Expert Systems with Applications 39 (2015) 9564–9570

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Detecting earnings management with neural networks

Henrik Höglund [⇑]

Hanken School of Economics, Handelsesplanaden 2, 65101 Vasa, Finland

article info

Keywords:
Earnings management
Discretionary accruals
Neural networks

abstract

A large body of studies has examined the occurrence of earnings management in various contexts. In most studies, the assumption has been that earnings are managed through accounting accruals. Thus, a range of accrual based earnings management detection models have been suggested. The ability of these models to detect earnings management has, however, been questioned in a number of studies. An explanation to the poor performance of the existing models is that most models use a linear approach for modeling the accrual process even though the accrual process has in fact proven non-linear in several studies. An alternative way to deal with the non-linearity is to use various types of neural networks. The purpose of this study is to assess whether neural network-based models outperform linear and piecewise linear-based models in detecting earnings management. The study comprises neural network models based on a self-organizing map (SOM), a multilayer perceptron (MLP) and a general regression neural network (GRNN). The results show that the GRNN-based model performs best, whereas the linear regression-based model has the poorest performance. However, the results also show that all five models assessed in this study estimate discretionary accruals, a proxy for earnings management, with some bias.

1. Introduction

A variety of studies has examined the occurrence of earnings management in several contexts, such as prior to initial public offerings (Teoh, Welch, & Wong, 1998), during financial distress (DeFond & Jiambalvo, 1994; Jaggi & Lee, 2002) and during changes in accounting standards (Van Tendeloo & Vanstraelen, 2005). In most studies, the assumption has been that earnings are managed through accounting accruals. Based on this assumption, various models that divide the accruals into non-discretionary accruals (expected accruals) and discretionary accruals (unexpected accruals) have been suggested. The discretionary accruals estimated with these models are considered a proxy for earnings management. The problem in using discretionary accrual estimation models is that earnings management is not directly measurable, not even ex-post. That is, earnings management activities are often difficult to distinguish from normal business activities. Thus, the assessment of the actual performance of these models can be problematic. The ability of the discretionary accrual estimation models to extract the discretionary part of accruals has been questioned in a number of studies. Thomas and Zhang (2000), for example, showed that most of the discretionary accrual estimation models perform worse than just the simple assumption that the non-discretionary accruals equal 5% of total assets. One explanation to the poor performance of the models is that the data usually is rather noisy. Another possible explanation is that most models use a linear approach for modeling the accrual process, which might impair the performance of the models as the results from several studies suggest that the accrual process in fact is non-linear (e.g. Dechow, Sloan, & Sweeney, 1995; Jeter & Shivakumar, 1999; Kothari, Leone, & Wasley, 2005). Despite of the obvious need for a non-linear approach for estimating discretionary accruals, there have so far been only a few suggestions for such models. Attempts to deal with the non-linearity have, for example, been done with performance matching (Kothari et al., 2005) and piecewise linear regression (Ball & Shivakumar, 2006; Jeter & Shivakumar, 1999). An alternative way to deal with the non-linearity is to use various types of neural networks. Although successfully used in a number of financial studies (see Vellido, Lisboa & Vaughan, 1999; Paliwal & Kumar, 2009 for reviews), neural networks have yet not been used for estimating discretionary accruals. When applied to regression analysis problems, neural networks have several appealing advantages compared to traditional statistical methods, such as multiple linear regression. First, neural networks can be used for modeling non-linear relationships. Second, a number of assumptions, which are necessary when using traditional statistical methods, can be disregarded when using neural networks. Third, neural networks are not as sensitive to outliers and missing data as traditional statistical models.

The purpose of this study is to assess whether neural network-based models outperform linear and piecewise linear-based models in detecting earnings management. The study comprises one model based on linear regression, one model based on piecewise linear regression and three models based on different types of neural networks.

⇑ Tel.: +358 403521.

E-mail address: henrik.hoglund@hanken.fi

0957-4174/$ - see front matter 2015 Published by Elsevier Ltd.
doi:10.1016/j.eswa.2015.02.096

The remainder of the study is organized as follows: different approaches for estimating discretionary accruals are presented in Section 2. Furthermore, previous studies where linear regression has been compared with neural networks are discussed in Section 2. The research design is presented in Section 3 and the results from the empirical study are presented in Section 4. Section 5 concludes the study.

2. Detecting earnings management

2.1. Definition of earnings management

A central issue when discussing earnings management is the trade-off between the relevance and reliability of the reported accounting data. Focusing entirely on reliability, the management would report only realized cash flows, whereas a focus on relevance would emphasize the current value of expected future cash flows (Dechow, 1994). However, even though relevance and reliability in reporting are not completely mutually exclusive, the reliability of the reported accounting data usually suffers as the relevance increases. The problem is that relevant reporting requires various estimations made by the company management, which in turn involve at least some degree of subjectivity. As the various estimates are difficult to verify, they become vulnerable to managerial manipulation (Sloan, 1999). Despite the problems with relevant reporting, a certain degree of both relevance and reliability is required by the reported accounting data. Thus, generally accepted accounting principles provide some degree of discretion in reporting. In general, earnings management can be viewed as the utilization of the discretion in reporting provided by generally accepted accounting principles.

2.2. Total accrual measure

The first step in using discretionary accrual estimation models is to determine the total accruals. In the literature, mainly two methods are used for calculating total accruals: the balance sheet approach and the statement of cash flows approach. With the more widely used balance sheet approach, total accruals (TACC) are calculated by subtracting total depreciation from the change in non-cash working capital:

$$TACC_t = \Delta CA - \Delta CL - \Delta Cash + \Delta STDEBT - DEP$$

where ΔCA = the change in current assets; ΔCL = the change in current liabilities; $\Delta Cash$ = the change in cash and cash equivalents; $\Delta STDEBT$ = the current maturities of long-term debt and other short-term debt included in current liabilities; DEP = depreciation and amortization expenses. In the second method for calculating total accruals, the statement of cash flows approach, cash flows from operations (CFO) is subtracted from earnings before extraordinary items (EXBI):

$$TACC_t = EXBI - CFO$$

Of these two methods, the statement of cash flows approach is to be preferred as previous studies have shown that mergers, acquisitions and divestitures and foreign currency translations might lead to a biased measure of total accruals (Collins & Hribar, 2002).

2.3. Discretionary accrual estimation models

The purpose of the discretionary accrual estimation models is to divide accruals in to non-discretionary and discretionary accruals. The non-discretionary part of the accruals is the part that the company management has no or little control over. The discretionary part of the accruals, on the other hand, is under the control of the company management. Thus, the discretionary accruals estimated with these models are considered a proxy for earnings management.

The first discretionary accrual estimation model was presented by Healy (1985). The logic behind Healy's model was that the non-discretionary accruals for the event year equal the long-run average of the total accruals. Another early aggregate accruals model was presented by DeAngelo (1986). DeAngelo's model is based on the assumption that the non-discretionary accruals equal the lagged total accruals. However, the most commonly used discretionary accrual estimation model is the model that was suggested by Jones (1991). The Jones model is based on a linear regression where change in sales (DREV) and property, plant and equipment (PPE) are regressed on total accruals (TACC):

$$\frac{TACC_t}{TA_{t-1}} = b_0\frac{1}{TA_{t-1}} + b_1\frac{\Delta REV_t}{TA_{t-1}} + b_2\frac{PPE_t}{TA_{t-1}}$$

In the regression model, the change in sales controls for the current accruals (working-capital accruals) whereas property, plant and equipment controls for the non-current accruals (depreciation and amortization). In the original Jones model the regression coefficients were estimated using company specific data from 1 year prior to the event year to at least 10 years prior to the event year, thus making the coefficients company specific. In more recent studies this approach has been more or less abandoned in favor of a cross-sectional approach (e.g. DeFond & Jiambalvo, 1994). With the cross-sectional approach the regression coefficients are estimated by using data from companies within the same industry and year as the companies for which earnings management is studied. Thus, the estimated regression coefficients are industry and year specific instead of company specific. Once the regression coefficients have been estimated, they are used for calculating the non-discretionary accruals for a specific company. Finally, the discretionary accruals are calculated by subtracting the non-discretionary accruals from the actual total accruals.

Several studies have indicated that the Jones model is a biased measure of earnings management. For example, Dechow et al. (1995) estimated discretionary accruals for companies with either extreme earnings performance or extreme cash flow from operations. The results showed that the companies with high earnings had high discretionary accruals whereas the companies with low earnings had low discretionary accruals. Similarly, the companies with high cash flow from operations had low discretionary accruals whereas the companies with low cash flows from operations had high discretionary accruals. These findings follow the fact that the companies with high (low) earnings tend to have high (low) accruals whereas the companies with high (low) cash flow from operations tend to have low (high) accruals. In other words, these results suggest that the earnings management detection models attribute parts of the non-discretionary accruals for companies with extreme financial performance as discretionary accruals. Similar results have also been reported by Kothari et al. (2005). This kind of bias can in the worst case lead to wrong conclusions being drawn when studying earnings management among companies with extreme financial performance. Thus, it seems warranted to augment the Jones model with explanatory variables which controls for performance.

A widely used measure of performance used together with the Jones model is cash flows from operations (CFO) (e.g. Kaznik, 1999; Rees, Gil, & Gore, 1996). An implicit assumption with the Rees et al. and Kasznik variants of the Jones model is that the relationship between the accruals and CFO is linear. Considering only the noise-reduction role of accruals, this assumption might hold. However, accruals do also have an asymmetrically timely loss recognition role (Ball & Shivakumar, 2006) which challenges the linear relationship between the accruals and CFO. This second role is in line with the accounting conservatism principle according to which unrealized losses are more likely to be recognized earlier than unrealized gains (Basu, 1997). According to Ball and Shivakumar (2006) the accounting conservatism implies that accrual models that are linear in cash flows are miss-specified. Instead, they argued that the correct specification most likely is piecewise linear. Thus, they proposed a piecewise linear regression approach when estimating the discretionary accruals with the Jones model:

$$\frac{TACC_t}{TA_{t-1}} = b_0\frac{1}{TA_{t-1}} + b_1\frac{DREV_t}{TA_{t-1}} + b_2\frac{PPE_t}{TA_{t-1}} + b_3\frac{CFO_t}{TA_{t-1}} + b_4 DCFO + b_5 DCFO\,\frac{CFO_{t-1}}{TA_{t-1}}$$

Ball and Shivakumar augmented the Jones model with cash flows from operations (CFO), a dummy variable indicating losses or gains (DCFO) and an interaction term of these two variables. The results from Ball and Shivakumar's study showed a noticeable increase in the regression R^2 value when applying the piecewise linear approach.

2.4. Linear regression versus neural networks

When applied to regression analysis problems, neural networks have several appealing advantages compared to traditional statistical methods, such as linear regression. DeTienne, DeTienne, and Joshi (2003) discussed the main advantages and disadvantages with neural networks when compared to linear regression. First, a considerable disadvantage with linear regression is that it cannot deal with non-linear relationships between the dependent variable and the independent variables. Neural networks, on the other hand, can effectively model non-linear relationships. Second, the performance of linear regression models hinges largely on various assumptions such as absence of multicollinearity, and normally distributed residuals with zero mean and constant variance. In contrast, these assumptions are not required with neural networks. Third, when using linear regression the underlying model must be specified in advance. With linear regression, for example, the model assumption is that the dependent variable is related to a linear combination of the independent variables (Warner & Misra, 1996). This is not required with neural networks as they are entirely data driven. Fourth, linear regression is relatively sensitive to missing and noisy data as well as outliers in the data. Neural networks, on the other hand, are less sensitive to these kinds of features in the data. Thus, neural networks are to be preferred over traditional parametric models when the data do not satisfy the assumptions required by the parametric models or when it is evident that the data contain large outliers (Coakley & Brown, 2000).

Even though neural networks provide several advantages over traditional statistical methods, previous research has shown that the neural networks are not superior to linear regression in all situations. Most studies show that neural networks perform worse or equally well compared with linear regression when the underlying data is correctly modeled (e.g. Denton, 1995; Marquez, Hill, Worthley & Remus, 1992; Warner & Misra, 1996). If, on the other hand, the underlying data has not been correctly modeled, the neural network in most cases outperforms the linear regression approach. Thus, it is preferable to use neural networks instead of linear regression if the underlying function of the data is unknown or if it is impossible to model it correctly. In addition to the issues regarding the underlying models, there are also some practical issues to consider. First, with neural networks there are no clear rules for how to define the architecture and the different parameters of the network. The most common way to determine the optimal architecture of the neural network is simply by trial and error (Zhang, Patuwo, & Hu, 1998). Second, with neural networks the training of the network is usually a lot more time-consuming than the estimation of the regression coefficients. For example, Marquez et al. (1992) report that the time required for estimating the regression coefficients with a sample containing 100 observations was only a few seconds whereas the neural network models required tens of minutes for training. Furthermore, the training time of neural networks usually grows exponentially as the number of nodes in the model increase (Coakley & Brown, 2000). Third, neural network-based models cannot be interpreted in the same way as linear regression based-models. More closely, the neural network weights cannot be interpreted in the same way as regression coefficients (DeTienne et al., 2003). It is also not possible to test the significances of the neural network inputs (Calderon & Cheh, 2002).

3. Research methodology

3.1. Research task

The purpose of this study is to assess whether neural network-based models outperform linear and piecewise linear-based models in detecting earnings management. The study comprises one model based on linear regression, one model based on piecewise linear regression and three models based on different types of neural networks. The performance of the five different models is assessed both using a random data set as well as a stratified random data set. The main measure of performance is how close to zero a given model estimates the mean and median discretionary accruals for the data sets.

3.2. Models based on linear regression

This study comprises two different models based on traditional statistical approaches. Both of these models are augmentations of the linear regression-based model suggested by Jones (1991). In the first of the two models the reciprocal of lagged total assets, the change in sales (DREV), gross property, plant and equipment (PPE) and cash flows from operations (CFO) are regressed on total accruals (TACC). In this study, the total accruals are calculated using the statement of cash flows approach. All variables in the regression equation are deflated by lagged total assets (TA_{t-1}). This variant of the Jones model was first suggested by Rees et al. (1996).

$$\frac{TACC_t}{TA_{t-1}} = b_0\frac{1}{TA_{t-1}} + b_1\frac{DREV_t}{TA_{t-1}} + b_2\frac{PPE_t}{TA_{t-1}} + b_3\frac{CFO_t}{TA_{t-1}}$$

Change in sales is assumed to explain current accruals resulting in the expected sign for the b_1 coefficient being positive. Property, plant and equipment are assumed to control for non-current accruals. As the non-current accruals mainly consist of depreciations, the expected sign for the b_2 coefficient is negative. Several prior studies (e.g. Dechow, 1994; Sloan, 1996) have shown a strong negative correlation between cash flows from operations and total accruals. Thus, the expected sign for the b_3 coefficient is also negative.

In addition to their noise reduction role the accruals also have an asymmetrically timely loss recognition role. This calls in question a linear relationship between cash flows from operations and total accruals. The second model takes this into consideration by estimating different slopes for positive and negative cash flows from operations. This piecewise linear regression approach was first suggested by Ball and Shivakumar (2006).

$$\frac{TACC_t}{TA_{t-1}} = b_0\frac{1}{TA_{t-1}} + b_1\frac{DREV_t}{TA_{t-1}} + b_2\frac{PPE_t}{TA_{t-1}} + b_3\frac{CFO_t}{TA_{t-1}} + b_4 DCFO + b_5 DCFO\,\frac{CFO_{t-1}}{TA_{t-1}}$$

The additional variables in the second model are a dummy variable (DCFO) assigned a value 1 if the cash flows from operations are negative and an interaction variable between the dummy variable and cash flows from operations. If the asymmetrically timely loss recognition role of accruals is dominant, the expected sign for the b_5 coefficient is positive.

The regression coefficients for both models are estimated using cross-sectional data. That is, the financial statement data used for estimation is from the same industry and year as the financial statement data for the companies for which earnings management is examined. The estimated regression coefficients are then used on financial statement data for the companies for which earnings management is examined in order to determine an expected level of accruals (non-discretionary accruals). Finally, the discretionary accruals, a proxy for earnings management, are calculated by subtracting the non-discretionary accruals from the actual total accruals.

3.3. Models based on neural networks

The performance of the two models based on traditional statistical approaches is compared with the performance of three different types of neural network-based models. All three models use total accruals as the dependent variable and the reciprocal of lagged total assets, change in sales, gross property, plant and equipment and cash flows from operations as independent variables.

The first of the three neural network models is based on a self-organizing map (SOM). The SOM uses an unsupervised learning algorithm. Therefore, it is not as such suitable for regression type of problems. In this study, the SOM is instead used for creating several local linear regression models (Vesanto, 1997; Whigham, 2005). When training the SOM, both the dependent variable and the independent variables are used for selecting the best matching units (BMU). Once the training is completed, a local linear regression model is created for each node. The number of observations required for the local models is determined by testing with different alternatives, ranging from 1% to 100% of the total estimation data set. The number of observations that yields the lowest mean squared error (MSE) for the test data set will be used in the final model. If a specific node does not have the required number of observations, the estimation data set for the local model is augmented with observations from the closest nodes. When calculating the non-discretionary accruals for an observation, the BMU for the data vector is first selected based on the independent variables. Once the BMU has been selected, the non-discretionary accruals are calculated by using the local linear regression model. Finally, the discretionary accruals are calculated by subtracting the non-discretionary accruals from the actual total accruals.

The second neural network model is based on a multilayer perceptron (MLP). Contrary to the SOM, the MLP is based on a supervised learning algorithm and is therefore better suited for regression problems. The basic structure of a MLP is one input layer, one or more hidden layers and one output layer. Each of these layers comprises one or several nodes. The structure of the input and output layers is rather straight forward. That is, the input layer has as many nodes as there are independent variables and correspondingly the output layer has as many nodes as there are dependent variables. The structure of the hidden layer, however, is more difficult to determine. In this study, one hidden layer is used and the number of nodes in this layer is determined by testing with values between 1 and 20. The transfer function for the nodes in the hidden layer is set to hyperbolic tangent and the transfer function for the output node is set to linear. The hidden layer structure with the lowest MSE for the test sample is used as the final model. The non-discretionary accruals are calculated by presenting the independent variables from a data vector to the MLP input layer. The discretionary accruals are then calculated by subtracting the non-discretionary accruals from the output node with the actual total accruals.

The third neural network model is based on a general regression neural network (GRNN). Similarly to the MLP, the GRNN is based on a supervised learning algorithm. The GRNN comprises an input layer, a hidden layer, a summation layer and an output layer. The input and output layers equals the independent variables and the dependent variable respectively. The hidden layer consists of as many nodes as there are trainings vectors whereas the summation layer consists of two nodes. The training process of the GRNN is performed in one sweep. When new data is presented to the GRNN, the distance between the input and the weight vectors is calculated. The distance is then passed through a radial basis function so that a shorter distance returns a larger output value. A central parameter for a GRNN is the spread. A low value for spread gives a steep radial basis function leading to only a few neurons contributing significantly to the output. A high value for spread, on the other hand, results in a smoother network function as several nodes contribute to the output. In this study, the optimal value for spread is determined by testing different values between 0 and 2 with an increment of 0.02. The spread with the lowest MSE for the test sample is used in the final model. The non-discretionary accruals are calculated by presenting the independent variables from a data vector to the GRNN input layer. The discretionary accruals are then calculated by subtracting the non-discretionary accruals from the output node with the actual total accruals.

3.4. Data set description

The data set comprises financial statement data for public US manufacturing companies (SIC 20xx–SIC 39xx) with complete data for years 2006 and 2007. The data used in this study are retrieved from Thomson One Banker. Companies with the absolute value of total accruals equal to or greater than the lagged total assets are removed from the data set. Furthermore, all five variables used in the models are trimmed at the 1st and 99th percentile. Based on these criteria, the total number of companies included in the data set is 2032.

The total data set is randomly separated into an estimation data set (75% of the total data set) and an evaluation data set (25% of the total data set) (see Fig. 1). The estimation data set is used for estimating the coefficients for the linear regression-based models and for training the neural network-based models. With the SOM- and GRNN-based models, one-third of the estimation data set is used as a test data set. The MLP-based model requires both a test data set and a validation data set, each comprising one sixth of the estimation data set. The validation data set is used for preventing the MLP from overtraining. The companies in both the test and the validation data set are randomly drawn from the estimation data set.

3.5. Strategy of analysis

A considerable problem when comparing different models for detecting earnings management is that the actual magnitude of

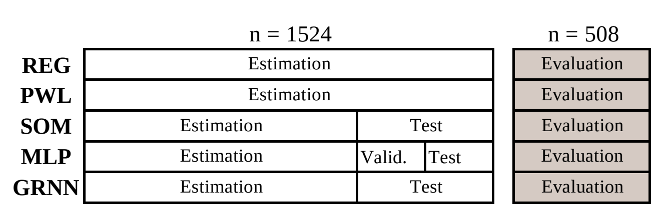

Fig. 1. Data set structure.

earnings management cannot be measured. Therefore, the performance of the models cannot be directly assessed. There are, however, alternative ways of determining the performance of the models. First, an assumption is that if the estimation and evaluation data sets are randomly drawn from the same population, the average discretionary accruals should not be significantly different from zero. Thus, the closer to zero a model estimates the discretionary accruals in the evaluation data set, the better the performance of the model. Second, the discretionary accruals estimated by the different models should not correlate with various variables measuring size and performance. In this study, the discretionary accruals are sorted based on five different variables (ROA, CFO, Size, P/B-value and P/E-value). The upper and lower quartiles for each variable and model are then examined. The closer to zero a model estimates the discretionary accruals for a specific quartile and variable, the better the performance of the model.

4. Results and discussion

4.1. Descriptive statistics

The total data set consists of 2032 public US manufacturing companies. The descriptive statistics in Table 1 show that there are considerable differences between the companies. The smallest company has total assets of only 200,000$, whereas the largest company has total assets of more than 55 billion dollars. Similarly, the sales range between zero sales and more than 95 billion dollars. The minimum and maximum values of the P/E and P/B ratios are fairly large as only the variables used in the models have been trimmed.

The regression coefficients for the linear regression (REG) and piecewise linear regression-based (PWL) models is presented in Table 2. With both models, the coefficients for change in sales (b_1) and property, plant and equipment (b_2) have the expected signs. The positive b_3 coefficient for the REG model implies that the timely loss recognition role of accruals is more dominant for the companies in the data set than the noise-reduction role. The coefficients for the PWL model show that for companies experiencing losses (negative cash flows from operations) the timely loss recognition role of accruals is more dominant, whereas for companies experiencing gains (positive cash flows from operations) the noise-reduction role is more dominant. Finally, the R^2 value for the PWL model is noticeably higher than the R^2 value for the REG model.

The estimation and test errors (mean-squared error, MSE) for the three different types of neural network-based models are presented in Figs. 2–4. The model performance is determined based on the MSE for the test sample. With the SOM model, the errors are

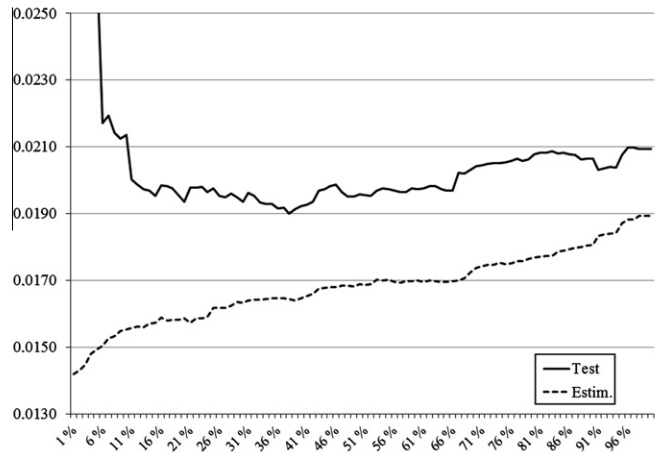

Fig. 2. Estimation and test errors for SOM-based model.

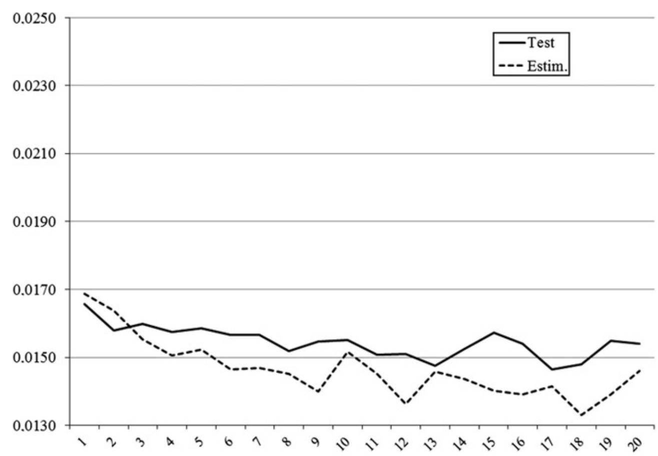

Fig. 3. Estimation and test errors for MLP-based model.

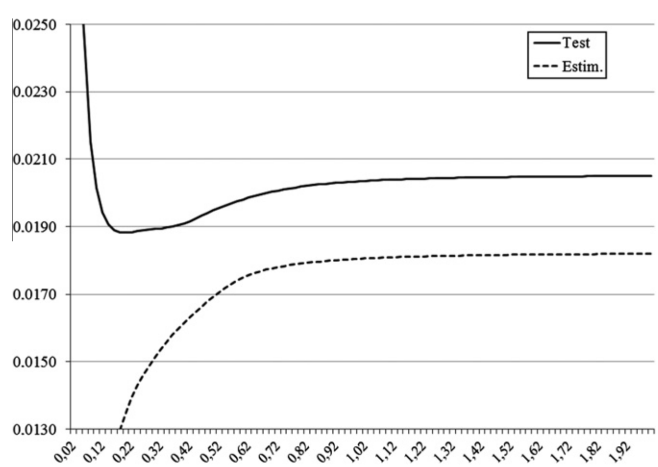

Fig. 4. Estimation and test errors for GRNN-based model.

Table 1
Descriptive statistics for data set companies.

n = 2032	Sales	Income	CFO	Total assets	P/E	P/B
Mean	1866.8	123.3	204.1	1917.6	5.1	9.0
Median	159.1	2.4	8.7	183.4	2.2	11.1
St. dev.	6135.6	554.4	756.8	5496.1	59.4	94.5
Min	0	3379.0	411.0	0.2	963.0	1743.0
Max	95,327.0	7333.0	12,625.0	55,370.0	1575.0	1185.4

All values (except for the P/E and P/B values) are given in million dollars. Income = income before extraordinary items and preferred dividens, CFO = net cash flows from operations, P/E = price-to-earnings ratio, P/B = price-to-book ratio.

Table 2
Regression coefficients.

	$1/TA_{t\ 1}$ b_0	$DREV_t$ b_1	PPE_t b_2	CFO_t b_3	$dCFO_t$ b_4	$CFO_t\ dCFO_t$ b_5	R^2
REG	0.219	0.011	0.087	0.079			0.196
	0	0.424	0	0			
PWL	0.134	0.062	0.006	0.458	0.049	0.576	0.301
	0	0	0.457	0	0	0	

measured with different data set sizes for the local linear regression models. The best performance is achieved when the data set size for the local linear regression models equal 38% of the total estimation data set size. With the MLP model with one hidden layer, different numbers of hidden layer nodes, ranging from 1 to 20, are assessed. The results show that the MSE for the test sample reaches its lowest value when the number of hidden layer nodes equals 17. Finally, the performance of the GRNN model is assessed with different values for the spread, ranging between 0 and 2. The GRNN model has the lowest MSE for the test sample with the spread value 0.20. Overall, the lowest MSE for the test sample is showed by the MLP model, followed by the GRNN model and the SOM model.

4.2. Random data set

The problem in evaluating the performance of discretionary accrual estimation models is that the actual discretionary accruals cannot be measured. One method to assess the performance of the models is to randomly draw an estimation data set and an evaluation data set from the same population. As both data sets are randomly drawn from the same population, the discretionary accruals should not be significantly different from zero if the models are well specified. The closer to zero a model estimates the mean and median discretionary accruals, the better the performance.

The results in Table 3 show that all models, except for the REG model, exhibit discretionary accruals for the evaluation data set which are not significantly different from zero. The REG model shows discretionary accruals significantly different from zero (1% level) at a mean (median) of 3.3% (1.7%) of lagged total assets. The model with mean discretionary accruals closest to zero is the PWL model at 1.3% of lagged total assets, whereas the model with the median discretionary accruals closest to zero is the GRNN model at 0.1% of lagged total assets. The difference in standard deviation of the discretionary accruals between the models is marginal, ranging between 0.140 (PWL model) and 0.145 (REG model).

4.3. Stratified random data set

If the earnings management detection models are well specified, the estimated discretionary accruals should not correlate with various variables measuring performance and size. In this study, the discretionary accruals estimated with the different models are sorted according to return on assets (ROA), cash flows from operations (CFO), lagged total assets (Size), the price-to-book ratio (P/B) and the price-to-earnings ratio (P/E). For each model and variable the median discretionary accruals in the lowest and highest quartiles are examined (see Table 4). The closer to zero the median discretionary accruals in a quartile are, the better the performance of the model.

The overall performance of the models is assessed by scoring the models for each variable and quartile based on how far from zero the median discretionary accruals are. That is, the model closest to zero is assigned the value 4, whereas the model furthest away from zero is assigned the value 0. The performance of the models is then ranked based on the summed scores for each of the ten quartiles. The results show that the highest scoring model is the GRNN model. The second highest scoring model is the MLP model, followed by the PWL model. The model with the lowest score is the REG model. Furthermore, the results presented in Table 4 show that all five models show considerable differences between the discretionary accruals in the lowest and highest quartiles. This is clear especially when sorting the discretionary accruals according to ROA. Even with the PWL model, which has the lowest difference in median discretionary accruals between the quartiles when sorting based on ROA, the difference is as high as 10.7% of lagged total assets.

4.4. Discussion

The results in Sections 4.2 and 4.3 are similar to the findings in previous studies. That is, if the underlying data is correctly modeled the models based on linear regression and the models based on neural networks show a similar performance. In the REG model the data is not correctly modeled, which can clearly be seen from the biased estimates of discretionary accruals. However, when the data is modeled with the piecewise linear approach in the PWL model the performance increases considerably. In fact, the PWL model outperforms the SOM model where several local linear regressions models are created based on the nodes in a self-organizing map. Thus, considering the complexity of the SOM model, it is questionable if this approach is feasible for estimating discretionary accruals. The two best performing models are the GRNN model and the MLP model. Of these two models, the GRNN model performs marginally better showing both the lowest median discretionary accruals for the random data set in Section 4.2 as well as the highest overall score for the stratified random data set in Section 4.3. Furthermore, the training process of the GRNN model is performed in one sweep making it faster than the training process of the MLP model. The difference in the duration of the training process can be considerable when the models are trained with large data sets.

Table 3
Discretionary accruals estimated with the evaluation data set.

	REG	PWL	SOM	MLP	GRNN
Mean	0.033	0.013	0.022	0.016	0.019
Median	0.017	0.005	0.004	0.002	0.001
p-Value[a]	0	0.742	0.785	0.957	0.554
St. dev.	0.145	0.140	0.146	0.142	0.145

[a] Wilcoxon signed rank test.

Table 4
Discretionary accruals estimated with the stratified random sample.

Model	ROA		CFO		Size		P/B		P/E		Score
	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper	
REG	0.111**	0.001*	0.025**	0.035**	0.019*	0.006*	0.023**	0.020*	0.028**	0.010	13
PWL	0.075**	0.033**	0.011	0.014	0.009	0.003	0.024**	0.022	0.031*	0.003	19
SOM	0.122**	0.044**	0.023**	0.011	0.007**	0.013**	0.012**	0.009	0.024**	0.013**	16
MLP	0.089**	0.025**	0.031**	0.000	0.006	0.006	0.017**	0.005	0.021*	0.009*	24
GRNN	0.090**	0.019**	0.011*	0.012**	0.020**	0.002	0.016**	0.003	0.015*	0.009	28

* Significant at 5% level.
** Significant at 1% level.

5. Conclusion

The purpose of this study was to assess whether neural network-based models outperform linear and piecewise linear-based regression models in detecting earnings management. The results showed that the linear regression-based model had clearly the poorest performance. This was expected as previous research has shown that the relationship between the dependent variable and some of the independent variables in the model is non-linear. The models based on a piecewise linear regression, a multilayer perceptron (MLP) and a general regression neural network (GRNN) clearly outperformed the linear regression-based model. Out of these three models, the GRNN-based model showed the highest performance followed by the MLP-based model. The model based on a self-organizing map (SOM) also outperformed the linear regression-based model but not as clearly as the other three models. Furthermore, considering the complexity of the SOM-based model, its feasibility for estimating discretionary accruals is questionable. The question is could MLP- or GRNN-based models replace linear and piecewise linear-based models in accounting research? They would probably yield more precise estimates of discretionary accruals but at the same time they are more difficult to implement and the presentation of the results less intuitive. Therefore, it is likely that accounting researchers will continue using models based on linear and piecewise-linear regression for estimating discretionary accruals.

References

Ball, R., & Shivakumar, L. (2006). The role of accruals in asymmetrically timely gain and loss recognition. Journal of Accounting Research, 44(2), 207–242.

Basu, S. (1997). The conservatism principle and the asymmetric timeliness of earnings. Journal of Accounting and Economics, 24(1), 3–37.

Calderon, T. G., & Cheh, J. J. (2002). A roadmap for future neural networks research in auditing and risk assessment. International Journal of Accounting Information Systems, 3(4), 203–236.

Coakley, J. R., & Brown, C. E. (2000). Artificial neural networks in accounting and finance. modeling issues. International Journal of Intelligent Systems in Accounting, Finance & Management, 9(2), 119–144.

Collins, P., & Hribar, D. W. (2002). Errors in estimating accruals: Implications for empirical research. Journal of Accounting Research, 40(1), 105–134.

DeAngelo, L. (1986). Accounting numbers as market valuation substitutes: A study of management buyouts of public stockholders. The Accounting Review, 61(3), 400–420.

Dechow, P. M. (1994). Accounting earnings and cash flows as measures of firm performance: The role of accounting accruals. Journal of Accounting and Economics, 18(1), 3–42.

Dechow, P. M., Sloan, R. G., & Sweeney, A. P. (1995). Detecting earnings management. The Accounting Review, 70(2), 193–225.

DeFond, M. L., & Jiambalvo, J. (1994). Debt covenant violation and manipulation of accruals. Journal of Accounting and Economics, 17(1–2), 145–176.

Denton, J. W. (1995). How good are neural networks for causal forecasting? Journal of Business Forecasting Methods and Systems, 14(2), 17–20.

DeTienne, K. B., DeTienne, D. H., & Joshi, S. A. (2003). Neural networks as statistical tools for business researchers. Organizational Research Methods, 6(2), 236–265.

Healy, P. M. (1985). The effect of bonus schemes on accounting decision. Journal of Accounting and Economics, 7(1–3), 85–107.

Jaggi, B., & Lee, P. (2002). Earnings management response to debt covenant violations and debt restructuring. Journal of Accounting, Auditing and Finance, 17(4), 295–324.

Jeter, D. C., & Shivakumar, L. (1999). Cross-sectional estimation of abnormal accruals using quarterly and annual data: Effectiveness in detecting event-specific earnings management. Accounting and Business Research, 29(4), 299–319.

Jones, J. J. (1991). Earnings management during import relief investigation. Journal of Accounting Research, 29(2), 193–228.

Kaznik, R. (1999). On the association between voluntary disclosure and earnings management. Journal of Accounting Research, 37(1), 57–83.

Kothari, S. P., Leone, A. J., & Wasley, C. E. (2005). Performance matched discretionary accrual measures. Journal of Accounting and Economics, 39(1), 163–197.

Marquez, L., Hill, T., Worthley, R., & Remus, W. (1992). Neural network models as an alternative to regression. In R. Trippi & E. Turban (Eds.), Neural networks in finance and investing (pp. 435–449). Chicago: Probus (1993).

Paliwal, M., & Kumar, U. A. (2009). Neural networks and statistical techniques: A review of applications. Expert Systems with Applications, 36(1), 2–17.

Rees, L., Gil, S., & Gore, R. (1996). An investigation of asset write-downs and concurrent abnormal accruals. Journal of Accounting Research, 34(3), 157–169.

Sloan, R. G. (1996). Do stock prices fully reflect information in accruals and cash flows about future earnings? The Accounting Review, 71(3), 289–315.

Sloan, R. G. (1999). Evaluating the reliability of current value estimates. Journal of Accounting and Economics, 26(1–3), 193–200.

Teoh, S. H., Welch, I., & Wong, T. J. (1998). Earnings management and the long-run market performance of initial public offerings. Journal of Financial Economics, 53(6), 1935–1974.

Thomas, J., & Zhang, X. (2000). Identifying unexpected accruals: A comparison of current approaches. Journal of Accounting and Public Policy, 19(4–5), 347–376.

Van Tendeloo, B., & Vanstraelen, A. (2005). Earnings management under German GAAP versus IFRS. European Accounting Review, 14(1), 155–180.

Vellido, A., Lisboa, P. J. G., & Vaughan, J. (1999). Neural networks in business: A survey of applications (1992–1998). Expert Systems with Applications, 17(1), 51–70.

Vesanto, J. (1997). Using the SOM and local models in time-series prediction. In Proceedings of workshop on self-organizing maps (WSOM'97) (pp. 209–214). Finland: Espoo.

Warner, B., & Misra, M. (1996). Understanding neural networks as statistical tools. The American Statistician, 50(4), 284–293.

Whigham, P. A. (2005). Local modelling by SOM partitioning and linear regression for ecological modelling. In A. Zerger & R. M. Argent (Eds.), MODSIM 2005 international congress on modelling and simulation, December 2005 (pp. 1319–1325). Modelling and Simulation Society of Australia and New Zealand.

Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with artificial neural networks: The state of the art. International Journal of Forecasting, 14(1), 35–62.