{k=1}^{N} [y(k) - \hat{y}(k)]^2$ is sum of squared model errors; DW is

Durbin-Watson statistic; $MAPE$ is mean absolute percentage error for one step-ahead forecasts; U is Theil coefficient

that measures forecasting characteristic of a model; α, β are appropriately selected weighting coefficients; $\hat{\theta}_i$ is

parameter vector for $i-th$ candidate model. A criterion of this type is used for automatic selection of the best candidate model. The criterion also allows operation of DSS in adaptive mode. Certainly, other forms of the complex criteria are possible. While constructing the criterion it is important not to overweigh separate members in right hand side.

Coping with uncertainties of a level (amplitude) type. The use of random (i.e. with random amplitude or a level) and/or non-measurable variables leads to necessity of hiring fuzzy sets for describing such situations. The variable with random amplitude can be described with some probability distribution if the measurements are available or they come for analysis in acceptable time span. However, some variables cannot be measured (registered) in principle, say amount of shadow capital that "disappears" every month in offshore, or amount of shadow salaries paid at some company, or a technology parameter that cannot be measures on-line due to absence of appropriate gauge. In such situations we could assign to the variable a set of possible values in linguistic form as follows: *capital amount* = { *very low*, *low*, *medium*, *high*, *very high* }. There is a necessary set of mathematical operations to be applied to such fuzzy variables. Finally fuzzy value could be transformed into exact form using known techniques.

Probabilistic uncertainties and their description. The use of random variables leads to necessity of constructing probability distributions and their application in inference procedures. Usually an observation value is known only approximately though we know the limits for the values. Probability distributions are very useful for describing such situations. When dealing with discrete outcomes, we assign probabilities to specific outcomes using a mass function. It shows how much "weight" (or mass) to assign to each observation or measurement. An answer to the question about the value of a particular outcome will be its mass. The Kolmogorov's axioms of probability are helpful for deeper understanding of what is going on. If two or more variables are analyzed simultaneously we come to necessity of constructing and use joint distributions. Joint distributions allow us calculate conditional probabilities using renormalization procedures when necessary. Very helpful for performing probabilistic computations is a notion of conditional independence: $P(x, y \mid z) = P(x \mid z)\, P(y \mid z)$, where x and y are independent events. The identities of independence are very handy though one should be careful when using them, i.e. the events should be really independent. The remarkable intuitive meaning of discrete Bayes' law, $P(A/B) = P(B/A)P(A)/P(B)$, is that it allows to ask reverse questions: "Given that event A happened, what is the probability that a particular event B evoked it?" [9]. The marginal probability, $P(B)$, can be computed from conditionals. The probability that B occurs in general ($P(B)$) can be obtained from the condition:

$$P(B) = P(B/A)\,P(A) + P(B/\overline{A})\,P(\overline{A}).$$

Probabilistic types of uncertainties regarding whether or not some event will happen can be taken into consideration with various probabilistic models. To solve the problem of describing and taking into account such uncertainties a variety of Bayesian models could be hired that are presented in the form of so called Bayesian Programming formalism. A set of the models includes Bayesian networks (BN) [10], dynamic Bayesian networks (DBN), Bayesian filters, particle filters, hidden Markov models (discrete and continuous), Kalman filters, Bayesian maps etc. The structure of Bayesian program includes the following elements (steps): (1) problem description and statement formulation with a basic question of the form: $P(Searched / Known)$ or $P(X_i / D, Kn)$, where X_i defines one variable only, i.e. what should be estimated (computed) using specific inference engine; (2) the use of prior knowledge Kn and experimental data D to perform model structure and parameters identification; (3) selection and application of pertinent inference technique to answer the question stated before; (4) testing quality of the final result. This scheme also works well in adaptation mode to adjust structure and parameters of a model being developed to new experimental data or expert estimates that could be used, for example, for estimation of prior distributions or BN (DBN) structure.

Example Of Dss Application

Numerous examples of model constructing have been solved with the DSS created. In this specific example we used the database consisting of 4700 records that was divided into learning sample (4300 records), and test sample (400 records). The default probabilities were computed and compared to actual data, and also errors of the first and second type were computed using different values of cut-off value. It was established that maximum model accuracy reached for Bayesian network was 0.787 with the cut-off value 0.3. The Bayesian network is "inclined to over insurance", i.e. it rejects more often the clients who could return the credit. The model accuracy and the errors of type I and type II depend on the cut-off value. The cut-off value determines the lowest probability limit for client's solvency, i.e. below this limit a client is considered as such that will not return the credit. Or the cut-off value determines the lowest probability limit for client's default, i.e. below this limit a client is considered as such that will return the credit. As far as the cut-off value of 0.1 or 0.2 is considered as not important, in practice it is reasonable to set the cut-off value at the level of about 0.25 – 0.30. Statistical characteristics characterizing quality of the models constructed are given in table 1.

Table 1: Quality of the models constructed

Model type	Gini index	AUC	Common accuracy	Model quality
Bayesian network	**0.719**	0.864	0.787	Very high
Logistic regression	0.685	0.858	**0.813**	Very high
Decision tree	0.597	0.798	0.775	Acceptable
Linear regression	0.396	0.657	0.631	Unacceptable

It follows from the table 1 that the best models for estimation of credit return probability turned out to be logistic regression and Bayesian network. The best common accuracy showed logistic regression (0.813) though Bayesian network showed higher Gini index (0.719). The decision tree used is characterized by Gini index of about 0.597, and $CA = 0.775$. It should be stressed that acceptable values of Gini index for developing countries like Ukraine are in the range 0.4 – 0.6. Bayesian network constructed and nonlinear regression showed rather high values of Gini index that are acceptable for the Ukrainian economy in transition.

The results of computing experiments lead to the conclusion that today scoring models and Bayesian networks belong to the set of the best instruments for banking system due to the fact that BN provide a possibility for detecting "bad" clients and to reduce financial risks caused by the clients. It also should be stressed that DSS constructed is very useful instrument for a decision maker that helps to perform quality processing of statistical data using different techniques, generate alternatives and to select the best one relying on a set of appropriate criteria. The system performs tracking of the whole computational process using separate sets of statistical quality criteria at each stage (level of hierarchy) of decision making: quality of data, models and forecasts (or risk estimates).

Thus, the systematic approach to forecasting proposed helps to develop DSS that possesses the features of directed search for the best forecasting model in respective spaces of structures and parameters, and consequently to enhance its quality. Preliminary computational experiments with actual data showed high usefulness for practical applications of the systemic approach to modeling and forecasting and necessity for it further refinement in the future studies. It is especially important to improve descriptions for the uncertainties mentioned and to use them for reduction degree of uncertainty in the process of model building and forecast estimation.

Conclusions

The methodology was proposed for constructing DSS for mathematical modeling and forecasting of economic and financial processes, and credit risk prediction (estimation) that is based on the following system analysis principles: hierarchical system structure, taking into consideration of probabilistic and statistical uncertainties, availability of adaptation features, generating multiple decision alternatives, and tracking of computational processes at all the stages of data processing with appropriate sets of statistical quality criteria.

The system proposed has a modular architecture that provides a possibility for easy extension of its functional possibilities with new parameter estimation techniques, forecasting methods, financial risk estimation, and generation of decision alternatives. High quality of the final result is achieved thanks to appropriate tracking of the computational processes at all data processing stages: preliminary data processing, model structure and parameter estimation, computing of short- and middle-term forecasts, and estimation of risk variables (parameters) as well as thanks to convenient for a user intermediate and final representation of results. The system is based on the ideologically different techniques of modeling and risk forecasting what creates a good base for combination of various approaches to achieve the best results. The example of the system application shows that it can be used successfully for solving practical problems of forecasting and risk estimation. The results of computing experiments lead to the conclusion that today scoring models, nonlinear regression and Bayesian networks are the best instruments for banking system due to the fact that they provide a possibility for detecting "bad" clients and to reduce financial risks caused by the clients. It also should be stressed that

DSS constructed turned out to be very useful instrument for a decision maker that helps to perform quality processing of statistical data using different techniques, generate alternatives and to select the best one. The system performs tracking of the whole computational process using separate sets of statistical quality criteria at each stage of decision making. The DSS can be used for supporting decision making process in various areas of human activities including development of strategy for banking system and industrial enterprises, investment companies etc.

Further extension of the system functions is planned with new forecasting techniques based on probabilistic techniques, fuzzy sets and other artificial intelligence methods. An appropriate attention should also be paid to constructing user friendly adaptive interface based on the human factors principles.

References

[1]   Fernandez G. Data mining using SAS applications. – New York: CRC Press LLC, 2003. – 360 p.

[2]   Bidiyk P.I.,Romanenko V.D., Tymoschuk O.L. Analiz chasovich riadiv. – Kiev : Politechnika, NTUU «KPI», 2013. – 600 s.

[3]   Zgurowskii M.Z., Pankratova N.D. Systemnyi analiz: problemi, metodologiya, prilogeniya. – Kiev: Naukova dumka, 2005. –   745 s.

[4]   Anfilatov V.S., Emelyanov A.A., Kukushkin A.A. Systemnyi analiz v upravlenii. – M.: Finansi I statistika, 2002. –368 c.

[5]   Zgurowskii M.Z., Podladchikov V.N. Analiticheskie metodi kalmanovskoi filtracii. –Kiev: Naukova dumka,1995.– 285 s.

[6]   Gibbs B.P.Advanced Kalman filtering, least squares and modeling. – Hoboken: John Wiley & Sons,Inc.,2011. – 627 p.

[7]   Haykin S. Adaptive filter theory. – Upper Saddle River (New Jersey): Prentice Hall, 2002. – 922 p.

[8]   Gilks W.R., Richardson S., Spiegelhalter D.J. Markov Chain Monte Carlo in practice. – New York: CRC Press LLC, 2000. – 486 p.

[9]   Jao C.S. Efficient decision support systems – practice and challenges from current to future. – Rijeka (Croatia): Intech, 2011. – 556 p.

[10] Zgurowskii M.Z., Bidyuk P., Terentyev A.M. Methods of constructing Bayesian networks based on scoring functions // Cybernetics and System Analysis, 2008, Vol. 44, No. 2, pp. 219 – 224.