doi: 10.36689/uhk/hed/2020-01-023

Fuzzy Sets as an Initial Analysis for the Prediction of the Bankruptcy Situation

Vítězslav HÁLEK and Majid ZIAEI NAFCHI *

University of Hradec Králové, Hradec Králové, Czech Republic; vitezslav.halek@uhk.cz; majid@seznam.cz
* Correspondence: majid@seznam.cz

Abstract: The financial analyses of a company are substantially complicated. Especially when it comes to the situations where there are high levels of uncertainty. In this paper, we address the issues regarding the prediction of a bankruptcy situation with the help of fuzzy sets and compare the results of different analytical methods and approaches. A significant part of fuzzy sets theory is related to optimization of the decision making and it is clearly it's best used in the situations where the use of prediction methods that relay on the past data is not possible due to the significance of error caused by the level of uncertainty. The construction of predictions through linear models is not yet offered in the literature. The main reason is the inconsistency of the methodology, which could find application in practice. Work on the use of fuzzy sets in financial analysis is part of a larger project to design a bankruptcy model or methodology that must be able to take into account possible errors (or its dispersion).

Keywords: fuzzy sets; bankruptcy; financial analysis

JEL Classification: C1; G0

1. Introduction

The word "fuzzy" means wispy, unclear, misty, vague, and uncertain (Zadeh 1965). It is possible to understand the fuzzy set as the complete universe, where only some elements are not definitely in it, and the membership degree to the fuzzy set is specified by mathematical functions (Kainz 2010).

The field of impact of the fuzzy sets in mathematic modeling of vagueness cannot be separated from the terms used in verbal description. The terms used in financial analysis do not and cannot have exact borders. It is possible to mention such terms as rapid onset of a bankruptcy situation, prospective development of financial management, reasonable debt level, a financially sound company etc. These terms are not clearly and exactly delimited and it is not adequate to use the probability either. However the related insecurity is not of stochastic, arbitrary type (Mareš 2002).

Mathematic operations used for the description of fuzziness are naturally different from operations with random events and variables. That is why the fuzzy sets theory has been created. A significant part of this theory is related to optimization of the decision making and evaluation of the fuzzy data files. In case of classical deterministic sets it is possible to clearly determine if each object belongs to a specific set or not (industrial company, debt sources, long-term property). Fuzzy sets can be used with objects (see above), whose membership cannot be definitively determined.

To get a clear description of a fuzzy sub-set "A", it is possible to generalize the characteristic-membership function μA in such a way, that for each element x applies:

$\mu A(x) = 1$ If x certainly belongs to A, $\mu A(x) = 0$ if x certainly does not belong to A,
$0 < \mu A(x) < 1$ If certain, that x belongs to A.

In the last case the value of $\mu A(x)$ will be the closer to 1, the more likely x can be considered an element of a fuzzy set A (Navara and Olšák 2002). From this point of view an expert opinion is a source of fuzziness of quantitative data; reasonable profit rate, relatively high correlation, an appropriate structure of capital resources, frequent decline in the economic sector etc. And thus if financial analysis uses verbal description of the economic reality, it is necessary to know, how to clearly distinguish the relative significance of descriptive characteristics. If fuzzy sets are used, the vagueness of the results increases enormously, however this theory enables to work efficiently with the fuzzy quantitative and qualitative data. In this expression a fuzzy set is an ordered pair $\tilde{A} = (U, \mu\tilde{A})$, where U is the basic set and $\mu\tilde{A}$ is the membership function defined as U $\mu\tilde{A}: U \rightarrow [0, 1]$ (Zadeh 1965).

The value $\mu\tilde{A}$ (x) then expresses the grade of membership of x to the fuzzy set $\tilde{A}$. It is possible to simplify the expression of a fuzzy set $\tilde{A}$ using grades of membership in the discrete case of universe (i.e. the basic set) U , so that the given elements are omitted and the fuzzy set $\tilde{A}$ is written as a row vector where the vector components are equal to grades of membership of individual elements. If a fuzzy set is expressed in this way, it is possible to work with it same as with a usual vector, the only difference being use of special operations (e.g. to set if the selected asset of the entity was acquired recently or not, a so called "new property". The vagueness of the description of the economic reality consists in the attribute "new". The value $\mu\tilde{A}$ (x) sets a date in the calendar. If x (meaning new property) is acquired on January 1st, then $\mu\tilde{A}$ (x) = 1, 31st of December $\mu\tilde{A}$ (x) = 0, other values (dates) are within the interval $0 < \mu A(x) < 1$. This division is based on the idea of a calendar year with 365 days. A more appropriate approach for calculations would be based on the length of different vectors, where e.g. the length of the value 1st of January is 1, 3 of 3rd February is 34 etc. From the point of view of vector operations, it is already evident, that the longer the vector is the older is the asset).

2. Results and Discussion

Within financial analysis such a way of expression of descriptive relations can be considered sufficient, even though for a user of outputs and conclusions, such an intermediate result of the analysis may be rather confusing. Moreover, such an entry is often not even considered a clear result of monitoring of two variables. In case of bankruptcy it is important to know what debt ratio; eventually the amount of debt finance is acceptable with the monitored entity. It is possible to use the recommended values, however even their boundaries are too wide.

In case of debt financing it is thus possible, that one processor of the financial analysis considers the amount of loan capital at the level CK1 adequate, another processor does not. In a similar way, at the level CK2>CK1 the data can be seen in a different way. In this case we often consider the verbal descriptions and financial analysis conclusions as inaccurate. (They are findings and expressions of acceptable amount of debt finance with different analysts and of different analyzed companies). Graphic expression of the above mentioned problem can certainly be as follows in figure 1:

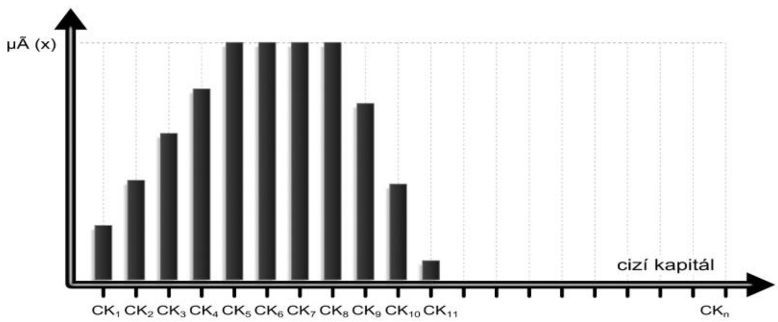

Figure 1: Foreign capital.

In case it is necessary to define a set of all levels of loan capital, that can be marked as acceptable, then the limit values of a selected company would be: CK1 e.g.: 1 million Czech crowns) < acceptable < CK2 (e.g. 10 million Czech crowns). The value of CK3 (e.g.: 500 000 Czech crowns) would not belong to the mentioned set (value log. 0), while the value CK2 (10 million Czech crowns) would (log. 1). How could the analyst classify the value CK4 (e.g. 999 000 Czech crowns)? In practice such level would probably be acceptable. If such level could be acceptable, how should we classify the level of 998,000 CZK and 997,000 CZK?

However, e.g. Boolean logic that is often used in standard bankruptcy models, does not know the above mentioned details. The mentioned problem is solved by the grade of membership. In case of debt resources it is the value of probability, that the given element (999,000 CZK of debts resources) belongs to the given set "acceptable". It can be written as follows: $\mu\tilde{A} = 0.9$ (999), $\mu\tilde{A} = 1$ (10), $\mu\tilde{A} = 0$ (2).

Within the CCB model, universe is estimated according to the economy area, in which the company operates. In that case the debt sources are: U= [0, 20] million Czech crowns. For all companies in general probably: U= [-20,100] million Czech crowns. [Here the CCB model emphasizes the accounting aspect, where it is possible to recognize negative values of debt sources (claims) and their maximum level is determined by the biggest company (from the point of view of balance sheet total) in the area].

A value from the interval [0, 1] that expresses the grade of membership to the given set "acceptable" is assigned to each level of debt sources. This assignment can be done with help of calculation types, charts (fields) or with definition of the membership function, the membership function of the set "acceptable". This membership is possible with use of calculation types, charts (fields) or with definition of the membership function. The membership function to a set could be of the following form:

$$A\ acceptable \begin{cases} 0, if\_CK < 0.9 mil.CZK \\ \frac{CK-0.9}{1-0.9}, if\_0.9 < CK < 1 mil.CZK \\ 1, if\_1 < CK < 10 mil.CZK \\ \frac{10-t}{20-10}, if\_10 < CK < 20 mil.CZK \\ 0, if\_CK > 20 mil.CZK \end{cases}$$

The fuzzy set is created by elements x selected from the set U, $x \in U$, where to each element is assigned a number $a \in [0, 1]$, i.e. the grade of membership of the element into the fuzzy set A. It is a set of ordered pairs – its grade of membership. It is possible to explicitly write the fuzzy set as:

$$A = \{a1 / x1, ..., an / xn\}, \tag{1}$$

where $x1, ..., xn \in U$ are elements with associated grades of membership $a1, ..., an \in (0,1]$, i.e. elements with grades of membership 0 are not included. If universe is not the final set and it is not possible to be written down as a list of elements, it is possible to write it down as follows:

$$A = \{a1 / x1 \mid i \in I\}, \tag{2}$$

where i is an index set or it is possible to specify the characteristics of xi and ai in a more detailed way. If for example the elements x are real numbers and the grades of membership are defined by a function, the fuzzy set may be written down as follows:

$$A = \{ f(x) / x \mid x \in R\}, \tag{3}$$

where R is a set of real numbers. Other specific sets (carrier, a-cut and core) are further used only in theoretical way to write down the analyses results, however the CCB uses this knowledge (recording the value of a selected ratio in the form of singleton.) The carrier of the fuzzy set A is a set of all elements of the universe, whose grade of membership into A is not zero.

This set is very important as it contains all elements that are interesting for those who elaborate the financial analysis. The elements whose grade of membership is zero are not interesting as they can be absolutely arbitrary.

$$Supp\ (A) = \{ x \mid A(x)>0\}, \tag{4}$$

A-cut is a set of elements whose grade of membership is at least (i.e. bigger or equal) the given grade a. This set can be obtained from the fuzzy set cutting all elements whose grade of membership is less than a.

$$Aa = \{ x \mid A(x) \geq a\}, \tag{5}$$

For a-cuts of fuzzy sets the following applies: if a≤b, then Ab $\subseteq$ Aa

Equality the so-called clause about representation of a fuzzy set means, that the grade of membership of the element x into the fuzzy set A equals the supreme of all indexes and cuts according to the image:

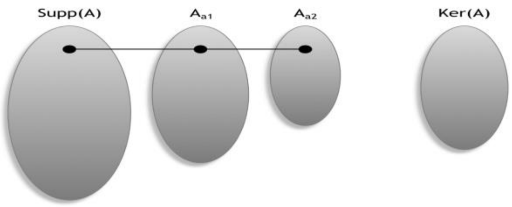

Figure 2: The supreme of all indexes. (Source: Author)

The core is a set of those elements that for sure belong in the fuzzy set A. They represent typical elements for the given fuzzy set. It will be linguistically typically "acceptable" from 1 million Czech crowns to 10 million Czech crowns.

$$Ker(A) = \{ x \mid A(x)=1\}, \tag{6}$$

A fuzzy one-element set (singleton) or a fuzzy unit {a /x} plays an important role too. The term of a fuzzy unit is very important especially in fuzzy regulation, because it is the way how to understand the result of a specific result, for example the ratio – the debt ratio (CK/VK) of 0,400 will be in the form {1/0.400} (Klepárník 2003).

According to Klepárník (2003) aspects of fuzzy approach to modeling can be:

- Modeling of vagueness and non-stochastic uncertainty of values and intensity of relations,
- Expert evaluation with use of grades of membership and linguistic terms (variables),
- Transformation of fuzzy models into classical models that can be solved with use of existing methods, eventually creating of new special methods.

Within financial analysis fuzzy sets, thus enable to solve the character of some "attributes" that are often vital for partial characteristics of the bankruptcy state. However, methodology of fuzzy sets (including work with their elements) does not allow determining concretely what sets are going to be analyzed.

This again confirms the presumption, that cooperation of two analysts is ideal for financial analysis elaboration. If all the analytical activity and prediction is done only e.g. with use of calculation software, that is able to connect both procedures, then it can be thought that it is much more important to select the data and its attributes (see "acceptable") than to do the mathematical work with the data.

The same point of view can be used even in case of fractional geometry that, as it is often stated, does not methodologically serve to predict (bankruptcy, financial health), but knowledge resulting from fractals must be seen as a new way of data elaboration. The selection of analyzed data and information depends again on the decision of the analyst or the person who presents the financial analysis. In relation with the hypothesis of the effective market it is good to remind here that for a long time it was generally assumed, that prices of securities reflect more or less arbitrary events in the world of business and that their fluctuation is arbitrary.

However, if that was so, the bankruptcy prediction, eventually reasoning based on time change of the security of the analyzed entity, would be absolutely useless. The first described point of view was introduced by Louis Bachelier, who worked on modeling price fluctuation. Later on it became apparent, that his pattern was very approximate and that it significantly underestimates the probability of a relatively big price fluctuation. Benoit Mandelbrot later proposes to express price fluctuation with use of the so called Lévy distribution. It is characterized especially by the fact that for big fluctuations it gets closer to power distribution in the form:

$$P(\Delta x) \approx (\Delta x) -1-\alpha, \tag{7}$$

The power distribution is almost always a manifestation of the mentioned fractional geometry. And a power price distribution indicates the fractional characteristics of price fluctuation. A more subtle manifestation of fractionality in economics is the so called scaling, that observes change of price with a simultaneous change of time. The characteristics scaling means: the probable distribution will not change if the prices scale does not change either. And thus the price distribution does not depend on two variables, time interval Δt and price change Δx, but on one variable only, and that is the selected combination (e.g. rise and decline of prices of different commodities develop in the same way in time - they decline in the time of crisis and rise in the time of recovery, however it does not matter if recession period is one day in a month, one month in a year or one year in a century. It can be expressed as follows:

$$P\Delta t(\Delta x) = (\Delta t)\text{-}H\ f\ (\ (\Delta t)\text{-}H\ \Delta x\ ), \tag{8}$$

where f is a function that decreases as power f (y) ≈ y−1−α, and H constant is the Hurst exponent. Bachelier´s idea of price fluctuation as a random walk meets exactly the characteristics of scaling with the value H = 1/2. The problem is, that real data indicate a significantly higher value around H = 2/3. Apart from the power form of the scaling function f the higher value of the Hurst exponent is one of the reasons why a usual random walk is absolutely inappropriate as a model of stock exchange fluctuations and it is necessary to search for something substantially better (Slanina 2006).

Any fractal object can be generally defined with use of two basic characteristics: (1) repeated similarity of the basic form that appears in one object in different sizes, (2) all near points are strongly correlated (for example during crisis the number of entities active in business decreases no matter if it is a global crisis, national crisis, crisis in the specific industry, regional crisis or a crisis between suppliers). These characteristics create appropriate conditions for prediction of time series that are important for financial analysis. Fractal geometry states, that even in an absolute chaos it is possible to find situations that repeat and describes more complex structures that use more simple structures that are their part (Sojka and Mendelík 2001). In an apparent chaos fractals represent the repeating elements (Moravec 2006). The market development has a number of chaos creating factors. R. N. Elliot was regularly observing hourly data on a New York stock exchange between 1935 and 1947. Thanks to the accumulated data he was able to describe market behavior. The so-called Elliot wave was then named after him.

2.1. The issue of data analysis with use of expert systems and neural networks

Data analysis tools are also often integrated into database and information systems. They are tools for the use of data warehouses and part of analytical elaboration (i.e. data mining issue). Further development in this area will probably mean combination of individual technologies in order to get optimal approaches for different types of data files. It can be, for example, a combination of genetic and neuron algorithms with decision trees. Nowadays, the term hybrid system in term of a combination of various algorithms can already be found in literature. Arminger et al. (1997) use problems with repayment to show the combination of the logistic discrimination analysis, classification tree and neuron network. As far as the program equipment is concerned, statistic program systems are the base. Apart from that, specialized products focused on decision trees or neuron networks are also offered. Among other types of software equipment with integrated data mining technology are relational database systems, systems supporting decision making. These procedures can be also used in the analysis of economic trends in the analyzed entity. There are different types of tasks (see Table 1) to be solved same as different procedures that can be used.

Table 1. Classification of data analysis.

Task	Method
Classification	Discriminant analysis, Logistic regression, Classification (decision) trees, Neuron networks
Predictions of values of the explained variable	Linear regression analysis, Non-Linear regression analysis, Neuron networks
Segmentation (clustering)	Cluster analysis, Genetic algorithms Neuron clustering (Kohonen maps)
Relation analysis	Association algorithm for derivation of rules like If X, then Y

When classifying and estimating values with some entities, it is possible to use values of explanatory variables as well as values of the explained variable. The aim is to analyze the influence of explanatory variables on the explained variable, so that it was possible to estimate the value of an entity with unknown value of the explained variable. In the terminology of neuron networks it is supervised learning. The same principle is also used with prediction of time series (Řezanková 2001). (The use of neuron networks is justified in case when either during a problem solution it is not possible to mathematically describe all relationships and contexts, that influence the observed process or in cases when it is possible to create the exact mathematical model, but it is so complicated, that it is almost impossible to algorithmize the task. They are especially complex and non-linear systems. Among the biggest advantages of artificial neuron networks is the ability to learn. It means to acquire knowledge by learning with use of a set of presented formula without necessity of knowing the algorithm of solution. Neuron networks may be used in financial analysis for prediction of time series and eventual consecutive decision making.

In segmentation (clustering) the data file is divided into groups and thus clusters of objects are created. The appropriate number of groups is usually identified in the course of the data analysis. Association algorithm is used in the relationship analysis to acquire rules, i.e. implications of IF (logical combination fact) THEN fact, where fact is an elemental logical statement. It is found out, what percentage of a specific logical combination of facts (antecedent) implies a fact on the right side of the rule and what percentage of records can be found in this association. Detection of deviations can be made with use of a graph with original (identified) values (a correlation graph XY) or statistical characteristics of the file. The basis for the data analysis is creation of a model that represents the data set. There are several modeling techniques, within which there are a lot of different approaches. The aim of modeling in case of classification decision trees is creation of a tree structure. There is a number of different algorithms for categorical data. These algorithms can be combined.

In this way it is possible to identify high-risk entities from a selected sector of economy (i.e. to select possible entities endangered by financial distress) on the basis of:

- Their property / capital structure,
- Size of the unit,
- Market orientation of the company,
- Active years in business.

The algorithm CART is in this case based on the fact, that from each nude that is not final (the last) lead two branches (other methods admit more branches, and the maximum number of branches depends on the number of categories of the variable, that serves as a predictor).

The procedure of creating the model can be divided into 3 steps:

1. Selection of variables from the data file. The user chooses the variable whose values are to estimate (financial distress, financial structure of the company). Then he/she chooses the explanatory variables used to carry out the estimation.

2. Creation of the groups of values. On the basis of statistical tests (chi square, see further) two groups of values are created for each explanatory variable. The principle of the test is to find out what are the two groups with the biggest variability between the groups and the smallest variability inside the groups. It is an iterative process.
3. Creation of a tree structure. A variable that contributes to the greatest extent to estimation of the explained variable is identified. On its basis the tree structure is created (Řezanková 2001).

Among the disciplines that are a basis for data mining are especially intuitive learning, machine learning and statistics. They are also basis for the models used in data mining. There are 7 types of models (table 2) used for solution of standard problems:

Table 2. Data mining models.

Model	Description of behavior	Prediction
Classification		×
Regression		×
Time series		×
Neuron networks		×
Clustering	×	
Exploratory analysis	×	
Association analysis	×	

It is obvious that, when using the techniques of data mining, it is necessary to use the appropriate software. Expert writings remind (Klímek 2005), that success of data mining depends not only on a good choice of the model and statistical method, but especially on a good formulation of the problem and use of correct data. That again confirms the assumption of the CCB model about the necessity of precise selection of data to be analyzed and then the use of methods for the data elaboration.

Recommendation about the form of the analytical tool for data mining is based on the following characteristics:

- Analytical abilities. The tools for data mining contain currently most used methods for data mining: decision trees, clustering and modeling with use of neuron networks and a number of other algorithms. There is different scope and possibilities of parameterization for different products, the products also support creating of proper models.
- Ease of use and analytical work. Creation of a specific model in the frame of data mining is often an iterative and complicated process. In clustering analysis it is usual to try an optimal method (e.g. K-means, Kohonen network, probability model) and testing of the optimal number of clusters. If a multi-layer neuron networks are used, the behavior can be radically changed due to the change of number of neurons, way of normalization of the input data etc. These and other reasons lead to a requirement of intuitive environment for creation, administration, connection and continuous evaluation of models and source and modified sets.
- Connectivity for input data, operability of work with data sets. All mentioned tools enable to flexibly import the input data with use of different sampling methods, to create their subsets and manipulate with them flexibly. Different kinds of functions for input data transformation are supported as well, e.g. filtering, normalization, compensation of values, change of distribution characteristics etc.
- Visualization and statistical evaluation of data and results of the models. Profiling, visualization and statistical elaboration of the input data and analyses results is always a part of the project of data mining (Klímek 2005).

3. Materials and Methods

Point estimates, as first of the group of statistical methods of financial analysis of a company, are used for a rough $\bar{x}$ estimate of a normal or comparative value of a specific indicator for a group of companies. The value of the point estimate quantitatively represents the whole file (arithmetic mean standard deviation σ). In case that certain statistical assumption do not apply (which is very typical in the $\bar{x}$ economic reality), the use of point estimates is not appropriate due to their sensitivity to remote data, that is usually present in financial data. In order to remedy robustness (sensitivity to remote data) point estimates of another class, the so-called ordinal statistics [median] are used, (Průcha 2005). In scientific literature it was already proven, that relative measures (ratio of two selected items) do not give the real image of the financial situation of the company, if they are compared with average values of the indicators of the so called corresponding undertakings. Mr. And Mrs. Kovanic state, that the assumption, that both numerator and denominator of a ratio indicator are directly proportional to the size of the company, is complied with only exceptionally. (Kovanicová and Kovanic 1996). If the text of this publication is primarily based on elemental statistic methods, we assume, that for the estimation of the financial situation of a company focus should be put especially on point and interval forecast, that however from the statistical point of view are based on absolutely different characteristics than the mentioned point estimates.

Forecast construction based on models with variable regimes that serve for forecast is generally characterized in the following A general non-linear model of the order one in the form:

$$Xt = G(Xt\text{-}1, \delta) + at, \tag{9}$$

where G is a non-linear function and δ is a vector of parameters. An optimal point forecast is the conditional median value. An optimal forecast with the horizon h made in time T can in general be expressed as:

$$XT(h) = E(XT{+}h \mid \Omega T), \tag{10}$$

where ΩT expresses history of the time series until time T included, i.e. a line of values XT, XT-1, XT-2, ... The optimal forecast with horizon one is in the form of:

$$XT(1) = E(XT{+}1 \mid \Omega T) = G(X1, \delta), \tag{11}$$

It is a forecast that is constructed in a similar way as in case of linear models. The situation becomes more complicated though, if the forecast horizon increases. The optimal forecast with horizon two is in the form:

$$XT(2) = E(XT{+}h \mid \Omega T) = E[G(Xt{+}1, \delta) \mid \Omega T], \tag{12}$$

The basic problem is, that in contrast with linear models, inequality applies $E[G(Xt+1, \delta) \mid \Omega T] \neq E[G(Xt+1 \mid \Omega T), \delta]$ which means, that the linear operator of the conditional median value E cannot be exchanged with non-linear operator G. It is possible to conclude, that in case of non-linear models there is no simple relationship between forecasts with different horizons, same as in case of linear models. In this context several possibilities of construction of point forecasts appeared. They may be based on the relationship $XT(2) = E(XT+2 \mid \Omega T) = E[G(Xt+1, \delta) \mid \Omega T) = E[G(G(Xt, \delta) + at+1, \delta) \mid \Omega T] = E[G(XT(1) + at+1, \delta) \mid \Omega T]$

Supposing, that at+1 = 0, it is possible to remove the conditional median value from the relationship and then the forecast can be written down as:

$$XT(2)(n) = G(XT(1), \delta), \tag{13}$$

In literature this type of forecast is denominated as naive (Kohout 2005). From this denomination it is evident, that it is possible to construct more appropriate forms of point forecasts. One of them is the closed form of forecast that can be expressed as follows:

$$XT(2)(n) = \int_{-\infty}^{\infty} G(X_{T+1}, \delta) f(X_{T+1} \vee \Omega_T) dX_{T+1}, \tag{14}$$

where f (XT+1 ∣ ΩT) is conditional density of probability XT+1 on condition ΩT. It is direct expression of point forecast J in form of calculation. One of the problems of this point forecast is, that the integral is not generally available in analytical form and it is necessary to approximate it numerically, this activity can be, especially in case of a forecast with high horizon, relatively time demanding. Another problem is lack of knowledge about distribution of the variable at. In order to achieve a complex analysis of forecasts of time series it is efficient to use interval forecasts (constructed on the basis of linear models) whose typical characteristics is symmetry around point forecasts.

This fact is based on the assumption of normality of conditional distribution of the variableXT+h with median value XT(h) that is the basis for these models. In case of non-linear models the conditional distribution can be asymmetric and multimodal. And thus it is a question, if the symmetric forecast interval is a good choice. In literature (Hyndman 1995) there already have been three proposals of the way of construction of forecast intervals for non-linear models. Symmetric model around median value:

$$S\alpha = (XT(h) - \Delta, XT(h) + \Delta), \tag{15}$$

where Δ shall be such, that P(XT+h Sα∣ΩT) = 1 – α. The interval between 100(α/2)% and 100(1- α/2)% of the quantile of the forecast distribution Qα = (qα/2, q1-α/2) (2.23)

Area with highest density:

$$HDR\alpha = \{XT+h \mid [f(XT+h \mid \Omega T) \geq f\alpha]\}, \tag{16}$$

Where fα is such, that P(XT+h HDRα∣ΩT) = 1 – α. In case of symmetric division with one mode the above mentioned forecast intervals are identical, in case of asymmetric distribution or distribution with more modes, the forecasts are not identical. The most natural is the third forecast interval, as it is the smallest of all possible intervals
100(1- α) % of forecast intervals and each point within this interval has conditional density of f (XT+h∣ΩT) at least the same as every point outside of the interval. It is interesting to compare the quality of forecasts constructed on the basis of linear and non-linear models.

However, the proposed bankruptcy model must be able to count with an eventual mistake (eventually its variance). Therefore, it is obvious, that despite the validity of conclusions of point and interval forecasts, the model must admit an error in case the horizon h increases.

We believe that the mere statement about the relationship of the accuracy of the forecast and horizon, that indicates its indirect dependence, is already sufficient here. We assume, that in case of any quantitative estimates of time series, that are a result of extension of development from the past till the present, the known assumptions are too unreal (e.g. Selection of the model, permanence of the model parameters). In this estimation the analyst examining the data is able to exhibit an error that is within the CCB model constructed at the level:

$$\hat{a}T+h = yT+h - \hat{y}T(h), \tag{17}$$

The error supposed by the model is further divided in two parts âT+h = ĉS1 + ĉS2 where each of the parts ĉS is explained by the CCB model in a different way ĉS1 refers to the model selection and ĉS2 is allocated to the estimates of parameters by the proposed model.

The relationship can thus be noted as âT+h = (yT+h – YT+h) + (YT+h – ŷT (h)), where yT+h – YT+h is dedicated to the error made by the model selection (in case an appropriate model is selected yT+h – YT+h = 0) and YT+h – ŷT (h) explains the error made by the estimate of the model parameters. For exact extrapolation the CCB model further requires, that the forecast was an undistorted and a solid estimate E {(âT+h)} = E {(yT+h – ŷT (h))} = 0 (2.27) E {(âT+h)2} = E {(yT+h – ŷT (h))2} = σ_p^2 ⇒min.

4. Conclusions

In our opinion, the financial analysis of a company would be considerably complicated, if the analyst did not know not only the purpose of its processing, but especially the object of its activity of the analyzed company. The object of the activity of the analyzed company will without doubt be influenced by the extent of extensive indicators. The extent of extensive indicators (tax liabilities, receivables from clients, basic capital, etc.) is certainly varied, depending on the object of the activity. In case of industrial companies this extent will be much bigger than for example in financial companies. E.g. expected return on debt from a group of textile manufacturing companies (and furthermore its development) should then reflect the situation in textile industry as a whole. Due to the uncertainty of point estimates, interval forecasts with a concrete horizon were used. This forecast precisely reflects the character of debt yields (and furthermore its development) of the present and past development, however the use of the same method for prediction is not possible anymore, as increase in error rate burdens the total result. Therefore, in this case it is not possible to use the past data to describe future development of the debt yield in textile industry as a whole.

Acknowledgments: The paper was written with the support of the specific project 2020 grant "Determinants of cognitive processes impacting the work performance" granted by the Faculty of Informatics and Management University of Hradec Kralove, Czech Republic and thanks to help of Ph.D. student František Hašek.

Conflicts of Interest: The authors declare no conflict of interest.

References

Arminger Gerhard, Enache Daniel, and Bonne Thorsten. 1997. Analyzing Credit Risk Data: A Comparison of Logistic Discrimination Tree Analysis, and Feed forward Networks. *Computational Statistics*: 12(2), 293-310.

Hálek Vítězslav. 2013. *Predikce finanční tísně podniku na základě vlastního bankrotního modelu CCB.* DonauMedia.

Hyndman Rob J. 1995. Highest-Density Forecast Regions for Nonlinear and Nonnormal Time Series. *Journal of Forecasting:* 14, 431-441. https://doi.org/10.1002/for.3980140503.

Kainz Wolfgang. 2010. Mathematical Principles of GIS. Available online: http://homepage.univie.ac.at/Wolfgang.Kainz/Lehrveranstaltungen/15th_Nordic_Summer_School/The_Mathematics_of_GIS_Draft.pdf (accessed on 18 January 2020).

Klepárník Jan. 2003. Návrh měřícího a regulačního systému pro solární systém ZTB. VUT, VUT-EU-ODDI-3302-02-03.

Klímek Petr. 2005. Úvod do problematiky data miningu. Praha: *Česká statistická společnost*: 16(3), 12-19.

Kohout Pavel. 2005. *Ekonomická analýza výnosových křivek*. Praha: PPF, Vol. 3. pp. 211-225.

Kovanicová Dana, and Kovanic Pavel. 1996. *Poklady skryté v účetnictví.* Part III. 1st ed. Praha: Polygon.

Mareš Milan. 2002. Algebraické vlastnosti fuzzy množin. Praha: UTIA AV ČR. Available online: http://www.statspol.cz/robust/ (accessed on 18 January 2020).

Moravec Zdeněk. 2006. Fraktály. Brno: Masarykova univerzita. Available online: http://is.muni.cz/th/106681/prif_b/fractalv08.pdf (accessed on 18 January 2020).

Navara Mirko, and Olšák Petr. 2002. *Základy fuzzy množin.* Praha, Skriptum ČVUT.

Průcha Ladislav. 2005. *Popisná statistika*. Praha: ČVUT. Available online: http://math.feld.cvut.cz/prucha/ubmip/u11.pdf (accessed on 18 January 2020).

Řezanková Hana. 2001. *Postupy používané při analýze dat*. Praha: VŠE, BADAME.

Slanina František. 2006. *Jak fyzika pomáhá ekonomii*. Praha: Fyzikální ústav AVČR, 2006, pp. 142-147. Available online: http://www.otevrena-veda.cz/ov/users/Image/default/C2Seminare/MultiObSem/113.pdf (accessed on 18 January 2020).

Sojka Zdeněk and Mandelík Petr. 2001. *Cenné papíry a burzy*. 2nd ed. Brno: PC-DIR.

Zadeh Lotfi A. 1965. Fuzzy Sets. *Information and Control*: 8(3), 338-353. https://doi.org/10.1016/S0019-9958(65)90241-X.