# Empowering cash managers to achieve cost savings by improving predictive accuracy

Francisco Salas-Molina [a,*], Francisco J. Martin [b], Juan A. Rodríguez-Aguilar [c], Joan Serrà [d], Josep Ll. Arcos [c]

[a] Hilaturas Ferre, S.A., Les Molines, 2, 03450 Banyeres de Mariola, Alicante, Spain
[b] BigML, Inc., 2851 NW 9th Suite D, Conifer Plaza Building, Corvallis, OR 97330, United States
[c] IIIA-CSIC, Campus UAB, 08913 Cerdanyola, Catalonia, Spain
[d] Telefonica Research, Ernest Lluch i Martín, 5, 08019 Barcelona, Catalonia, Spain

## ARTICLE INFO

## ABSTRACT

Cash management is concerned with optimizing a company's short-term funding requirements. To this end, various different optimization strategies have been proposed for minimizing costs, using daily cash flow forecasts as the main input to the models. However, the effect of the accuracy of such forecasts on cash management policies has not been studied. This article uses two real data sets from the textile industry to show that the predictive accuracy is highly correlated with cost savings when daily forecasts are used in cash management models. A new method is proposed to help cash managers estimate whether their efforts in improving the predictive accuracy are rewarded by proportional cost savings. Our results indicate the need for an analysis of the potential cost savings derived from improving the predictive accuracy. On that basis, the search for better forecasting models is in place in order to improve cash management.

## 1. Introduction

Cash flow management is concerned with the efficient use of a company's cash as a critical task in working capital management. Decision making in cash flow management focuses on maintaining the balance between what the company holds in cash and what has been placed in short-term investments, such as deposit accounts or treasury bills. Various different models have been designed to answer these questions, and reviews are provided by da Costa Moraes, Nagano, and Sobreiro (2015), Gregory (1976), and Srinivasan and Kim (1986). However, to the best of our knowledge, little attention has been paid to the utility of cash flow forecasts, with the exception of Gormley and Meade (2007) and Stone (1972). Both studies showed the utility of forecasting in cash management, but neither researched the importance of the predictive accuracy of the forecasts used in their respective models. Therefore, it is unknown whether even small improvements in predictive accuracy may lead to savings that could perhaps amount to millions of euros in total.

Cash planning deals with the forecasting of cash balances, short-term investment and short-term borrowing (Fabozzi & Masonson, 1985; Stone, 1973). The cash management problem was first addressed from an inventory control point of view by Baumol (1952) in a deterministic way. Miller and Orr (1966) introduced a simple stochastic approach by considering a symmetric Bernouilli

---

\* Corresponding author.
  *E-mail addresses:* francisco.salas@hifesa.com (F. Salas-Molina),
martin@bigml.com (F.J. Martin), jar@iiia.csic.es (J.A. Rodríguez-Aguilar),
joan.serra@telefonica.com (J. Serrà), arcos@iiia.csic.es (J.L. Arcos).

process in which both the inflow and the outflow were of exactly the same size and had probability 1/2. Later, while Girgis (1968) considered continuous net cash flows with both fixed and linear transaction costs, Eppen and Fama (1969) focused on discrete net cash flows with only variable transaction costs. The use of forecasts in the corporate cash management problem was first introduced by Stone (1972) as a way of smoothing cash flows. More recently, Gormley and Meade (2007) used the model proposed by Penttinen (1991) as a benchmark in order to demonstrate the utility of cash flow forecasts in the cash management problem. They proposed a dynamic simple policy for minimizing transaction costs, under a general cost structure, and developed a time series model for providing forecasts. Surprisingly, even though their model is based on cash flow forecasts, they ignored the possibility of exploring alternative forecasting methods. We claim that this step is mandatory, especially when improvements in forecasting accuracy may be correlated with cost savings. This hypothesis was suggested by the authors but was not verified.

In cash flow forecasting, Maier, Robinson, and Vander Weide (1981), Miller and Stone (1985), Stone (1976), Stone and Miller (1981), Stone and Miller (1987) and Stone and Wood (1977) presented different useful linear models. A measure of the quality of any forecasting technique is its predictive accuracy, and, from an economic perspective, the predictive accuracy must be mapped to estimated cost savings. This analysis assesses how much companies can save by improving their predictive models, and, consequently, the cost of not predicting, i.e., the missed savings minus the cost of implementing the model. For example, if a 32% reduction in forecasting error produced savings of €320.000 per year, it can be stated that, on average, each percentage point of predictive accuracy is €10.000 worth.

Overall, one can conclude that the predictive accuracy of the forecasts used in cash management models has attracted little attention from the research community, neglecting its significance and implications. As a consequence, our discussion seeks to assess the quality of alternative forecasting methods. In this paper, we present and compare various different forecasting methods, including both linear and non-linear models. In this sense, we expect that non-linear models will be able to deal with cash flow time series in a cost-saving manner. For the sake of simplicity, we restrict our analysis to transaction and holding costs for cash balances in a single currency. Using two real data sets from companies in the textile industry in Spain, this paper: (i) shows empirically that the forecasting accuracy is highly correlated with savings in cash management, and, thus, performs a comparison in terms of accuracy and savings between different forecasting models; (ii) argues that the effect of the forecasting accuracy on cash management can be estimated in advance and, thus, proposes a new methodology for estimating this effect.

The rest of this paper is organized as follows. We begin by describing our real cash flow data sets in Section 2. We then enumerate different forecasting models in Section 3: linear models such as autoregressive and regression models, and non-linear models such as the radial basis function, random forests and seasonal indicator models.

These forecasting models are then ranked according to their predictive accuracies in the evaluation in Section 4. Section 5 empirically verifies that a better forecast produces a better policy. Moreover, we estimate what savings (if any) can be obtained by using the cash policies produced by an improvement in forecasting accuracy. Finally, Section 6 concludes.

## 2. Description and data preprocessing

This section describes the two real cash flow data sets used in this paper. Data sets 1 and 2 gather the net daily flow on workdays from two different companies in the textile industry. Both sets of observations are in the domain of real numbers, and the distributions of their values show bell-like shapes but excess kurtosis. An additional transformation is performed to deal with anomalies. More specifically, following the recommendations of Gormley and Meade (2007) and Hyndman (2016), any observation that is more than five times the standard deviation is considered an outlier and replaced by a linear interpolation. After time series cleaning, the empirical distributions of our cash flow data sets are shown in Fig. 1. In order to cover a wider range of realistic industrial company cases, a number of cash flow data sets are derived from the two real data sets as follows:

- Real cash flow: data sets 1 and 2.
- Stable cash flow: data set 3 is derived from data set 1 and applies to companies that are in a more stable environment, with daily cash flows that are characterized by low variances. In this case, observations that are greater than three times the standard deviation are replaced by values of exactly three times the standard deviation.
- Unstable cash flow: data set 4 is also derived from data set 1 and applies to companies with high variances in their daily cash flows, which may occur for various different reasons, such as lower numbers of customers or suppliers, as is the case with small companies. In this case, observations that are greater than three times the standard deviation are replaced by values of exactly double the original observation.
- Random shock cash flow: data sets 5 and 6 are derived from the original data sets 1 and 2 respectively, before anomaly detection, and aim to cover the likely occurrence of unexpected changes in industrial markets. In this case, 5% of the observations are chosen at random and replaced by randomly selected values from the previous set of labeled outliers. This transformation simulates a disturbance that, to some extent, may break existing patterns and increase the deviations from the expected cash flow by introducing a number of outliers in the original data. In addition, it also helps to estimate the impact of outliers in accuracy.

A summary of the characteristics of each data set is presented in Table 1.

It is important to note at this point that we are dealing with the net cash flow without component separation. Mixing different cash flows may compromise the attainable accuracy due to the different statistical properties of cash flow components, such as major and