

## Cash Flows Indicators in the Prediction of Financial Distress

Michal Karas, Maria Reznakova

*Brno University of Technology*  
Kolejni 2906/4, Brno, Czech Republic  
E-mail. karas@fbm.vutbr.cz; reznakova@fbm.vutbr.cz

**crossref** <http://dx.doi.org/10.5755/j01.ee.31.5.25202>

*We argue that the conventional approach to bankruptcy modelling, which relies on accrual-based ratios, is vulnerable to the earnings management of a company threatened by insolvency. This fact may pose significant limits on the possibilities of distress prediction. Business distress is defined as cashflow insufficiency, and cashflow indicators are less vulnerable to earnings management. For these reasons we assume that cashflow ratios are theoretically more suitable for predicting distress. In our research we analysed the usefulness of cashflow-based ratios as potential predictors of bankruptcy. During the research, the cashflow-based ratios take the form of many variants of operating, financial, investment and free cashflow into a firm in combination with total assets, sales, liabilities and other indicators. The research was carried out on a sample of 4,350 Czech manufacturing SMEs operating during the period between 2013 and 2018. We employ the previously published approach of hybrid modelling to create the prediction model, though we propose a modification for the purposes of this paper. The modified hybrid model employs Classification and Regression Trees and Logistic Regression, while we use the Principal Component Analysis method to deal with the problem of multicollinearity. The results showed that operating cashflow ratios play a significant role in financial distress, especially when combined with short-term debts.*

Keywords: *Financial Distress; Prediction Models; Hybrid Model; SMEs; Logistic Regression.*

### Introduction

The current state of the art of financial distress could be divided into two main streams – accounting-based models and structural models. The structural model approach is considered by many to be superior to the accounting-based approach. However its application requires data from capital markets, which limits its application on a sample of private companies or limits application in environment with an underdeveloped capital market. We aim to contribute to research into the accounting-based distress model, and specifically in two ways. Firstly, by analysing the usefulness of cashflow-based indicators as predictors of distress, while addressing several different types of cashflow, which is, to our best knowledge, a rather novel approach. There are three main reasons behind our choice of analysing the cashflow indicators. First, the authors of accounting-based analyses often rely on traditional financial ratios (particularly profit-based ratios), which are often easy to manipulate by distressed businesses, with such businesses often having a tendency to do so in order to cover up their unfavourable situation. Second, from a theoretical point of view, there is a strong link between cashflow insufficiency and business distress, which highlights the importance of such analysis. Thirdly, the cashflow approach is a widely recognised way of estimating the business value, in case of businesses which are able to satisfy the going concern assumption, i.e. they are financially viable.

We also aim to contribute to the methodological issues of deriving a distress prediction model by suggesting a modified approach to deriving the model. The motivation behind this is our conviction that the potential for the model's

further improvement lies in fundamental factors that are often neglected by mainstream research. Firstly, the attention paid to the assumptions of the applied classification algorithms is not always sufficient and results in weak potential for generalising the model. As our focus is on a relatively specific group of variables, there is a high probability of the presence of multicollinearity in the initial sample; to deal with this issue we propose a modification to the previously published approach of hybrid distress modelling.

The aim of this paper is to introduce a new hybrid model that incorporates solely cashflow-based indicators, while we suggest a modified approach to hybrid model. In course of the research, three versions of the model were derived, while the ROC curves and corresponding Area Under Curve (AUC) served as measures for comparing the model's overall quality.

The following review of the literature provides a picture of the applied ratios used in the financial distress literature to date, as well as applied classification methods and their combinations (hybrid models).

### An Overview of Distress Predictors

As mentioned by Mai et al. (2019), researchers usually employ accounting-based variables and market-based variables (e.g. Altman, 1968; Deakin, 1972; Martin, 1977; Altman, Haldeman & Narayanan, 1977; Altman, 2000; Ohlson, 1980; Taffler, 1982; Zmijewski, 1984; Tam & Kiang, 1992; Shumway, 2001; Sanchez-Lasheras *et al.*, 2012 and many others). In addition, research in recent years has addressed the importance of governance indicators (see Liang, Lu, Tsai & Shih, 2016) and country characteristics (Duompos

*et al.*, 2017). Mai et al. (2019) introduced a model employing a combination of textual data and accounting-based and market-based variables. Li and Faff (2019) created a hybrid model combining accounting-based variables with market-based variables, with the loadings on both types of variable being non-monotonic, i.e. the hybrid models employ a regime-switching approach. Profit-based ratios play a significant role in the case of accounting-based ratios, while cash-flow-based ratios have been neglected by the mainstream of the literature. Altman (1968) summarised the importance of asset profitability (EBIT/total assets) in the following way: “Since a firm’s ultimate existence is based on the earning power of its assets, this ratio appears to be particularly appropriate for studies dealing with corporate failure”. Among Altman’s variables the return on assets (EBIT/total assets) is regarded as the strongest predictor (see Shumway, 2001). Jones (2016) pointed out that the return on assets (based on EBIT) is similar to cash-flow returns (operating cash flow over total assets) and that none of Altman’s studies tests cash-flow indicators (see Altman, 2002). He also stressed the fact that earnings are often subject to systematic management by companies, while operating cash flows are relatively more difficult to manipulate as they do not involve accruals or deferrals of any kind (see Jones and Belkaoui, 2010). Bledinger and Michalski (2018) were concerned about the way in which the long-term value added competitiveness can be measured and conclude that there should be a focus on cash-flow factors in addition to EVA, ROCE, ROE and ROA. The situation is even more complicated in the context of bankruptcy prediction as “*distressed companies have a high propensity to engage in earnings management*” (see Jones, 2011). Suhaily, Rashidah, and Mahenthiran (2013) add, specifically for the Malaysian environment, that financial distress is significantly and positively related to fraudulent financial reporting. Etemadi and Tariverdi (2006) conclude that the final result of company operation is not profit but cash flow, and added that “*while profit is an artificial concept, cash flow is objective and real*” (Etemadi and Tariverdi, 2006; in Kordestani et al., 2011). Further arguments highlighting the importance of cash-flow-based indicators lie in the following facts: 1) Financial distress occurs when the business is unable to meet its mature obligations or, in other words, when the “reservoir” of liquid assets is exhausted, while the cash flow from operations can be viewed as the net inflow of liquid assets into the “reservoir”. The larger the inflows, the lower the probability of failure (see Beaver, 1966). This applies particularly to operating cash-flow-based indicators. 2) Another definition of distress has used the fair value of business assets to describe the situation. In line with this definition, distress arises “*when the total liabilities exceed a fair valuation of the firm’s assets with value determined by the earning power of the assets*” (see Altman, 1968). The business value is often described in terms of the discounted cash-flow method, in which the business value is given by the present value of free cash flow (see, for example, Pohl, 2017).

Furthermore, these arguments apply particularly to the SME segment, which is the subject of investigation in this research. The reason for this is that, from the financial perspective, SMEs can be considered a special segment of business that are particularly sensitive to economic downturns or face more obstacles over the course of their growth. Much

academic attention has been dedicated to estimating the financial constraints of firms in recent years (see, for example, Ullah, 2019; Erdogan, 2018; McGuinness *et al.*, 2018). Erdogan (2018) showed that firms with high growth potential may have to rely on limited internal funds, which in turn constrains their ability to make investments, with such obstacles having adverse effects on the growth of SMEs as a consequence. McGuinness et al. (2018) examine whether trade credit has helped financially constrained SMEs survive the recent financial crisis, while highlighting SMEs’ dependency on bank finance and their vulnerability to financing constraints. According to Jin et al. (2018) businesses experiencing higher financing constraints are more dependent on external funds and are therefore more sensitive to fluctuations on credit markets.

Cash-flow-based indicators are often mentioned as powerful predictors, especially in relation to total debt. Beaver (1996) was among the first to explore the ratio potential cash flow over total debt. However, the cash flow was defined only as the sum of net income and depreciation and amortisation. Ong et al. (2011) also came to the conclusion that cash flow over total debt is a powerful predictor of bankruptcy in the case of Malaysian companies; in their work the cash flow is defined in terms of EBITDA. EBITDA is often applied as a simplified surrogate of operating cash flow (see Mulford and Comiskey, 2002). The study by Welc (2017) provides a comparison of the power of EBITDA versus cash flow in bankruptcy prediction. Welc’s study mentioned several drawbacks of both types of measure. For example, the omission of working capital changes is often mentioned as a pitfall of EBITDA (Fridson and Alvarez, 2002). On the other hand, the cash flow also has drawbacks, such as sales of receivables accounts in factoring transactions or liquidation of inventories in “fire sales” (see Welc, 2017 for more detail).

For the above reasons, we believe that the potential of accounting-based ratios has not been fully exhausted, while cash-flow-based indicators are often neglected by the mainstream of bankruptcy prediction literature or the cash-flow value is often limited to merely the sum of net income and depreciation and amortisation.

### **Classification Algorithms and their Combinations Applied in Distress Prediction Models**

Historically, various algorithms have been employed to devise models of bankruptcy. The first was the linear discrimination analysis (LDA) method (Altman, 1968). According to Alaka et al. (2018) many further studies have simply adopted LDA without considering the assumptions that are to be satisfied for LDA models to be valid. Other algorithms were later applied in reaction to its shortcomings. A number of parametric methods exist, such as the probit model (Zmijewski, 1984), the logistic regression or logit model (Martin, 1977; Ohlson, 1980) and Cox’s model (Henerby, 1996; Shumway, 2001). Great interest has been seen in non-parametric methods, for example artificial neural networks with a back-propagation algorithm (ANN), e.g. Tam and Kiang (1992). Aziz and Dar (2006) provide an overview of other popular methods used in this field, of which we could name the Support Vector Machine (SVM), Rough Sets (RS), Case Base Reasoning (CBR), Decision Trees (DT) and Genetic Algorithms (GA). We can also find the application of

Data Envelopment Analysis (DEA) methods (Cielen, Peeters and Vanhoof, 2004; Ding, Song and Zen, 2008) and Classification and Regression Trees (CART), see Li, Sun and Wu (2010), etc. According to Brezinger-Masten and Masten (2012), the common feature of these methods is “a fully nonparametric specification of both the distributional form of variables and functional relations among them”. The importance of non-parametric methods in predicting bankruptcy is also highlighted in the study by De Andrés et al. (2011), who consider these methods especially useful for distress prediction due to the specific properties of financial data, such as non-normality and heteroscedasticity.

Mai et al. (2019) claim there is a trend in the recent literature for studying the combination of models, with the models being combined horizontally using ensemble techniques (see, for example, Kim & Kang, 2010) or vertically (see, for example, Du Jardin, 2016). Such combinations have come to be known as hybrid approach models. De Andrés et al. (2011) specify hybrid approach models as the combination of several classification methods achieving greater accuracy than single models.

Ahn and Kim (2009) suggest the combination of Case Base Reasoning (CBR) and Genetic Algorithms (GA), while the novelty of their approach lies in simultaneous optimisation of feature weighting and the instance selection of CBR using GA. De Andrés et al. (2011) combine a fuzzy c-means clustering method with Multivariate Adaptive Regression Splines (MARS) to create a hybrid model. The justification of employing clustering lies in respecting the idea that there are types of failing processes. Fuzzy c-means clustering allows data to belong to two or more clusters at once, while belonging to a cluster is represented by probability coefficients. The advantage of MARS methods lies in their high generalisability and the fact that they do not impose any assumptions about data features. Du Jardin (2016) argues, in line with De Andrés et al. (2011), that the process anticipating bankruptcy is not the same for all companies, but that there are several profiles, while the common weakness of the traditional approach is that it relies on identification of a given profile only. Du Jardin (2016) suggests a two-stage hybrid model. The given profile is identified by a Kohonen map in the first stage (see Kohonen, 2001) and the model specially derived for the profile is applied later.

Another way of combining classifiers is boosting, the principle of which is combining several “weak” models into one strong classification rule. The boosting algorithm was introduced by Friedman (2001). Using the boosting algorithm raises the accuracy of the classification algorithm, to which it is applied by progressively reducing the error term (Braun and Mues, 2012; Friedman, 2001). The resultant classification rule represents a set of many “weak” learners. The boosting algorithm is most often applied to CART, but an ANN application may also be encountered (Kim and Kang, 2010).

Alaka et al. (2018) provided a systematic review of methods applied in distress prediction research and showed that the method selection of many studies is not based on the capabilities of the given tool, but on popularity or professional background. Moreover, they proposed a guideline for choosing the proper method for the given data set features. Alaka et al. (2018) pointed out the “black box” nature of methods such as ANN or SVM and the fact that the coefficients assigned to the model’s variable are illogical and

extremely hard to interpret. The same interpretation drawback may also be seen in otherwise popular methods of LDA. On the other hand, the same study mentioned logistic regression methods (LR) as easy to interpret methods, whereas the importance of the selected variable “*is transparent and helps users identify key areas of the problem facing a failing firm*”. Furthermore, the LR model does not impose any assumptions about the distribution of independent variables (as LDA models, for example, do), though it should be stressed that LR models are extremely sensitive to multicollinearity (see Balcean and Ooghe, 2006).

From the viewpoint of interpretability, the model of Brezinger-Masten and Masten (2012) represents an example of an easy-to-interpret hybrid approach model. The approach of this study will be further discussed, as we used this approach with some modifications. Brezinger-Masten and Masten (2012) introduced two versions of their model. The first version incorporates only variables identified by the CART method, which were plugged into the logit model in the form of dummy variables. The second version of the model combines dummy CART selected variables with conventionally selected variables. The process of conventionally selecting variables could be described in the following steps in the mentioned study. In the first step, a logit model was created for each of the analysed variables (i.e. the approach of selecting variables was univariate in the first step) in such a way that significant predictors were selected. Secondly, the identification of correlated groups of variables was carried out, after which the PCA method was used to identify the main components of each group. The final step lies in using the logistic step-wise procedure. As a result the final classification rule can profit from some features of CART and LR, while the rule is easy to interpret. The procedure of creating a model in this paper was inspired by the above-mentioned approach. However, we slightly modified this approach during the course of the research. Our study deals with a set of similarly defined ratios, which is a cause of potentially significant multicollinearity. To deal with this issue, we apply the PCA method as the first step of reducing the initial set of variables, while the appropriateness of using the PCA method is tested *a priori* by the Kaiser-Meyer-Olkin test. Although PCA is a method of unsupervised learning it has been tested by several authors as a method of distress predictor selection, although the results showed that it has been superseded by other approaches (see, for example, Tsai, 2009). The point of using PCA lies in dealing with multicollinearity which would otherwise negatively influence the step-wise logistic procedure results and, thereby, the whole model.

The aim of this paper is to derive a distress prediction model that incorporates only cash-flow-based indicators and to address the methodology of creating the model by proposing a slightly modified approach to hybrid modelling.

### **The Sample and Ratios Under Analysis**

The analysis was performed on a sample of 4,350 small and medium-sized Czech companies, of which 40 companies were defaulted (bankrupt) and 4,310 non-defaulted. The analysed companies operated in manufacturing (NACE Rev. 2 Main Section C). The data was obtained from AMADEUS (Analysis Major Database for European Sources). The last

report on defaulted companies comes from between 2013 and 2018. The reason for focusing on small and medium-sized companies is that these companies are more vulnerable in the case of an economic recession than large companies or multinational companies (see Jin et al., 2018). The original sample was randomly divided into a learning sample (70 % of

the observations) and a test sample (30 % of the observations). For the research a set of 31 cash-flow-based ratios was collected, of which 17 ratios were selected by literature review and the remainder proposed by the authors of this paper.

Table 1

List of Ratios Under Analysis

Abbrev.	Description <sup>(source)</sup>	Abbrev.	Description <sup>(source)</sup>
OCF/CL	Operating cash flow/current liabilities <sup>(2, 6)</sup>	FCFF/TA	Free cash flow/total assets <sup>(10)</sup>
OCF/INT	(Operating cash flow + interest expenses + tax)/interest expenses <sup>(2, 6, 9)</sup>	FCFF/TL	Free cash flow/total liabilities <sup>(10)</sup>
OCF/S	Operating cash flow/sales <sup>(2, 6)</sup>	FCFF/CL	Free cash flow/current liabilities <sup>(10)</sup>
OCF/TA	Operating cash flow/total assets <sup>(2)</sup>	FCFF/S	Free cash flow/sales <sup>(4)</sup>
EBIT/OCF	EBIT/operating cash flow <sup>(6, 7, 8)</sup>	FCF/OCF	Financial cash flow/cash flow <sup>(4)</sup>
OCF/TL	Operating cash flow <sup>1</sup> /total liabilities <sup>(4)</sup>	ICF/CF	Investment cash flow/cash flow <sup>(5)</sup>
OCF/CF	Operating cash flow/cash flow <sup>(10)</sup>	ICF/(OCF+CF)	Investment cash flow/(operating cash flow + financial cash flow) <sup>(5)</sup>
OCF/OC	Operating cash flow/operating cost <sup>(10)</sup>	CF/NWC	Cash flow/net working capital <sup>(10)</sup>
OCF/E	Operating cash flow/equity <sup>(10)</sup>	PCF/CL	Potential cash flow/current liabilities <sup>(10)</sup>
CF/S	Cash flow/sales <sup>(10)</sup>	PCF/INT	Potential cash flow/interest expenses <sup>(1)</sup>
CF/TA	Cash flow/total assets <sup>(10)</sup>	PCF/S	Potential cash flow/sales <sup>(1)</sup>
CF/TL	Cash flow/total liabilities <sup>(10)</sup>	PCF/TA	Potential cash flow/total assets <sup>(1)</sup>
CF/NW	Cash flow/net worth <sup>(3)</sup>	PCF/TL	Potential cash flow/total liabilities <sup>(10)</sup>
CF/CL	Cash flow/current liabilities <sup>(5)</sup>	PCF/NWC	Potential cash flow/net working capital <sup>(10)</sup>
CF/INT	Cash flow/interest expenses <sup>(10)</sup>	PCF/E	Potential cash flow/equity <sup>(10)</sup>
FCFF/OCF	Free cash flow/operating cash flow <sup>(10)</sup>		

Source: 1 – Beaver (1966), 2 – Bhandari and Johnson-Snyder (2018), 3 – Brezigar-Masten and Masten (2012), 4 – Kadarova, Bajus and Rajnoha (2015), 5 – Shanmugham and Mahalakshmi (2015), 6 – Thomas Ng, Wong and Zhang (2011), 7 – Blum (1974), 8 – Martens et al. (2008), 9 – Li and Sun (2009), 10 – proposed.

A definition of employed types of cash flow is given below:

Table 2

Cash Flow Definitions

No.	Type	Abbrev.	Formula
1	Operating cash flow	OCF	EBIT*(1-T) + depreciation - (non-cash NWC (t) – non-cash NWC (t-1))
2	Total cash flow	CF	cash (t) - cash (t-1)
3	Free cash flow into the firm	FCFF	EBIT*(1-T) - (non-cash NWC (t) – non-cash NWC (t-1)) - (FA (t) - FA (t-1))
4	Potential cash flow	PCF	EAT + depreciation
5	Financial cash flow	FCF	long-term debt (t) - long-term debt (t-1) - interest expenses + shareholder’s funds (t) - shareholder’s fund (t-1) - EAT (t)
6	Investment cash flow	ICF	-((FA (t) - FA (t-1) + depreciation)

### The Methodology for Deriving the Model and the Modified Approach Proposed

As our research focuses on specific types of financial ratios, namely cash-flow-based indicators, a possible multicollinearity problem may occur as we have to deal with similarly defined indicators. The aim of the research is to create a reduced-form default prediction model. We employ Principal Component Analysis (PCA) to reduce the initial set of variables. Nevertheless, the PCA method is a tool of unsupervised learning procedures whose goal is not to seek a relationship to a target variable (the occurrence of default in the case presented here). Further methods of supervised learning have to be employed (in our case CART and logistic regression) for this reason. To derive the model, we employ the methodology inspired by Brezigar-Masten and Masten (2012) who proposed plugging CART variables into a logistic regression model and thereby created a hybrid model. However, our approach differs in two ways. Firstly, we focus on the usefulness of specific types of ratios only (cash-flow-based), and secondly, we employ the PCA method to reduce the initial set of variables.

### Principal Component Analysis (PCA)

PCA was used to select the variables for the model (see McNamara and Duncan, 1995; Issah and Antwi, 2017). The PCA method is a general technique for finding data patterns. This method is used to reduce the size of high-dimensional feature vectors. PCA feature reduction can be explained as follows. Suppose that  $M$  is a  $t$ -dimensional data set. The  $n$  principal axes  $G_1, G_2, \dots, G_n$ , here  $1 \leq n \leq t$ , are orthonormal axes on which the retained variance is the maximum in the projected space (see Avci and Turkoglu, 2009; Polat and Gunes, 2007). Commonly  $G_1, G_2, \dots, G_n$  can be given by the  $n$  leading eigenvectors of the sample covariance matrix

$$C = \left(\frac{1}{L}\right) \sum_{k=1}^L (x_k - m)^\top (x_k - m) \tag{1}$$

in here  $x_k \in M$ ,  $m$  is the mean of samples.  $L$  is the number of samples. According to this:  $UG_k = v_k G_k$ ;  $k \in 1; \dots; n$ , here  $v_k$  is the  $k^{\text{th}}$  largest eigenvalue of  $U$ .

### Classification and Regression Trees (CART)

The CART method was first used for the prediction of corporate default by Frydman et al. (1985). It is suitable for the selection of predictors (see, for example, Brezingar-Masten and Masten, 2012) and for the formulation of a decision-making rule (see, for example, Gepp and Kumar, 2015; Liang, Lu, Tsai and Shih, 2016). The main advantages of CARTs include: 1) the resulting classification rule is easy to interpret (Brezingar-Masten and Masten, 2012); 2) this is a non-parametric method which is also able to capture complex relationships between variables (Brezingar-Masten and Masten, 2012); 3) the method is very robust with regard to the existence of outliers in the sample (Di Marco and Nieddu, 2014). When applying the CART method, a frequently discussed topic is the stability of the created tree, i.e. the dependence of its structure on the studied sample. It is common that the created model is able to define companies in the sample of the companies that were used to create the model extremely correctly, though the accuracy is significantly lower outside the learning sample. We solved this problem by using k-fold cross validation with  $k = 10$ , similarly to Liang, Lu, Tsai and Shih (2016). In this approach, the studied sample is split into ten learning and test sets used to train and test the model, and the whole process is repeated

ten times. The final rule is derived as an average of all ten results obtained.

### Logit Model

For the purposes of classification, we employ the logit model, where the probability of the bankruptcy of the  $i^{\text{th}}$  company ( $\pi_i$ ) is given by the formula:

$$\pi_i = \frac{\exp(x_i)}{1 + \exp(x_i)} \tag{2}$$

where  $x_i$  is a vector of covariates and  $\beta$  is a vector of regression coefficients. The parameters  $\beta$  are estimated by maximising the following log-likelihood function:

$$\log L(\beta) = \sum \{y_i \log(\pi_i) + (n_i - y_i) \log(1 - \pi_i)\} \tag{3}$$

Maximisation is performed either by Fisher's scoring procedure or the Newton-Raphson method.

### Results

The Kaiser-Meyer-Olkin test (further referred to as the KMO test) and the Bartlett's test measure of sampling adequacy were used to examine the appropriateness of the PCA method.

Table 3

KMO and Bartlett's Test Results

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.608
Bartlett's Test of Sphericity	Approx. Chi-Square	213930.766
	df	465
	Sig.	0.000

Taking a 95 % level of significance, the PCA is valid as  $p < 0.05$ , for which reason we reject the null hypothesis  $H_0$  and accept the alternate hypothesis ( $H_1$ ) that there may be a statistically significant interrelationship between variables. The KMO statistic of 0.608 is greater than 0.50, hence PCA is considered an appropriate technique for further analysis of the data. The initial step of employing the PCA method was to find the appropriate set of principal components. The

contribution of each component to the total variance of the data is as follows – the first three components explain 42.4 % of the total data variance, while the first five explain 62.2 % of the total data variance. Furthermore, the contribution of additional components is quite low. The number of components was set to 3. For interpreting the PCA results the component matrix was used, which shows the correlation between the original variables and the component.

Table 4

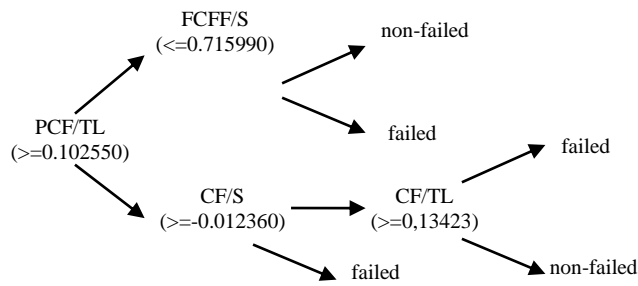
Component Matrix

Variable/ correlation coefficient	Component			Variable/ correlation coefficient	Component		
	1	2	3		1	2	3
CF/TL	<b>0.877</b>	0.338	0.010	CF/S	0.374	<b>-0.822</b>	0.415
FCFF/TL	<b>0.876</b>	0.337	0.014	OCF/S	0.367	-0.771	0.481
PCF/TL	<b>-0.823</b>	-0.302	0.086	OCF/OC	-0.124	0.681	-0.167
CF/CL	0.788	0.315	0.066	CF/TA	-0.161	0.364	<b>0.904</b>
PCF/CL	-0.785	-0.242	0.150	OCF/TA	-0.164	0.364	<b>0.902</b>
FCFF/CL	0.727	0.304	0.064	PCF/TA	-0.164	0.364	<b>0.902</b>
PCF/S	-0.388	<b>0.879</b>	-0.246	FCFF/TA	0.167	-0.364	<b>-0.901</b>
FCFF/S	0.375	<b>-0.823</b>	0.408				

For further analysis we chose only those variables that exhibit a correlation coefficient greater than 0.8 (such coefficients are highlighted by bold numbers). The 10 highlighted ratios were subjected to further analysis.

### CART Model Results

The derived tree model was pruned on the basis of the deviance measure. The final tree consists of four non-terminal nodes and five terminal nodes. We employ the k-fold cross-validation method at  $k=10$  to ensure coefficient stability. The node split criteria were further utilised to define a new set of dichotomous variables for the logit model, which is a procedure proposed by Brezingar-Masten and Masten (2012).



**Figure 1.** Estimated CART Model

Source: The author's own processing based on calculations on the Amadeus dataset

Four adjusted variables were defined in this way:  $dPCF/TL = 1$  if  $PCF/TL > 0.102550$  and 0 otherwise;  $dCF/S = 1$  if  $CF/S > -0.012360$  and 0 otherwise;  $dFCFF/S = 1$  if  $FCFF/S > 0.715990$  and 0 otherwise;  $dCF/TL = 1$  if  $CF/TL > 0.13423$  and 0 otherwise. Three models were derived during the course of the research. The parameters of Model 1 were estimated on the original sample of data with the use of the stepwise regression procedure. The parameters of Model 2 were derived on a subset of original ratios reduced by the results of PCA procedure application. Finally, Model 3 was derived on the same subset of ratios as Model 2, though the original continuous variables were replaced by their dichotomous versions in the sample. Another selection of the variables was needed as the original set of variables was modified, and this time stepwise logistic regression was utilised. Both forward selection and backward elimination were applied – both based on the Wald statistics.

Table 5

**Goodness of Fit Measures for the Derived Models**

Measure/Sample	Model 1		Model 2		Model 3	
	Learn	Test	Learn	Test	Learn	Test
AIC	1213.8821	495.6359	1450.038	634.817	188.028	132.699
BIC	1247.0875	523.8918	1467.445	649.760	222.841	162.585
Cox-Snell R <sup>2</sup>	0.5248	0.5491	0.548837	0.551520	0.7313	0.7203
Nagelkerke R <sup>2</sup>	0.6997	0.7321	0.7318	0.7354	0.9751	0.9604
Log-likelihood	-600.9410	-241.8180	-722.019	-314.409	-88.0139	-60.3494

The information criteria values (AIC and BIC) attained by Model 3 are the lowest of the derived models, which indicates the most preferable model. The same results are

given by the *pseudo R squares measures*, with the Nagelkerke R<sup>2</sup> for Model 3 attaining 0.97 on the learn sample and 0.96 on the test sample.

Table 6

**Hosmer-Lemeshow Test Results**

Model/statistics	Model 1		Model 2		Model 3	
	Learn	Test	Learn	Test	Learn*	Test
H-stat.	676.5733	280.1826	813.4932	336.8045	8.6377	41.1645
p-val.	0.000000	0.000000	0.000000	0.000000	0.373777	0.000002

According to the results of the Hosmer-Lemeshow test, only Model 3 can be considered well calibrated on the learn sample at a 5 % level of significance.

Model 1 was estimated on an original set of variables, while the variables entering the model were chosen by a stepwise regression procedure. Model 2 was derived on a subset of the original set of variables selected by application of the PCA procedure. Model 3 was utilised on a modified data set in which the original continuous variables were replaced by their categorical or rather dichotomous alternatives (while the cut-off was selected by the CART method).

**Estimation of Model Parameters**

All the mentioned models were derived by means of stepwise regression. Three models (Model 1, Model 2 and Model 3) were derived during the course of the research.

Table 7

**The Models' Estimated Parameters**

Model	Variable	Par. Estimate	Standard Error	Wald. (Stat.)	Lower CI (95%)	Upper CI (95%)	p-val.
1	PCF/TA**	-11.3682	1.456416	60.9276	-14.2227	-8.5137	0.000000
	OCF/TA**	-5.9277	0.919717	41.53971	-7.7303	-4.12508	0.000000
	FCFF/OCF*	-0.0197	0.008897	4.89542	-0.0371	-0.00225	0.026928
	PCF/CL**	-1.8423	0.395204	21.72971	-2.6168	-1.06766	0.000003
	CF/CL**	0.679	0.22877	8.81005	0.2306	1.12741	0.002996
	OCF/TL*	0.9909	0.389816	6.46155	0.2269	1.75492	0.011023
2	CF/TL**	1.2856	0.059643	464.6236	1.1687	1.40251	0.000000
	PCF/S*	0.0880	0.036601	5.7825	0.0163	0.15975	0.016187
	PCF/TL**	-10.0796	0.467225	465.4100	-10.9954	-9.16387	0.000000
3	d FCFF/S**	-5.58192	0.417925	178.3903	-6.40104	-4.76280	0.000000
	CF/TA*	-3.24856	1.312542	6.1257	-5.82109	-0.67602	0.013323
	FCFF/TA**	-2.58552	0.863463	8.9662	-4.27787	-0.89316	0.002750
	d CF/TL	0.50544	0.337794	2.2389	-0.15662	1.16751	0.134574
	d CF/S*	-0.66201	0.330559	4.0108	-1.30989	-0.01413	0.045210
	d PCF/TL**	0.88394	0.277385	10.1549	0.34027	1.42760	0.001439

Note: \*\*significant at the 1 % level, \*significant at the 5 % level



Model 1 contains six variables, all of which are statistically significant at the 5 % level at least. Model 2 contains only three variables, all of which are statistically significant at least at the 5 % level. Model 3 also contains the adjusted dichotomous variables, all four of these variables able to enter the model, and three of them (d FCFF/S, d CF/S, d PCF/TL) are significant at the 5 % level, while the last one – d CF/TL – is not significant at any standard level. A possible explanation for this may lie in the gradual way in which the CART procedure selected its variables, with the most significant variable first being set to split the variables space,

and the created subspace then being further split by another variable. Two continuous variables also enter the model – FCFF/TA and CF/TA, with the first mentioned variable being significant at the 1 % level and the second at the 5 % level.

### Model Testing Results

The ROC curves and corresponding Area Under Curve (AUC) were employed to test the models. Both the learn and test subsamples were analysed in this way.

Table 8

**Model Testing Results**

Sample	Model 1	Model 2	Model 3
Learn	0.73	0.81	0.93
Test	0.82	0.78	0.52

Model 3 exhibits the highest AUC on the learn sample, though the result obtained on the test sample is the lowest. On the other hand, the highest AUC value on the test subsample was obtained by Model 1.

### Discussion

Many financial distress models have been proposed to date and a variety of methods have been employed. The current trend in the financial distress literature clearly lies in exploring the possibilities of creating a hybrid model. A hybrid model can benefit from the features of the incorporated methods. Several authors (e.g. Alaka *et al.*, 2018) point out that hybrid methods are often hard to interpret. In this respect we used the modelling approach suggested by Brezingar-Masten and Masten (2012), with some proposed modifications. However, we did not focus on a number of areas of financial ratios, such as solvency, profitability and asset management, which is an approach that has become a paradigm in financial distress studies since the pioneer paper by Altman (1968). We argue that cash-flow-component ratios represent an area of potential predictors that has been rather neglected by the mainstream of research, although these ratios are strongly theoretically related to financial distress as they directly describe a business’s ability to generate inflows into the liquid asset reserve whose existence protects the business from insolvency (or distress), as has already been noted by Beaver (1966). The inability to generate such inflows results in an inability to meet short-term obligations, which is one of the most frequently mentioned reasons for going bankrupt (see Deakin, 1972; Gilson, 1989). Another relevant argument highlighting the practical usefulness of cash-flow-based indicators lies in the fact that these indicators are less vulnerable to earning management (see Jones and Belkaoui, 2010) for the reason that cash flow is not accrual-based like profit. The increase of net accruals is a possible way of earnings manipulation (see Jones, 1991). This phenomenon occurs frequently in cases of distressed businesses (see Jones, 2011 or Suhaily, Rashidah and Mahenthiran, 2013). The situation is further complicated by the fact that profit-based indicators are often considered the most significant predictors of distress (see Altman, 1968 or Shumway, 2001). Unlike many previous papers, we not only address the cash-flow indicator in terms of the sum of EBIT and depreciation, but also incorporate operating cash flow, financial cash flow,

investment cash flow or free cash flow. These indicators reflect both fixed assets investments and changes of net working capital (specifically in FCFF indicators or ICF indicators), which may be considered sufficient compensation for asset management indicators. Furthermore, operating cash flow over total assets (or rather sales) indicators are augmented alternatives to profitability ratios. Capital structure changes are also reflected in the cash-flow-component ratios utilised in this research, specifically in financial cash flow indicators. For the above-mentioned reasons we assume that the application of cash-flow-based ratios is possible and even desirable. The results obtained confirmed these assumptions.

In general, the idea of a reduced-form model is very common in financial distress studies (see, for example, Lin, Liang & Chen, 2011; Wang & Lee, 2008; Niemann *et al.*, 2008; Tseng & Hu, 2010; Psillaki, Tsolas & Margaritis, 2009; Cheng, Chen & Fu, 2006). Reducing the initial set of ratios is related to a risk pointed out by Scott (1981). Scott showed that such a reduced-form model may be less effective when applied to an alternative sample. This idea has also been supported by other researchers (see Platt & Platt, 1990; Grice & Dugan, 2001; Carling *et al.*, 2007; Wu, Gaunt & Gray, 2010). Comparison of the goodness of fit measures (such as information criteria – AIC, BIC or Area Under Curve) of the created model will lead to the conclusion that the suggested approach leads to a more significant model. Model 2 is more significant than Model 1, while Model 3 exceeds Model 2, although this is true only in in-sample evaluation, not in out-of-sample evaluation where the situation is opposite. In our specific case (unlisted SMEs, solely cash-flow indicators) the procedure of Brezingar-Masten and Masten (2012) leads to better results only in terms of the in-sample results. The same applies to the addition of PCA preselection of variables, which is promising only for in-sample results. The data on distressed companies was drawn one year prior to distress. According to Henerby (1996) cash-flow indicators are most significant three years prior to distress. Incorporating the time factor into the models may lead to enhanced model accuracy. During the course of the research a set of 31 cash-flow-based ratios was analysed, of which 14 have been proposed for the presented research. Focusing on significant ratios, there are six ratios in Model 1. Three of these ratios compare different levels of cash flow with the company’s debt, i.e. they consider business solvency (CF/CL, OCF/TL, PCF/CL). The ratio of total cash flow and current liabilities (CF/CL) was also

incorporated into the paper by Brezingar-Masten and Masten (2012). In their research this ratio entered the model only in its continuous form, not as its modified dichotomous variable, and this is in line with our results. On the other hand, the ratio of total or potential cash flow to total liabilities (CF/TL or PCF/TL) enters Model 3 in its modified dichotomous form. This may suggest that the ability to pay short-term debts is crucial in terms of business viability, though there may not be a sharp border distinguishing distress from vital businesses. If we change for total liabilities, however, a significant sharp border can be found. From the perspective of value creation, the ratios of cash flow to liabilities could be interpreted as cash flow to creditors or rather the ability to generate cash for loan providers. The role of external financing issues (represented by TL or CL factors in the above-mentioned ratios) is in line with expectations on financial constraints faced by SMEs (see Jin *et al.*, 2018) or the dependency of their survival on the extension of additional trade credit (as noted by McGuinness *et al.*, 2018).

The operating cash flow over total liabilities (OCF/TL) was also an indicator researched by Martens *et al.* (2008) who considered the prediction of a going concern auditor's opinion on a large sample of US listed companies; this indicator was not, however, part of their final model. The same applies to the study by Thomas Ng, Wong and Zhang (2011) which focused on predicting the distress of listed Chinese construction companies. Although the OCF/TL ratio was among the ratios analysed, it did not enter the final model. As the OCF/TL suggests the business's ability to generate cash to creditor for the business main operating activity, the other significant ratio of Model 1 – operating cash flow over total assets – may suggest the cash-generating ability to all stakeholders. Bhandari and Johnson-Snyder (2018) added that financially healthy companies are better able to generate cash more efficiently from their assets obtained through creditor and investor financing. This ratio is similar to the conventional return on assets (ROA) in that it uses total assets in the denominator; rather than using the net income as the numerator, as in the ROA ratio, this ratio uses operating cash flow. ROA is often mentioned as a significant distress predictor (see Altman, 1968; Li & Sun, 2009; Psillaki, Tsolas & Margaritis, 2009; Shumway, 2001). Nevertheless, the cash flow “*information has significant information content over accrual information in assessing the predicted probability of failure*” (Sharma, 2001 in Bhandari and Johnson-Snyder, 2018). From the methodological perspective, such a narrow focus on a given area of financial ratios poses a challenge to modelling as it necessarily results in high multicollinearity in

the initial sample of ratios. As both the logit model and step-wise logit procedure results are highly sensitive to the presence of such phenomena (see Balcean & Ooghe, 2006) in the feature variable space, the need for resolving this issue is unquestionable. For this reason we modified the procedure of feature selection suggested by Brezingar-Masten and Masten (2012), though we otherwise follow their methodology. The modification lies in the application of the PCA method as a first step of feature preselection. The advantage of this approach is that we avoid the arbitrary exclusion of correlated variables, while on the other hand the feature space reduction is not performed with any reflection on the discriminant power of the variables. The aim of the PCA method is to reduce the dimensionality of the feature variable space, while preserving as much original variance as possible.

## Conclusions

Accounting-based financial ratios are commonly used in predicting the financial distress of businesses. Distressed businesses often have a tendency towards earning manipulation covering their situation from stakeholders. Cash-flow-based ratios are less vulnerable to earning manipulation, as they do not employ accruals of any kind. The created models showed that cash flow can be utilised to predict business distress. The main difference from previously published studies lies in the fact that we address the term cash flow not only in the very simplified meaning of the sum of EBIT and depreciations and amortisation, which does not solve the issue of earning manipulation. We worked with cash-flow components such as free cash flow and operating cash flow, which takes into account changes of net working capital and fixed assets. We employed a hybrid model approach suggested by a previous study to create the model. We chose the hybrid approach because such an approach represents the current trend in this area of research. Our study deals with one type of ratio only, so we had to deal with potentially severe multicollinearity which was reason to modify the previously published approach to better fit the specific features of the analysed dataset.

There are some limitations to the results as is the case for every research result. Firstly, the analysed period was a relatively stable period of economic growth, which results in a rather low number of bankruptcies. Moreover, we focus on manufacturing SMEs and there is no guarantee that the results will be applicable to other industries. On the other hand, we present a methodology that could be useful for applications analysing businesses from other branches of the economy.

## Acknowledgement

This paper was presented at the 6<sup>th</sup> Annual Conference of the Academy of International Business Central and Eastern European (AIB-CEE) Chapter ‘International Business in the Dynamic Environment: Changes in Digitalisation, Innovation and Entrepreneurship’ (25–27 September 2019, Kaunas, Lithuania)

## References

- Alaka, H. A., Oyedele, L. O., Owolabi, H. A., Kumar, V., Ajayie, S. O., Akinade, S. O., & Bilal, M. (2018). Systematic review of bankruptcy prediction models: Towards a framework for tool selection. *Expert System with Application*, 94, 164–184. <https://doi.org/10.1016/j.eswa.2017.10.040>
- Altman, E. I., Haldeman, R. G., & Narayanan, P. (1977). ZETA Analysis. A new model to identify bankruptcy risk of corporations. *Journal of Banking and Finance*. 1(1), 22–54. [https://doi.org/10.1016/0378-4266\(77\)90017-6](https://doi.org/10.1016/0378-4266(77)90017-6)



- Altman, E. I. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *The Journal of Finance*, 23(4), 589–609. <https://doi.org/10.1111/j.1540-6261.1968.tb00843.x>
- Altman, E. I. (2000). Predicting financial distress of companies: Revisiting the Z-score and Zeta® models. Available from internet: <http://pages.stern.nyu.edu/~ealtman/Zscores.pdf> <https://doi.org/10.4337/9780857936080.00027>
- Altman, E. I. (2002). *Bankruptcy, Credit risk, and High Yield Junk Bonds*. Malden, MA: Blackwell Publishers. <https://doi.org/10.1111/1468-0300.00084>
- Avci, E., & Turkoglu, I. (2009). An intelligent diagnosis system based on principle component analysis and ANFIS for the heart valve diseases. *Expert Systems with Applications*, 36(2), 2873–2878. <https://doi.org/10.1016/j.eswa.2008.01.030>
- Aziz, M., & Dar, H. (2006). Predicting corporate bankruptcy: where we stand? *Corporate Governance*, 6(1), 18–33. <https://doi.org/10.1108/14720700610649436>
- Balcaen, S., & Ooghe, H. (2006). 35 Years of Studies on Business Failure: An Overview of the Classic Statistical Methodologies and Their Related Problems. *The British Accounting Review*, 38(1), 63–93. <https://doi.org/10.1016/j.bar.2005.09.001>
- Beaver, W. H. (1966). Financial Ratios as predictors of Failure. *Journal of Accounting Research*, 4, 71–111. <https://doi.org/10.2307/2490171>
- Bhandari, S., & Johnson-Snyder, A. J. (2018). A generic model of predicting probability of success-distress of an organization: a logistic regression analysis. *The journal of applied business research*, 34(1), 169–182. <https://doi.org/10.19030/jabr.v34i1.10107>
- Blendinger, G., & Michalski, G. (2018). Long-term competitiveness based on value added measures as part of highly professionalized corporate governance management of german dax 30 corporations. *Journal of Competitiveness*, 10(1), 5–20. <https://doi.org/10.7441/joc.2018.02.01>
- Blum, M. (1974). Failing company discriminant analysis. *Journal of Accounting Research*, 12(1), 1–25. <https://doi.org/10.2307/2490525>
- Braun, I., & Mues, C. (2012). An experimental comparison of classification algorithms for imbalanced credit scoring data sets. *Expert Systems with Applications*, 39(3), 3446–3453. <https://doi.org/10.1016/j.eswa.2011.09.033>
- Brezigar-Masten, A., & Masten, I. (2012). CART-based selection of bankruptcy predictors for the logit model. *Expert Systems with Applications*, 39(11), 10153–10159 <https://doi.org/10.1016/j.eswa.2012.02.125>
- Carling, K., Jacobson, T., Linde, J., & Rozsbaach, K. (2007). Corporate credit risk modelling and the macroeconomy. *Journal of Banking & Finance*, 31(3), 845–868. <https://doi.org/10.1016/j.jbankfin.2006.06.012>
- Cerny, C. A., & Kaiser, H. F. (1977). A study of a measure of sampling adequacy for factor-analytic correlation matrices. *Multivariate Behavioural Research*, 12(1), 43–47. [https://doi.org/10.1207/s15327906mbr1201\\_3](https://doi.org/10.1207/s15327906mbr1201_3)
- Cheng, C. B., Chen, C., L., & Fu, C. J. (2006). Financial Distress Prediction by a Radial Basis Function Network with Logit Analysis Learning. *Computers and Mathematics with Applications*, 51(3/4), 579–588. <https://doi.org/10.1016/j.camwa.2005.07.016>
- De Andres, J., Lorca, P., De Cos Juez, F. J., & Sanchez-Lasheras, F. (2011). Bankruptcy forecasting: A hybrid approach using Fuzzy c-means clustering and Multivariate Adaptive Regression Splines (MARS). *Expert Systems with Applications*, 38(3), 1866–1875. <https://doi.org/10.1016/j.eswa.2010.07.117>
- Deakin, E. B. (1972). A Discriminant Analysis of Predictors of Business Failure. *Journal of Accounting Research*, 10(1), 167–179. <https://doi.org/10.2307/2490225>
- Di Marco, L., & Nieddu, L. (2014). Trigger factors that influence bankruptcy: a comparative and exploratory study. *Rivista Italiana di Economia Demografia e Statistica*, 68 (3/4), 191–198.
- Ding, Y., Song, X., & Zen, Y. (2008). Forecasting financial condition of Chinese listed companies based on support vector machine. *Expert Systems with Applications*, 34(4), 3081–3089. <https://doi.org/10.1016/j.eswa.2007.06.037>
- Doumpos, M., Andriosopoulos, K., Galariotis, E., Makridou, G., & Zopounidis, C. (2017). Corporate failure prediction in the European energy sector: a multicriteria approach and the effect of country characteristics. *European Journal of Operational Research*, 262 (1), 347–360. <https://doi.org/10.1016/j.ejor.2017.04.024>
- Du Jardin, P. (2016). A two-stage classification technique for bankruptcy prediction. *European Journal of Operational Research*, 254(1), 236–252. <https://doi.org/10.1016/j.ejor.2016.03.008>
- Erdogan, A. I. (2018). Factors affecting SME access to bank financing: an interview study with Turkish bankers. *Small Enterprise Research*, 25(1), 23–35. <https://doi.org/10.1080/13215906.2018.1428911>
- Etemadi, H., & Tariverdi, Y. (2006). The effect of cash flow statement on the predictions of professional investors, *Iranian Accounting and Auditing Review*, 45, 69–86.
- Fridson, M., & Alvarez, F. (2002). *Financial Statement Analysis. A Practitioner's Guide*. New York: John Wiley & Sons,
- Friedman, J. (1991). Multivariate Adaptive Regression Splines, *The Annals of Statistics*, 19(1), 1–141. <https://doi.org/10.1214/aos/1176347973>

- Gepp, A., & Kumar, K. (2015). Predicting Financial Distress: A Comparison of Survival Analysis and Decision Tree Techniques, *Procedia Computer Science*, 54(1), 396–404. <https://doi.org/10.1016/j.procs.2015.06.046>
- Gilson, S. C. (1989). Management turnover and financial distress. *Journal of Financial Economics*, 25(2), 241–262. [https://doi.org/10.1016/0304-405X\(89\)90083-4](https://doi.org/10.1016/0304-405X(89)90083-4)
- Grice, J. S., & Dugan, M. T. (2001). The limitations of bankruptcy prediction models: Some cautions for the researchers. *Review of Quantitative Finance and Accounting*, 17(2), 151–166. <https://doi.org/10.1023/A:1017973604789>
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd ed. New York: Springer. <https://doi.org/10.1007/b94608>
- Henerby, K. L. (1996). Do Cash Flows Variables Improve the Prediction Accuracy of a Cox Proportional Hazards Model for Bank Failure? *The Quarterly Review of Economics and Finance*, 36(3), 395–409. [https://doi.org/10.1016/S1062-9769\(96\)90023-X](https://doi.org/10.1016/S1062-9769(96)90023-X)
- Issah, M., & Antwi, S. (2017). Role of macroeconomic variables on firms' performance: Evidence from the UK. *Cogent Economics & Finance*, 5(1), 1–18. <https://doi.org/10.1080/23322039.2017.1405581>
- Jin, Y. Luo, M., & Wan, C. (2018). Financial constraints, macro-financing environment and post-crisis recovery of firms. *International Review of Economics and Finance*, 55 (C), 54–67 <https://doi.org/10.1016/j.iref.2018.01.007>
- Jones, J. (1991). Earnings Management Following Import Relief Investigations. *Journal of Accounting Research*, 29(2), 193–228. <https://doi.org/10.2307/2491047>
- Jones, S., & Belkaoui, R. H. (2010). *Financial Accounting Theory* (3rd Edition), Sydney: Cengage.
- Jones, S. (2016). A cash flow based model of corporate bankruptcy in Australia. *Journal of Applied Management Accounting Research*, 3(2), 21–35.
- Kadarova, J., Bajus, R., & Rajnoha, R. (2015). Optimal financing of the industrial enterprise. *Procedia Economics and Finance*, 23, 953–958. [https://doi.org/10.1016/S2212-5671\(15\)00380-9](https://doi.org/10.1016/S2212-5671(15)00380-9)
- Kaiser, H. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31–36. <https://doi.org/10.1007/BF02291575>
- Kim, M. J., & Kang, D. K. (2010). Ensemble with neural networks for bankruptcy prediction. *Expert Systems with Applications*, 37(4), 3373–3379. <https://doi.org/10.1016/j.eswa.2009.10.012>
- Kohonen, T. (2001). *Self-Organizing Maps*. Springer Science & Business Media, 501 p. <https://doi.org/10.1007/978-3-642-56927-2>
- Kordestani, G., Biglari, S. V., & Bakhtiari, M. (2011). Ability of combinations of cash flow components to predict financial distress. *Business: Theory and Practice*. 12(3), 277–285 <https://doi.org/10.3846/btp.2011.28>
- Li, H., & Sun, J. (2009). Predicting business failure using multiple case-based reasoning combine with support vector machine. *Expert Systems with Applications*, 36(6), 10085–10096. <https://doi.org/10.1016/j.eswa.2009.01.013>
- Li, H., Sun, J., & Wu, J. (2010). Predicting business failure using Classification and Regression Tree: An empirical comparison with popular classical statistical methods and top classification mining methods. *Expert Systems with Applications*, 37(8), 5895–5904. <https://doi.org/10.1016/j.eswa.2010.02.016>
- Li, L., & Faff, R. (2019). Predicting corporate bankruptcy: What matters? *International Review of Economic and Finance*, 62, 1–19. <https://doi.org/10.1016/j.iref.2019.02.016>
- Liang, D., Lu, C., Tsai, C., & Shih, G. (2016). Financial ratios and corporate governance indicators in bankruptcy prediction: A comprehensive study. *European Journal of Operational Research*, 252(2), 561–572. <https://doi.org/10.1016/j.ejor.2016.01.012>
- Lin, F., Liang, D., & Chen, E. (2011). Financial ratio selection for business crisis prediction. *Expert Systems with Applications*, 38(12): 15094–15102. <https://doi.org/10.1016/j.eswa.2011.05.035>
- Mai, F., Tian, S., Lee, C., & Ma, L. (2019). Deep learning models for bankruptcy prediction using textual search. *European Journal of Operational Research*, 274(2), 743–758. <https://doi.org/10.1016/j.ejor.2018.10.024>
- Martens, D., Bruynseels, L., Baesens, B., Willekens, M., & Vanthienen, J. (2008). Predicting going concern opinion with data mining. *Decision Support Systems*, 45(4), 765–777. <https://doi.org/10.1016/j.dss.2008.01.003>
- Martin, D. (1977). Early warning of bank failure: A logit regression approach. *Journal of Banking & Finance*, 1(3), 249–276. [https://doi.org/10.1016/0378-4266\(77\)90022-X](https://doi.org/10.1016/0378-4266(77)90022-X)
- McGuinness, G., Hogana, T., & Powell, R. (2018). European trade credit use and SME survival. *Journal of Corporate Finance*, 49, 81–103. <https://doi.org/10.1016/j.jcorpfin.2017.12.005>
- McNamara, R., & Duncan, K. (1995). Firm performance and macro-economic variables. *School of Business Discussion Papers*. Paper 66.
- Mulford, C. W., & Comiskey, E. E. (2002). *The Financial Numbers Game: Detecting Creative Accounting Practices*. John Wiley & Sons, New York.

- Niemann, M., Schmidt, J. H., & Neukirchen, M. (2008). Improving performance of corporate rating prediction models by reducing financial ratio heterogeneity, *Journal of Banking & Finance*, 32(3), 434–446. <https://doi.org/10.1016/j.jbankfin.2007.05.015>
- Ohlson, J. A. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*, 18(1), 109–131. <https://doi.org/10.2307/2490395>
- Ong, S. Yap, V., & Khong, R. (2011). Corporate failure prediction: a study of public listed companies in Malaysia, *Managerial Finance*, 37(6), 553–564. <https://doi.org/10.1108/03074351111134745>
- Platt, D. H., & Platt, M. B. (1990). Development of a Class of Stable Predictive Variables: The Case of Bankruptcy Prediction. *Journal of Business Finance & Accounting*, 17(1), 31–51. <https://doi.org/10.1111/j.1468-5957.1990.tb00548.x>
- Pohl, P. (2017). Valuation of a Company using Time Series Analysis. *Journal of Business Valuation and Economic Loss Analysis*, 12(1), 1–39. <https://doi.org/10.1515/jbvela-2015-0004>
- Polat, K., & Gunes, S. (2007). Principles component analysis, fuzzy weighting pre-processing and artificial immune recognition system based diagnostic system for diagnosis of lung cancer. *Expert systems with applications*, 34(1), 214–221. <https://doi.org/10.1016/j.eswa.2006.09.001>
- Psillaki, M., Tsolas, I. T., & Margaritis, M. (2010). Evaluation of credit risk based on firm performance, *European Journal of Operational Research*, 201(3), 873–881. <https://doi.org/10.1016/j.ejor.2009.03.032>
- Sanchez-Lasheras, F., De Andres, J., Lorca, P., & De Cos Juez, F. C. (2012). A hybrid device for the solution of sampling bias problems in the forecasting of firms' bankruptcy. *Expert Systems with Applications*. 39(8), 7512–7523. <https://doi.org/10.1016/j.eswa.2012.01.135>
- Scott, J. (1981). The probability of bankruptcy: a comparison of empirical predictions and theoretical models. *Journal of Banking & Finance*, 5(3), 317–44. [https://doi.org/10.1016/0378-4266\(81\)90029-7](https://doi.org/10.1016/0378-4266(81)90029-7)
- Shanmugham, R. & Mahalakshmi, A. (2015). The best predictors of industrial sickness – accrual ratios or cash flow ratios? *International Journal of Management and Social Science Research Review*, 1(10), 202-209.
- Shumway, T. (2001). Forecasting Bankruptcy More Accurately: A Simple Hazard Model. *Journal of Business*, 74(1), 101–24. <https://www.jstor.org/stable/10.1086/209665>
- Suhaily, H., Rashidah, A. R., & Mahenthiran, S. (2013). Management Motive, Weak Governance, Earnings Management, and Fraudulent Financial Reporting: Malaysian Evidence. *Journal of International Accounting Research*, 12(1), 1–27. <https://doi.org/10.2308/jiar-50353>
- Taffler, R. J. (1982). Forecasting company failure in the UK using discriminant analysis and financial ratio data. *Journal of the Royal Statistical Society*, 145(3), 342–358. <https://doi.org/10.2307/2981867>
- Tam, K., & Kiang, M. (1992). Managerial applications of neural networks: the case of bank failure prediction, *Management Science*, 38(7), 926–947. <https://doi.org/10.1287/mnsc.38.7.926>
- Thomas Ng, S. T., Wong, J. M. W., & Zhang, J. (2011). Applying Z-score model to distinguish insolvent construction companies in China. *Habitat International*, 35(1), 599–607. <https://doi.org/10.1016/j.habitatint.2011.03.008>
- Tsai, C. (2009). Feature selection in bankruptcy prediction. *Knowledge-Based Systems*, 22(2), 120–127 <https://doi.org/10.1016/j.knosys.2008.08.002>
- Tseng, F. M., & Hu, Y. C. (2010). Comparing four bankruptcy prediction models: Logit, quadratic interval logit, neural and fuzzy neural networks. *Expert Systems with Applications*, 37(3), 1846–1853. <https://doi.org/10.1016/j.eswa.2009.07.081>
- Ullah, B. (2019). Financial constraints, corruption, and SME growth in transition economies. *The Quarterly Review of Economics and Finance*. In press. <https://doi.org/10.1016/j.qref.2019.05.009>
- Wang, Y. J., & Lee, H. S. (2008). A clustering method to identify representative financial ratios. *Information Sciences*, 178: 1087–1097. <https://doi.org/10.1016/j.ins.2007.09.016>
- Welc, J. (2017). EBITDA vs. Cash Flows in Bankruptcy Prediction on the Polish Capital Market. *European Financial and Accounting Journal*, 12(2), 91–103. <https://doi.org/10.18267/j.efaj.183>
- Wu, Y., Gaunt, C., & Gray, S. (2010). A comparison of alternative bankruptcy prediction models. *Journal of Contemporary Accounting & Economics*, 6(1), 34–45. <https://doi.org/10.1016/j.jcae.2010.04.002>
- Zmijewski, M. E. (1984). Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research*, 22, 59–82. <https://doi.org/10.2307/2490859>

The article has been reviewed.

Received in January 2020; accepted in December 2020.



This article is an Open Access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 (CC BY 4.0) License (<http://creativecommons.org/licenses/by/4.0/>).