International Conference on Business Research
University of Moratuwa, Sri Lanka
December 04, 2024

https://doi.org/10.31705/ICBR.2024.15

Paper ID: 18

OPTIMISING FINANCIAL FORECASTING: IMPLEMENTING A PREDICTIVE CASH FLOW PLATFORM FOR BANK BRANCHES

S.N. Hettiarachchi*and T.S. de Silva

Department of Decision Sciences, University of Moratuwa, Sri Lanka
*hettiarachchisn.20@uom.lk**

ABSTRACT
This study addresses the challenge of managing daily cash flows in bank branches, which often face excess or deficiency of cash, disrupting daily operations. The research focuses on developing a predictive model that can accurately forecast daily cash inflows and outflows across 175 bank branches in Sri Lanka, covering all provinces and districts. The aim is to create a robust tool that enhances financial efficiency by reducing idle cash balances while ensuring smooth operations. Two models were developed for this purpose: a multi-branch model utilizing Artificial Neural Networks (ANN) and a single-branch model using a Random Forest Regressor. The multi-branch model, which features separate sub-models for cash inflow and outflow, attained accuracies of 85.45% and 85.50%, respectively. In contrast, the single-branch model, tested on the Grandpass branch, demonstrated performance with accuracies of 60.06% for cash inflow and 70.29% for cash outflow predictions. The multi-branch model's superior performance underscores its ability to provide consistent and reliable predictions across a broader range of branches. The final models have been integrated into a web-based user interface, offering a user-friendly platform for real-time cash flow predictions. Overall, the results highlight the multi-branch model as a robust solution for effective cash flow management across bank branches.

Keywords: ANN, Cash Flow Management, Machine Learning, Multi-Branch Model, Single-Branch Model

1. Introduction
Effective cash flow management is essential for the smooth operation and financial stability of bank branches. In Sri Lanka, many bank branches face the recurring problem of having either an excess or a shortage of cash at the end of certain days, which can disrupt daily operations and lead to financial inefficiencies. Excess cash results in idle funds that could otherwise be used for investment or profit generation,

while cash shortages can delay transactions, leading to customer dissatisfaction and potential loss of business. The unpredictability of these cash flows poses a significant challenge for branch managers who need to accurately forecast cash requirements to ensure operational efficiency and financial stability.

The primary objective of this research is to predict daily cash inflows and outflows accurately to assist bank branches in managing their cash flow more effectively. By reducing idle cash balances and improving financial efficiency, this research aims to address the core problem faced by bank branches: the frequent occurrence of excess or deficient cash levels. Furthermore, this study seeks to develop a user-friendly platform that will help branch managers make real-time cash flow predictions, enabling better decision-making and minimizing cash flow unpredictability.

Accurate forecasting of cash flows has long been a topic of interest in both academic and practical settings. The Financial Accounting Standard Board (FASB) has emphasized the importance of predicting future cash flows since 1978, recognizing that historical earnings, while important, do not provide a complete picture for forecasting purposes. Previous research has largely focused on using company operational data, economic indicators, or daily transaction data to predict cash flows. Among these, accrual-based prediction methods are more reliable than cash-based methods, as accruals help reduce the noise in operating cash flows, making them more predictive of future trends. On the other hand, other business stakeholders may use historical cash flows to forecast potential cash flow investment decisions (Ali & Ali, 2021).

In this research, we developed two distinct predictive models to address the cash flow management challenges faced by bank branches in Sri Lanka. The first is a multi-branch model, utilizing Artificial Neural Networks (ANN) and Random Forest (RF) models with different scaling techniques to predict cash inflows and outflows separately. This model was applied to 175 bank branches across all provinces and districts of Sri Lanka, providing a comprehensive tool for use across the entire banking network. The ANN sub-models for cash inflow and outflow predictions achieved accuracies of 85.45% and 82.27%, respectively. Meanwhile, the Random Forest model with log transformation and Min-Max scaling performed well, achieving 73.37% accuracy for cash inflow and 85.50% for cash outflow.

The second model is a single-branch model using a Random Forest Regressor, which achieved accuracies of 60.06% for cash inflow and 70.29% for cash outflow. While the single-branch model had a higher performance for cash outflow predictions, its overall accuracy for cash inflow was lower. This suggests that, although the Random Forest approach can be effective, further optimization is necessary to improve its performance at the individual branch level.

In addition to developing these models, a web-based platform was created to make cash flow predictions accessible and easy to use for bank managers. The platform allows managers to input relevant data and receive real-time predictions, enhancing decision-making and addressing cash flow management challenges at the branch level. This study's findings, therefore, contribute to improving cash flow forecasting in the banking industry, providing a practical tool to tackle the persistent issue of cash flow unpredictability.

2. Literature Review

Future cash flows are critical for the survival of corporations. Reliable and accurate cash flow forecasting is important for academics and practitioners. (Shi, Pang, Shi, & Zhao, 2022) The Financial Accounting Standard Board (FASB) has mentioned forecast operating cash flows since 1978 which also emphasized the stronger interest of investors in historical earnings than historical cash flows to predict future cash flows (Nguyen & Nguyen, 2019).

The previous researchers mainly focused on either company operational data economic data or day data. If the researchers are mainly concerned about company operational data, they mainly focus on accrual-based and cash-based prediction methods. The accrual-based accounting method is stronger than a cash-based accounting one. (Noury, Hammami, Ousama, & Zeitun, 2020) Operating cash flows are a garbled or noisy measure of operating earnings with the noise being reduced by accruals. (Ball & Nikolaev, 2020) However, operating cash flows are of particular importance in the prediction of future operating cash flow, particularly when combined with disaggregated accruals. (Noury, Hammami, Ousama, & Zeitun, 2020) Furthermore, cash flow information has more predictive ability than earnings information in predicting future cash flows. (Mahmoud, 2021) Thus, accrual regression models are more efficient and some features such as firm size are more functional in the effectiveness (Sarraf, 2020).

The suitability of accrual-based forecasting systems within the banking industry requires a specialized analysis, as banks operate under different financial dynamics compared to other industries. Cash flow models are critical for predicting future financial performance, especially in sectors where financial operations are complex, such as banking (Ali & Ali, 2021). However, in a banking environment, accrual-based systems offer an enhanced perspective by accounting for anticipated revenues and expenses, which are vital for financial stability**.**

In the modern world, most researchers are willing to do these types of research using machine learning and deep learning techniques. There are several types of ML & DL techniques like Random Forest, SVM, Decision Tree, LSTM, ANN, etc. Furthermore, the best models are random forest and fine-tuned LSTM, depending on forecast horizons. (Riabykh, Suleimanov, Surzhko, Konovalikhin, & Ryazanov) The structure of the neural networks is tuned, and this design produces the most successful results in experiments. (Serengil & Ozpinar, 2019) Multilayer perceptron neural network is a suitable model for predicting cash flow and according to previous research findings based on multi-variable linear regression model superiority (Sarraf, 2020).

When looking at the non-linear dynamic approach models, The Grey-box model performs better in longer-term predictions. The grey-box model retains a simple white-box model structure, while their parameters are modeled as a black box with a Padé approximant as a functional form (Shi, Pang, Shi, & Zhao, 2022).

3. Methodology

Objective Alignment: Each model was developed with specific research objectives in mind. The multi-branch model aimed to predict cash inflows and outflows across multiple branches of a selected financial institution, while the single-branch model concentrated on detailed predictions for the Grandpass branch. This clear alignment with research goals ensured that the models directly addressed the institution's financial forecasting needs, allowing for tailored insights that enhance cash management strategies across varying locations.

The population for this study included all cash inflow and outflow data from the financial institution since the inception of the data collection up to the current date. For this analysis, a sample covering 1.5 years was selected. The initial dataset consisted of 13 columns and 85,632 rows, representing the multi-branch model's sample size. After applying one-hot encoding and scaling techniques, the dataset's shape transformed to 286 columns while maintaining the same number of rows. In contrast, the Grandpass branch data comprised 12 columns and 300 rows, forming the sample for the single-branch model. This extensive data preparation process ensured the dataset's integrity and suitability for machine learning algorithms, allowing for accurate predictions and insightful analysis.

	Branch	Province	District	Year	Month	Day	Holiday	Day of the week	Core Inflation(NCPI)	SDFR	AWDR	Cash Inflow	Cash Outflow
0	NUWARAELIYA	CP	NUWARAELIYA	2022	January	1	No	Saturday	12.9	5.0	4.94	6038006.00	16206.11
1	MATARA CITY	SP	MATARA	2022	January	1	No	Saturday	12.9	5.0	4.94	5753743.31	1221083.24
2	KIRIBATHGODA MAKOLA	WP	COLOMBO	2022	January	1	No	Saturday	12.9	5.0	4.94	4086377.80	404677.40
3	KIRIBATHGODA	WP	COLOMBO	2022	January	1	No	Saturday	12.9	5.0	4.94	16585023.88	4191379.79
4	VALACHCHENAI	EP	BATTICALOA	2022	January	1	No	Saturday	12.9	5.0	4.94	5032284.41	1093235.40

Figure 1: The Dataset after Preprocessing.

The following table provides a detailed summary of the key descriptive statistics for the dataset, including variables such as Year, Day, Headline Inflation (NCPI), Standing Deposit Facility Rate (SDFR), and Average Weighted Deposit Rate (AWDR). These statistics include the mean, standard deviation, minimum, and maximum values, offering insights into the overall data distribution and variability across the observed period.

	Year	Day	Headline Inflation(NCPI)	SDFR	AWDR
count	85632.000000	85632.000000	85632.000000	85632.000000	85632.000000
mean	2022.318584	15.779697	48.514377	13.202202	11.064615
std	0.465930	8.801656	18.876438	3.229876	3.801983
min	2022.000000	1.000000	16.800000	5.000000	4.940000
25%	2022.000000	8.000000	33.600000	13.500000	6.990000
50%	2022.000000	16.000000	53.200000	14.500000	12.600000
75%	2023.000000	23.000000	65.000000	14.500000	14.740000
max	2023.000000	31.000000	73.700000	15.500000	15.230000

Figure 2: Summary Statistics.

3.1. Multi-Branch Model

The multi-branch model is designed to predict cash flows across 175 bank branches in Sri Lanka, addressing the challenge of varying cash flow patterns across multiple locations. This model comprises two distinct sub-models: one for predicting cash inflows and another for cash outflows. Each sub-model is tailored to forecast specific aspects of cash flow based on a comprehensive set of variables.

The model uses a range of variables to enhance prediction accuracy:

- Branch: The specific branch under consideration.
- District: The district where the branch is located.
- Province: The province where the branch operates.
- Holiday: Indicator for public holidays which may affect cash flow.
- Year, Month, Day: Temporal variables to account for seasonal and daily variations.
- Day of the Week: To capture weekly patterns in cash flow.
- NCPI Core Inflation Rate: A measure of inflation that can impact cash flow trends.
- SDFR (Sri Lanka Deposit Facility Rate): Interest rate which influences branch operations and cash handling.
- AWLR (Average Weighted Lending Rate): Another interest rate variable affecting cash management.

Before training the model, data scaling is applied using either log transformation combined with Min-Max scaling or only Min-Max scaling, depending on the specific sub-model, to ensure inputs are on a comparable scale, which is crucial for machine learning performance.

To build the multi-branch model, we first split the data into training and testing sets. This allows us to evaluate the model's

performance on unseen data and ensures that the model generalizes well. The data is divided as follows:
- Training Set: 80% of the data used for training the model.
- Testing Set: 20% of the data used for evaluating model performance.

The sub-models for cash inflow and outflow predictions are built using Artificial Neural Networks (ANN). The choice of ANN is based on its ability to model complex, non-linear relationships in data. The structure of the ANN is as follows:
- Input Layer: Accepts input features from the preprocessed data.
- Hidden Layers: Two hidden layers with 64 and 32 neurons, respectively, using the ReLU (Rectified Linear Unit) activation function. This activation function introduces non-linearity, enabling the model to learn more complex patterns.
- Output Layer: A single neuron output for regression tasks, predicting the cash inflow or outflow value.

The ANN models are compiled using the Adam optimizer, which adjusts learning rates dynamically and mean squared error as the loss function, which measures the average squared difference between predicted and actual values. The models are trained for 50 epochs with a batch size of 32, and a portion of the training data is used for validation to monitor the model's performance and prevent overfitting.

The use of ANN in the multi-branch model is crucial due to its capability to capture and learn from the intricate patterns and relationships in the data. Unlike traditional statistical methods that may struggle with complex, non-linear relationships, ANNs offer flexibility and efficiency in handling diverse and large datasets. The ReLU activation function, in particular, helps in learning non-linear patterns without suffering from vanishing gradient problems, making the ANN well-suited for this forecasting task.

In comparison to previous research which may have employed simpler models or fewer variables, this approach leverages the power of neural networks to integrate a wide range of factors and interactions, improving the accuracy and robustness of cash flow predictions. The comprehensive feature set and advanced ANN architecture enable the model to adapt to various conditions across different branches, ultimately leading to more reliable and actionable forecasts.

The next section will detail the methodology for the single-branch model, which focuses on individual branches and employs different modeling techniques to further enhance prediction accuracy.

3.2. Single-Branch Model

The single-branch model was developed to provide a more focused prediction for cash flows at the Grandpass branch, allowing for a detailed analysis of cash management at an individual location. This model aims to improve the accuracy of cash flow forecasts by using the Random

Forest Regressor, a robust and versatile machine learning technique known for its high performance in regression tasks.

The single-branch model uses the same set of variables as the multi-branch model to ensure consistency and relevance in predictions:
- Holiday: Indicator for public holidays affecting cash flow.
- Year, Month, Day: Temporal features for capturing seasonal and daily trends.
- Day of the Week: Weekly patterns that influence cash flow.
- NCPI Core Inflation Rate: Economic variable impacting cash flow.
- SDFR (Sri Lanka Deposit Facility Rate): Interest rate affecting branch operations.
- AWLR (Average Weighted Lending Rate): Additional interest rate variable.

The data for the Grandpass branch was split into training and testing sets to evaluate the model's performance:
- Training Set: 80% of the data used to train the model.
- Testing Set: 20% of the data used for performance evaluation.

The `StandardScaler` was applied to standardize the feature values, ensuring that they are on a comparable scale. This step helps in improving the convergence and stability of the model during training.

The Random Forest Regressor was chosen for this model due to its ability to handle high-dimensional data and its robustness to overfitting. This ensemble learning method constructs multiple decision trees during training and outputs the mean prediction of the individual trees to improve accuracy and control overfitting.

The model was trained using the following parameters:
- Number of Estimators: 100 decision trees.
- Random State: 42 for reproducibility.

The performance of the Random Forest model was evaluated using the following metrics:
- Mean Squared Error (MSE): Measures the average squared difference between the predicted and actual values. Lower values indicate better model performance.
- R^2 score: Indicates the proportion of variance in the dependent variable that is predictable from the independent variables. An R^2 score close to 1 suggests that the model explains a large portion of the variance.

The model's performance was assessed on the test set, providing insights into its predictive accuracy and reliability. The metrics obtained from this evaluation help in understanding how well the Random Forest model generalizes to new, unseen data.

The Random Forest Regressor was selected for its robustness and ability to handle complex data relationships without extensive parameter tuning. Unlike traditional regression models, Random Forest can effectively manage non-linearities and interactions between variables, which is particularly useful for cash flow prediction where relationships are often intricate and non-linear. This technique's ensemble approach, averaging multiple decision trees, helps in achieving higher accuracy and reducing variance, making it well-suited for the forecasting task at the Grandpass branch.

4. Results/Analysis and Discussion

The study presents a comparative analysis of two predictive models for cash flow management at both a multi-branch and a single-branch level. The primary objective was to enhance the accuracy of cash flow forecasts to improve operational efficiency and financial stability in bank branches.

4.1. Multi-Branch Model Results

The multi-branch model, developed to serve 175 bank branches across Sri Lanka, demonstrated strong performance using both Random Forest (RF) and Artificial Neural Networks (ANN). The accuracies achieved by the models for cash inflow and cash outflow predictions are as follows:

Table 1: Accuracies of Multi-Branch Cash Flow Prediction Models.

	Cash Inflow	Cash Outflow
RF with log and min max	73.37%	85.50%
RF with min max	65.33%	53.35%
ANN	85.45%	82.72%

For cash inflow predictions, the ANN model achieved the highest accuracy at 85.45%, significantly outperforming the Random Forest models. However, for cash outflow predictions, the Random Forest model with log and Min-Max scaling outperformed the ANN, achieving a higher accuracy of 85.50% compared to the ANN's 82.27%.

These results demonstrate that while ANN is more effective for cash inflow predictions, Random Forest with log and Min-Max scaling is better suited for cash outflow forecasting. This highlights the need to carefully select models based on the specific nature of the prediction task, as different models may excel in different areas of cash flow forecasting.

4.2. Single-Branch Model Results

In contrast, the single-branch model, focused on the Grandpass branch, used a Random Forest Regressor and reported lower accuracies, with 60.06% accuracy for cash inflows and 70.29% for cash outflows. While these figures indicate that the Random Forest model performed reasonably well, its predictive accuracy for the Grandpass branch was lower compared to the multi-branch model.

Table 2: Accuracies of Single-Branch Cash Flow Prediction Models.

	Cash Inflow	Cash Outflow
Random Forest Reggresor	60.06%	70.29%

Despite the lower accuracy, we opted not to use Artificial Neural Networks (ANN) or log transformations for the single-branch model. This decision was based on several factors:
1. Model Simplicity: The Random Forest model is simpler and less computationally intensive compared to ANN, making it easier to implement and faster to train, especially with smaller datasets like those from a single branch.
2. Data Complexity: The Grandpass branch data lacks the diversity and volume of the multi-branch dataset. The complexity of an ANN may not provide significant benefits when working with limited data from a single branch, as overfitting can become a concern.
3. Interpretability: Random Forest models offer better interpretability than ANN models, which is important when trying to explain predictions in a real-world banking environment. The transparency of Random Forest models can help bank managers understand the key factors influencing predictions.
4. Log Transformation: While log transformation can enhance performance in some cases, it may not always be necessary. In this case, the model performed reasonably well without it, and adding such a transformation might complicate the model without yielding meaningful improvements.

Although the single-branch model has lower accuracy compared to the multi-branch model, its Random Forest approach provides a more straightforward, interpretable, and practical solution for predicting cash flows in individual branches, without the need for the added complexity of ANN or log transformations.

4.3. Discussion

This study demonstrates that advanced modeling techniques significantly enhance the accuracy of cash flow predictions. The results reveal that the Artificial Neural Network (ANN) model achieved an impressive accuracy of 85.45% for predicting cash inflows, positioning it as superior to traditional methods. This outperformance underscores the necessity of employing sophisticated approaches, as highlighted by Ali and Ali (2021), who noted the limitations of conventional cash flow models. Our research sets a new benchmark for future studies in cash flow forecasting by surpassing traditional models.

While Ball and Nikolaev (2020) argue that earnings often surpass cash flows in predictive capability, our findings challenge this notion by demonstrating that cash flow models—particularly when using advanced techniques—can achieve remarkable accuracy. Specifically, the cash outflow predictions with the ANN model reached 82.27%, far exceeding the performance of simpler models, such as the Random Forest (RF) with only Min-Max scaling, which recorded just 53.53%. This illustrates that our research not only confirms previous claims but also provides evidence that innovative methodologies can yield superior outcomes.

Moreover, our study's results resonate with the findings of Serengil and Ozpinar (2019), who acknowledged the potential of ANN in enhancing prediction accuracy. Unlike their work, our research not only validates the use of ANN but also highlights its dominance over other models in practical applications. The significant accuracy achieved by the ANN model, in contrast to the modest performances of traditional models, positions our research as a leading contribution in the field.

Furthermore, the successful integration of historical data with advanced techniques aligns with the views of Mahmoud (2021) and Nguyen and Nguyen (2019), who emphasize the importance of accrual-based accounting in cash flow forecasting. Our findings go a step further by demonstrating the practical superiority of these approaches. In light of this, our research paves the way for future studies to explore hybrid models, as suggested by Noury et al. (2020), thereby further enhancing the precision of cash flow forecasts.

4.4. Limitations and Future Research Opportunities

While this research offers valuable insights into cash flow management through predictive modeling, it also has several limitations that should be acknowledged. Firstly, the models were developed using historical data from a specific time frame (January 1, 2022, to May 7, 2023) and may not account for significant economic changes or external factors that could affect cash flows in the future. This temporal limitation may impact the models' adaptability to evolving market conditions, seasonal variations, or changes in consumer behavior. Additionally, the study focused primarily on cash inflows and outflows within the selected financial institutions and may not encompass other factors influencing cash management, such as macroeconomic indicators or regulatory changes. The reliance on historical data also poses challenges in predicting unprecedented events, such as economic crises or natural disasters, which could significantly disrupt cash flow patterns.

Future research could explore the integration of more diverse datasets, including real-time economic indicators, customer transaction patterns, and broader financial metrics, to enhance the predictive capabilities of the models. Investigating the application of advanced machine learning techniques, such as ensemble methods or deep learning approaches, could also yield improved accuracy and reliability in forecasts. Moreover, expanding the study to include a wider range of financial institutions across different regions may provide a more comprehensive understanding of cash flow dynamics and contribute to developing universally applicable cash management strategies.

By addressing these limitations and exploring these future research opportunities, scholars can further enhance the robustness of cash flow management models and provide more effective tools for financial institutions.

5. Conclusion and Implications

This study developed and assessed two predictive models for managing cash flows in Sri Lankan bank branches: a multi-branch model using Artificial Neural Networks (ANN) and a single-branch model with a Random Forest Regressor. The multi-branch ANN model performed well, achieving 85.45% accuracy for cash inflows and 85.50% for cash outflows. This demonstrates that advanced machine learning techniques can effectively handle the complexities of cash flow management across various branches, leading to better financial efficiency and reduced idle cash.

However, the single-branch model, which focused on the Grandpass branch, showed lower accuracies of 60.06% for inflows and 70.29% for outflows. This suggests that while localized models offer useful insights, their performance can be limited by branch-specific data issues and variability. These lower accuracies highlight the need for further improvements to enhance the reliability of single-branch predictions.

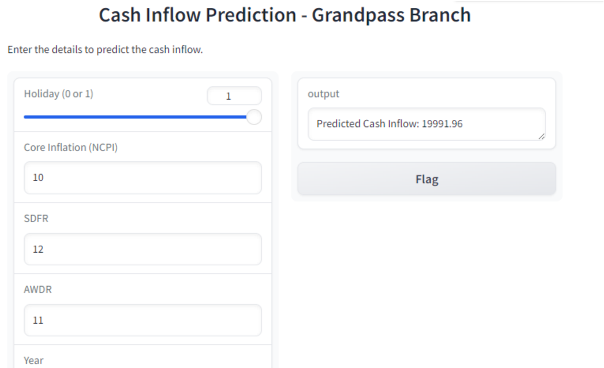

Figure 3: Example for Cash Inflow Prediction – Grandpass.

The study confirms that machine learning can significantly improve financial forecasting, with the multi-branch model providing a practical tool for broad application. For future work, refining the single-branch model with additional features and techniques could enhance its accuracy and utility. This would help bridge the gap between broad and localized models, leading to more effective cash flow management across all branches.

Acknowledgment
I would like to express my sincere gratitude to the financial institution for providing the critical data that made this research possible. Their generous support was essential for the success of this study, significantly contributing to advancements in cash flow forecasting and enhancing our understanding of effective financial management practices.

References

Ali, M. M., & Ali, K. M. (2021). USING A CASH FLOW MODEL TO PREDICT FUTURE CASH FLOW FROM HISTORICAL CASH FLOW. Academy of Accounting and Financial Studies Journal.

Ball, R., & Nikolaev, V. (2020, December). FASB was right: Earnings beat cash flows when predicting future cash flows.

Mahmoud, H. H. (2021). Earnings Versus Cash Flows in Predicting Future Cash Flows. Alexandria Journal of Accounting Research, 5(1).

Nguyen, H. A., & Nguyen, T. H. (2019). The prediction of future operating cash flows using accrual-based and cash-based accounting. Management Science Letters.

Noury, B., Hammami, H., Ousama, A., & Zeitun, R. (2020). The prediction of future cash flows based on operating cash flows, earnings, and accruals in the French context. Journal of Behavioral and Experimental Finance.

Riabykh, A., Suleimanov, I., Surzhko, D., Konovalikhin, M., & Ryazanov, V. (n.d.). ATM Cash Flow Prediction Using Local and Global Model Approaches. APPLIED PROBLEMS.

Sarraf, F. (2020). Cash flow forecasting by using simple and sophisticated models. Iranian Journal of Finance.

Serengil, S. I., & Ozpinar, A. (2019). ATM Cash Flow Prediction and Replenishment Optimization with ANN. International Journal of Engineering Research and Development.

Shi, S., Pang, Y., Shi, Y., & Zhao, Y. (2022). A nonlinear dynamic approach to cash flow forecasting. Review of Quantitative Finance and Accounting.