Cash Flow Prediction Using a Grey-Box Model

Yang Pang[a]; Kwaku Opong[a]; Luiz Moutinho[a]; Yun Li[b]

a. Adam Smith Business School; b. School of Engineering

University of Glasgow

Glasgow, United Kingdom

Corresponding Email: y.pang.1@research.gla.ac.uk

Abstract— **This paper tackles the problem of financial forecasting by extending methods developed in automation, engineering and computing science. Current methods existing in the literature for firm-level cash flows are first analysed. Then a grey-box modelling method is developed to elevate the performance of cash-flow prediction. Linear panel data modelling is used as a benchmark model. Experiments with out-of-sample tests are used to validate the grey-box approach. Encouragingly, nonlinear grey-box modelling outperforms linear panel data modelling in both one-period-ahead and multi-period-ahead predictions.**

Keywords-Prediction; cash flow; panel data; grey-box

I. INTRODUCTION

Making prediction is an inevitable and crucial part in economic and financial analysis, as well as in business practice. Predictions usually lead to decision making. For instance, sales prediction will affect a firm's inventory management; forecasts made by financial analysts are used to construct portfolios, monetary and fiscal policies of a country are also made with respect to the forecast of the country's future economic state etc. Various methods and techniques have been developed to enhance predictive power of models used to forecast economic variables. In both economics and finance, academic research is important in explaining the associations and interactions between particular variables. Seeking high accuracy in prediction is of less concern. As a result, studies in these areas tend to adopt simple and parsimonious models, usually in the linear form. The Nobel winning capital asset pricing model (CAPM) is one example (see e.g. [1]), which is used to model expected return of equity market. Similarly, a stock return predictive model that is widely cited is also a linear model [2].

Applications in science and engineering, however, place greater effort in reaching high accuracy in predictions. Thus methods in engineering are generally more complicated. The border between these different disciplines is becoming blurry in some aspect. Many models that are widely applied in economics and finance were originally developed and used in the area of engineering, mathematics or other disciplines of science. For instance, Kalman filter [3] were developed to deal with problems in signal processing and it is now commonly applied in econometrics for estimating state-space models; neural network in the area of artificial intelligence has also been widely been applied to work in financial practice [4]. Besides, it is a common phenomenon that finance borrows ideas from or even merges with other fields, forming inter-disciplinary studies.

This paper investigates various modelling methods in cash flow prediction. This is an application with analysis of firm-level financial or accounting information. There have been previously developed prediction models to predict cash flows in accountancy and finance. This paper applies the grey-box model developed in [5] and investigates the potential dynamic and nonlinear features of the model to cash flows that have been overlooked in previous modelling procedures.

The structure of this paper will be as follows. Section 2 introduces the background relating to cash flow prediction. Section 3 discusses the models that will be applied in this piece of work and section 4 discusses the research designs and data sources. Section 5 will present the empirical results by analysing data with the proposed models. In the final, section 6 will be the conclusive part.

II. BACKGROUND

Cash, not merely as a number, is the true income of business transactions. Firms will pay their debts and expenses with cash so they will fail without cash incomes no matter how good their earnings number looks. Earnings are not and should not necessarily be treated as equivalent to cash flow. Before the introduction of statement of cash flow, income statement reflected a firm's revenue, expense and profit, but the numbers meant nothing if they could not be eventually translated into the same amount of cash. Due to trade credit, a transaction could occur when revenues and expenses are recognised in accounting even without immediate cash settlement. There is the risk that these pre-recognised amounts may not fully end up with the equivalent cash because of default for instance. As a result, it is uncertain whether accounting revenues and the cash income will be equal. Similarly, the accounting expenses are not necessarily equal to the amount of cash paid out in the current period. In addition, there are non-cash items being recorded as expenses such as depreciation, which affect earnings but not cash flow. Earnings that are the difference between revenues and expenses are therefore often unmatched with cash income. If cash is genuinely the measure of profit, accounting earnings do not accurately reflect it.

In the U.S, a statement of cash flow has been a compulsory part of financial reports since 1987 when statement of financial accounting standards (SFAS) No. 95 was published. Before the U.S, Canada was the first country which required cash flow disclosure in 1985. Thereafter, the time series property of cash flow started to be studied along with earnings. Before then cash flow was indirectly estimated by deducting accrual terms and

non-cash items from earnings. There were early studies on the time series property of cash flow and the tool used in this procedure is mainly under the ARIMA framework (see e.g. [6]). The ARIMA model is univariate model, therefore it ignores potential influence of exogenous variables. [7] states importantly that unlike earnings, univariate time series model may not be sufficient to model cash flow because there are predicable components in cash flows brought on by accrual terms (denote their model as DKW for short). It would be better for cash flow prediction to include more variables in addition to lagged cash flow. [8] therefore adjusted DKW model by disaggregating earnings into cash flow and accrual term components as multiple predictors other than using aggregated number of earnings (denote their model as BCN model). Based on BCN model, there were further extensive studies (see e.g. [9]).

III. METHODOLIOGES AND MODELS

As the disclosure of firms' financial information mainly comes from their annual report, the frequency of data in practice is relatively low (even for firms disclosing their information quarterly, the frequency is still not high enough). Therefore, the length of sample data is usually short. This limitation of dataset has restricted the application of many advanced and complicated time-series methods that rely heavily on sufficiency of data. In economics, people often deal with panel data that is to put time series observations of many individuals together, where certain assumptions on the distribution of parameters apply. In this way, particular model parameters could still be estimated on a pool of individuals with short data samples. The DKW model is in the form of:

$$CF_{i,t+k} = \gamma_{i,0} + \gamma_{i,1}CF_{i,t} + \gamma_{i,2}EARN_{i,t} + \varepsilon_{i,t} \quad (1)$$

where CF denotes net operating cash flow and $EARN$ denotes earnings. BCN model suggests that these accrual terms could have different effect in the model and therefore it extends (1) as:

$$CF_{t+1} = \beta_0 + \beta_1 CF_t + \beta_2 \Delta INV_t + \beta_3 \Delta AP_t + \beta_4 \Delta AR_t$$
$$+ \beta_5 DEP_t + \beta_6 AMORT_t + \beta_7 OTHER_t + \varepsilon_{t+1} \quad (2)$$

where ΔINV denotes changes in inventory, ΔAP denotes changes in account payable, ΔAR denotes changes in account receivable, DEP denotes depreciation, $AMORT$ denotes amortisation and $OTHER$ denotes other accruals. Clearly (2) has more details than (1) and thus requires more data points to estimate the many parameters. In the two papers, DKW and BCN models were indeed estimated in different ways. As DKW model has 3 parameters to be estimated for each firm, it is possible to undertake estimation procedure individually for sample firms. BCN model requires estimation of 8 parameters; therefore the model was estimated using pooled regression in the original paper, which restricted the parameters to be identical across different firms.

Despite the different ways in the estimation of parameters, the parameters in (1) and (2) are static. However, from DKW model's assumption, the cash flow process could be re-derived,

which would take a dynamic form instead of (1) or (2). The derivation is shown as in the following.

Net operating cash flow is the difference of cash received and cash paid out, which is represented as in DKW's paper:

$$CF_t = (SALES_t - \Delta AR_t) - (PURCHASE_t - \Delta AP_t)$$
$$= (SALES_t - \Delta AR_t) - (COST_t + \Delta INV_t - \Delta AP_t)$$
$$= EARN_t - \Delta WC_t \quad (3)$$

where the definitions of each term are as follow:

CF : Net operating cash flow

$SALES$: Sales

ΔAR : Changes in account receivable

$PURCHASE$: Purchase that is equal to the sum of cost and changes in inventory

ΔAP : Changes in account payable

$COST$: Cost

ΔINV : Changes in inventory

$EARN$: Earnings that is the difference of sales and cost

ΔWC : Changes in working capital that is equal to $\Delta AR_t + \Delta INV_t - \Delta AP_t$

Assume cost, account receivable, account payable and inventory to be constantly proportional to sales, therefore earnings and working capitals are also constant proportion of sales. Assign two constant α and β so that:

$$EARN_t = \alpha SALES_t \quad (4)$$

$$WC_t = \beta SALES_t \quad (5)$$

Define r to be the growth rate of sales and the following relation holds:

$$SALES_t = (1 + r_t)SALES_{t-1} \quad (6)$$

With some manipulations, there will be recursive relationships for earnings and working capital as:

$$EARN_t = \alpha SALES_t$$
$$= \alpha(1 + r_t)SALES_{t-1}$$
$$= (1 + r_t)EARN_{t-1} \quad (7)$$

$$\Delta WC_t = \beta \Delta SALES_t$$
$$= \beta r_t SALES_{t-1}$$
$$= \beta r_t (1 + r_{t-1})SALES_{t-2}$$
$$= (\frac{r_t}{r_{t-1}} + r_t)\Delta WC_{t-1} \quad (8)$$

Rewrite (3) as:

$$CF_t = EARN_t - \Delta WC_t$$
$$= (1+r_t)EARN_{t-1} - (\frac{r_t}{r_{t-1}} + r_t)\Delta WC_{t-1}$$
$$= (1+r_t)EARN_{t-1} - (1+r_t)\Delta WC_{t-1} + (\frac{r_{t-1}-r_t}{r_{t-1}})\Delta WC_{t-1}$$
$$= (1+r_t)CF_{t-1} + (\frac{r_{t-1}-r_t}{r_{t-1}})\Delta WC_{t-1}$$

(9)

Combining (8) and (9), cash flow and changes in working capital will be described as:

$$E_{t-1}\left(\begin{bmatrix} CF_t \\ \Delta WC_t \end{bmatrix}\right) = E_{t-1}\left(\begin{bmatrix} 1+r_t & \dfrac{r_{t-1}-r_t}{r_{t-1}} \\ 0 & \dfrac{r_t}{r_{t-1}}(1+r_{t-1}) \end{bmatrix} \times \begin{bmatrix} CF_{t-1} \\ \Delta WC_{t-1} \end{bmatrix}\right)$$

(10)

Therefore, there are potentially nonlinearity and dynamics in the unobservable parameters. To take account of such effects, this paper will propose the application of Grey-box model in enhancing the simple linear model form.

Grey-box was developed in [5], where 'grey' means the combination of black box and clear box. Take (2) for instance, the model is in a clear linear form, and differences in estimation method will only alter the value of parameters but do not change the model to be black box. Model (10) is also a clear box model because it is a theoretical model based on ideal assumptions. Empirically, the parameters of cash flow model (1) and (2) will not exactly comply with the form described in (10). Their evolvement, instead, is not clearly explainable due to complex environment and inevitable omissions of model assumptions. Therefore, processes in social sciences cannot be precisely captured by perfect mathematical models because there are interactions between infinitely many variables and extremely high uncertainty. For such processes that cannot be captured by clear box, black box will be a powerful alternative option. Assume parameters in (2) are nonlinear and time varying:

$$\beta_{i,t} = F(\mathbf{z}_t)$$

(11)

where $F(\mathbf{z}_t)$ is a nonlinear function of some variables $\mathbf{z}$. The form of function F is unknown; therefore it would be numerically approximated with a black box model. There are several options for such functions. For instance, a neural network (NN) is considered as a universal approximation that is able to approximate any function [10]. Similarly, Taylor series and Fourier series are two more examples that can approximate functions with any degree of accuracy. This paper adopts Padé approximant [5] for the nonlinear function as the method is efficiently accurate with only a few coefficients to determine.

In this paper, two forms of grey-box model (GM) will be simultaneously examined.

GM1:

$$CF_{i,t+1} = \beta_{i,t,0} + \beta_{i,t,1}CF_{i,t} + \beta_{i,t,2}\Delta INV_{i,t} + \beta_{i,t,3}\Delta AP_{i,t} + \beta_{i,t,4}\Delta AR_{i,t}$$
$$+ \beta_{i,t,5}DA_{i,t} + \beta_{i,t,6}OTHER_{i,t} + \varepsilon_{i,t+1}$$

(12)

$$\beta_{i,t} \approx \frac{c_0 + c_1 r_t + c_2 r_t^2}{1 + d_1 r_t + d_2 r_t^2}$$

GM2:

$$CF_{i,t+1} = \beta_{i,t,0} + \beta_{i,t,1}CF_{i,t} + \beta_{i,t,2}\Delta INV_{i,t} + \beta_{i,t,3}\Delta AP_{i,t} + \beta_{i,t,4}\Delta AR_{i,t}$$
$$+ \beta_{i,t,5}DA_{i,t} + \beta_{i,t,6}OTHER_{i,t} + \varepsilon_{i,t+1}$$

(13)

$$\beta_{i,t} \approx \frac{c_0 + c_1 AGE_{i,t} + c_2 AGE_{i,t}^2}{1 + d_1 AGE_{i,t} + d_2 AGE_{i,t}^2}$$

where DA is depreciation and amortisation (which are aggregated for convenience in data processing); AGE denotes firm age, which is calculated as the difference of time t and the initial time t_0 when the first observation appears in the sample; c s and d s are the coefficients of Padé approximants to be determined. GM2 uses age as a proxy variable for firms' growth path based on empirically observed pattern that firms' growth is related to firms' ages.

In comparison, model (2) with static parameters will be estimated as benchmark (BM). This paper use two estimation methods of model (2) to allow differences in individual intercept terms:

BM:

$$CF_{i,t+1} = \beta_{i,0} + \beta_1 CF_{i,t} + \beta_2 \Delta INV_{i,t} + \beta_3 \Delta AP_{i,t} + \beta_4 \Delta AR_{i,t}$$
$$+ \beta_5 DA_{i,t} + \beta_6 OTHER_{i,t} + \varepsilon_{i,t+1}$$

(14)

The model is named fixed effect (or random effect if different assumption is made) model in micro-econometrics. Allowing individual effect as represented by the intercept terms would increase the models practical predictive power. The parameters of β_1 to β_6 are still assumed to be homogeneous among firms. They will be estimated by two methods: demean and first difference (for more details please see the textbooks [11]), and the intercept term is calculated as:

$$\hat{\beta}_{i,0} = \overline{CF}_{i,t}$$
$$- (\beta_1 \overline{CF}_{i,t-1} + \beta_2 \overline{\Delta INV}_{i,t-1} + \beta_3 \overline{\Delta AP}_{i,t-1}$$
$$+ \beta_4 \overline{\Delta AR}_{i,t-1} + \beta_5 \overline{DA}_{i,t-1} + \beta_6 \overline{OTHER}_{i,t-1})$$

(15)

IV.   RESEARCH DESIGN AND DATA

The models (GM1, GM2, BM demean and first difference) will be examined in the performance of practical prediction for cash flow. The comparison procedure will be in three stages: in-sample fitness, out-of-sample one period prediction and out-of-sample multiple period prediction. The out-of-sample test is adopted to take account of the fact that complicated models

tend to over-fit data. Over-fitting models would have very good performance in-sample but poorer performance out-of-sample, therefore models that are not tested by out-of-sample data might cause harm when put into practical use.

The dataset used for empirical analysis is the U.S.A annual data collected from the WRDS Compustat dataset. The period of the sample spans from 1988 to 2012. All available firms in the U.S. will be included in the sample as long as the required variables in the models are available except financial service firms (SIC codes from 6000 to 6999). Outliers' exclusion procedure follows that in Barth et al. (2001), leaving the sample with 99845 firm-year observations. All variables are deflated by average total assets. Data before year 2005(inclusive) is used to estimate model parameters and data thereafter are used for out-of-sample comparison.

The basic criterion for the model comparison is sum squared errors (SSE). With panel data that include different individuals, SSE that is a general measure aggregating over the whole sample may not be sufficient to clearly indicate different models' performance. Therefore, average rank is also used to compare the performance of models as a second test. For each observation, the model generating the minimum magnitude of prediction error will be deemed as the best and hence ranked 1st and the opposite for the worst model that produces the largest error. The ranks of each model will be averaged over all observations, indicating the average performance of the models in making individual prediction. The model with the smallest average rank would be considered to be better performing in general.

In out-of-sample prediction, all 4 models could be applied in one-period-ahead prediction. The multiple-period prediction will be implemented in the spirit of vector autoregressive (VAR) model (see textbook [12]) that is to make predictions for all predictive variables, and then use the predicted variables recursively to predict into further periods in the future. In multiple-period prediction, BM parameters are static and thus there is no difference to one-period prediction. However, the GM have dynamic parameters, which will also be recursively determined over time. GM1 relies significantly on prediction of future sales growth rates, where recursion hence cannot be naturally extended from one-period to multiple-period ahead. Therefore, in multiple-period prediction comparison, GM1 is excluded.

V.   EXPERIMENTAL RESULTS

In the dataset, all observations that are before 2005 inclusive are pooled together for estimating the factor loadings in (14). Then the intercept term is calculated individually using (15). For the implementation of (15), there is naturally a requirement that the individual firm need to have at least 2 observations in the period to calculate mean values for the variables. For firms with only one observation, the intercept term could be still calculated, which will be equivalent as the prediction error. In the in-sample comparison, firms with one observation are excluded, but which are included back for out-of-sample prediction. Therefore, in the in-sample stage, there are 39631 firm-year observations for comparison.

The parameters of β_1 to β_6 in (14) estimated by demean and first difference methods deviate significantly from each other as Table 1 shows, especially for the autoregressive (AR) parameter β_1. The AR coefficient estimated using the first difference method is negative, which is substantially due to the estimation procedure, which incurs negative autocorrelation bias. As a result, the intercept terms calculated based on the two methods are not close either. It is expected that the demean methods would give more reliable results than the first difference method.

It is noteworthy that in comparison with the BM model (14), the intercept term involved in GM1 and GM2 is not assumed to be constant within an observed individual. GM models hence are more flexible than the BM model as they do not place requirement on the number of observations of individual firms to calculate the intercept terms.

The Padé approximants coefficients in GM1 and GM2 are estimated by minimizing the sum squared error of the sample observations. After the coefficients are determined, the relationship of sales growth rates and firm ages with the cash flow model parameters can be calculated. Fig. 1 and Fig. 2 plot the relationship of sales growth rates and firm ages with β_1 respectively for demonstration.

In both figures, the nonlinearity is apparent, which suggest that the two selected variables, i.e. sales growth rates and firm ages tend to have some explanatory power to the model parameters, otherwise the figures would show flat straight lines. The curve in Fig.1 is generally monotonic. It suggests that higher growth in sales indicates higher AR parameter. Fig. 2 in general indicates that as time goes by, the AR parameter tends to increase except for the first two years.

Table 1

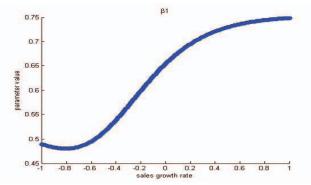| Estimation Methods	Estimated Parameters in the Linear Model					
	β_1	β_2	β_3	β_4	β_5	β_6
Difference	-0.2013	0.1709	-0.3062	0.3081	0.4493	0.2337
Demean	0.5054	0.3194	-0.4872	0.4757	0.7514	0.3657

Figure 1: The relationship of AR parameter with sale growth rates.

The in-sample fitness of the four models are compared in Table 2. The first two rows denote the models in comparison. The third row contains the mean squared errors (MSE) generated by each model. BM model estimated by demean method fits the data better because of the consideration of individual effects. However BM model estimated by difference method has the poorest fit with the data, suggesting that the parameters estimated in this way contains greater bias for practical prediction and poorer fit. The GM1 and GM2 have similar data fitting ability, but poorer than the BM demean model because they do not calculate intercept terms for individual firms. The numbers in the last row in Table 2 are the average ranks of the four models. The conclusion drawn from this criterion is not different from the MSE: the BM demean method have the lowest average rank so it has the best in-sample data fitness of all these four models. BM difference method is again the worst in this measure and the two GM models are similar to each other, both lying between the two BM models.

The parameters determined in-sample are used in out-of-sample test without re-estimation. In the one-period-ahead sample, the firms that do not appear in the estimation sample are not predictable using BM models because there is no way to decide their intercept terms. However the GM models could adapt well to this situation. The one-period-ahead results are shown in Table 3. The measures are calculated based on 17965 firm-year observations. Both MSE and average rank measures show a significant outperformance of grey-box models over the benchmark models in one-period-ahead prediction. Despite the better performance of demean model over the grey-box models in the in-sample data fitting capacity, the latter have shown stronger out-of-sample performance by the two criteria. GM2
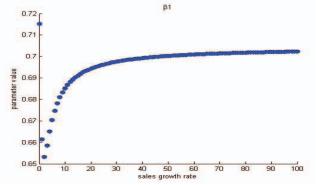

Figure 2: The relationship of AR parameter with firm ages.

Table 2

Measure of Performance	In-Sample Performance of the Models			
	GM1	GM2	BM	BMª
MSE	0.0075	0.0075	0.0053	0.0078
Average Rank	2.5247	2.5215	2.3760	2.5778

a.   Difference method; the other BM is default to be by demean method.

that uses firm age as input to black box has the lowest MSE among the four models and GM1 has the second lowest MSE. The difference in the MSE by the two grey-box models is negligible. The benchmark models have higher MSE, especially that estimated by the first difference method. The comparison of average rank in the last row also favours the grey-box models. Both GM have lower average rank than the two BM, which suggest that the grey-box models do not only provide more accurate predictions for the sample in general, but also have better performance in predicting on individual level. It is noteworthy that GM1 and GM2 does not dominate each other as GM1 has the lowest average rank while GM2 has the lowest MSE.

In the multiple-period-ahead prediction test, all forecast are made based on the observations in year 2005. Result of year 2006 is excluded because it is not multi-period prediction. Therefore, this test could examine the predictive power of the GM2 and BM models for at most 7 years. The results are summarised in Table 4. The results are impressive and encouraging for the application of GM2, i.e. grey-box model with firm age. It has outperformed the two benchmark model again in both criteria. The MSE of GM2 is smaller than 0.02 whereas that of the benchmark models are much higher, especially the first difference model. GM2 also has the lowest average rank of the three models and thus it dominantly outperforms the benchmark models. The demean method has resulted in better practical performance than the first difference method, according to the two criteria.

VI.   CONCLUSION

This paper has introduced the grey-box modelling technique to financial forecasting. This method has been shown to be powerful in engineering and its application to financial industry is therefore investigated, such as for firm-level cash flow prediction. Cash flow is not easy to predict. There are theories suggesting that it could be predicted with earnings components. In previous studies, simple linear models are developed for this application. However, it has been shown that the process of cash flow could be more complicated than linear process. Nonlinearity and dynamics that have been overlooked

Table 3

Measure of Performance	One-period Out-of-Sample Prediction Performance			
	GM1	GM2	BM	BMª
MSE	0.0111	0.0109	0.0140	0.0376
Average Rank	2.1283	2.2367	2.7716	2.8634

a.   Difference method; the other BM is default to be by demean method.

Table 4

Measure of Performance	Multiple-period Out-of-Sample Prediction Performance		
	GM2	BM	BM[a]
MSE	0.0199	0.0244	0.0278
Average Rank	1.6530	1.9273	2.4197

a. Difference method; the other BM is default to be by demean

are the main issue addressed in this paper.

The grey-box model incorporates a black-box model to fit nonlinear data into a clear box model which explains the theoretical mechanism of the target variables. There are two grey-box models implemented in this paper in comparison with the benchmark linear models. With the assistance of Padé approximants, the grey-box model has captured some nonlinearity in the modelling system. The modelling performance is examined for out-of-sample prediction. The results have shown great improvement of grey-box model in making practical predictions of cash flows. The grey-box models outperform the benchmark models in making both one-period-ahead and multi-period-ahead predictions. The conclusion is consistent between the two criteria adopted in this paper. Therefore, the results encourage the application of grey-box modelling in the business world where nonlinear interactions are inevitable. Grey-box model could incorporate more variables that are considered to contain useful information without changing the original clear-box model structure. Therefore it could add extra power to simple linear models.

REFERENCES

[1] Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk*. *The Journal of Finance*, *19*(3), 425-442.

[2] Campbell, J. Y., & Shiller, R. J. (1988). Stock prices, earnings, and expected dividends. *The Journal of Finance*, *43*(3), 661-676.

[3] Kalman, Rudolph Emil. "A new approach to linear filtering and prediction problems." *Journal of basic Engineering* 82, no. 1 (1960): 35-45.Sdf

[4] Trippi, R. R., & Turban, E. (1992). Neural Networks in Finance and Investing: Using Artificial Intelligence to Improve Real World Performance. McGraw-Hill, Inc..

[5] Tan, K. C., & Li, Y. (2002). Grey-box model identification via evolutionary computing. *Control Engineering Practice*, *10*(7), 673-684.Asdf

[6] Hopwood, W. S., & McKeown, J. C. (1992). Empirical evidence on the time-series properties of operating cash flows. *Managerial Finance*, *18*(5), 62-78.

[7] Dechow, P. M., Kothari, S. P., & L Watts, R. (1998). The relation between earnings and cash flows. *Journal of Accounting and Economics*, *25*(2), 133-168.

[8] Barth, M. E., Cram, D. P., & Nelson, K. K. (2001). Accruals and the prediction of future cash flows. *The Accounting Review*, *76*(1), 27-58.

[9] Cheng, C. A., & Hollie, D. (2008). Do core and non-core cash flows from operations persist differentially in predicting future cash flows?. *Review of Quantitative Finance and Accounting*, *31*(1), 29-53.

[10] Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of control, signals and systems*, *2*(4), 303-314.

[11] Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics: methods and applications*. Cambridge university press.

[12] Greene, W. H. (2008). *Econometric analysis*. Granite Hill Publishers.