U.Porto Journal of Engineering, 9:1 (2023) 1-15
ISSN 2183-6493
DOI: 10.24840/2183-6493_009-001_000976

Received: 16 July, 2021
Accepted: 26 October, 2022
Published: 23 January, 2023

Assessing Budget Risk with Monte Carlo and Time Series Bootstrap

André Pereira[1], A. Miguel Gomes[2]
[1]Faculdade de Engenharia, Universidade do Porto, Rua Dr. Roberto Frias, 4200-465 PORTO, Portugal (up201606906@g.uporto.pt) ORCID 0000-0002-4067-0469
[2]INESC TEC, Faculdade de Engenharia, Universidade do Porto, Rua Dr. Roberto Frias, 4200-465 PORTO, Portugal (agomes@fe.up.pt) ORCID 0000-0002-6251-2496

Abstract
Budgets are important management tools recognized for their help in planning, communication, monitoring the expense performance, and even motivating collaborators. However, recently there has been criticism of the traditional Budgeting Process due to its cumbersomeness, long duration, and eventual diversion of the focus from the day-to-day activities. Thus, improving the Budgeting Process by incorporating Expense component uncertainties is of uttermost importance to accelerate its approval.
This paper presents a methodology for companies to assess their budget risk based on their historical Expense data by applying Monte Carlo Simulation and Time Series Bootstrapping Techniques. Besides, some state-of-the-art sensitivity Importance Measures are also implemented to help evaluate the relative importance of the Expense components. The methodology proposed, based on a real case study with data from a major Portuguese retailer, has the advantage of being objective and supported by data, thus not being subject to bias from the management.

Author Keywords. Risk Management, Risk Assessment, Monte Carlo Simulation, Time Series Bootstrap, Budget Risk; Uncertainty Importance Measures

Type: Research Article
Open Access Peer Reviewed

1. Introduction

Budgets are management tools widely used to support planning and help managing resources and activities of organizations (Lidia 2014). There are several studies pointing out the advantages of Budgeting in an organization. Lidia (2014) underlines the positive impact on managers by giving them a sense of safety and certainty when operating in an uncertain environment. Libby and Lindsay (2010) refer to the support they offer in translating the strategy into more objective actions whilst interacting with the company's divisions. Other advantages expressed are its positive role in planning, resource allocation, performance evaluation, and even in motivating collaborators.

However, recently some authors have also criticized the usage of these tools and have urged its improvement. Common drawbacks of its implementation are its high cost and the fact that it is very time-consuming, which can hinder the focus on the core activities and innovation (Lidia 2014). Moreover, the fact that the majority of budgets are deterministic and do not consider uncertainties in the Budgeting Process often results in inadequate plans (Lord 1977; Rubin and Patel 2017; Hager, Yadavalli, and Webber-Youngman 2015; Crum and Rayhorn 2019). Some even recognize that they can be easy to manipulate through controlling the budgetary slack. That is, "the intentional overestimation of expenses and/or underestimation of revenue during budget setting" (Eaton 2005, 7). This is where the application of risk management techniques have an important role in helping improve budgets by considering the uncertainties in the Budgeting Process and helping reduce the budgetary slack existent by defining appropriate contingencies, leading to more accurate and effective budgets (Elmassri and Harris 2011; Robinson, Chariri, and Prabowo 2018). For this matter, Simulation can be a useful tool to consider different scenarios and support the decision-making regarding budget approval.

This paper contributes to scientific knowledge with a methodology for companies to assess their budget risk. It is an alternative to assuming a specific distribution to Expense components and relying heavily on the opinion of experts to estimate limit values, widely used in Project Management. The distributions are then obtained from the historical data Simulation by combining Time Series Bootstrapping techniques with Monte Carlo. This is a novel approach from an academic standpoint as from the literature reviewed, the Time Series Bootstrapping techniques are more applied in forecasting studies and not risk analyses. Besides, in this study, Monte Carlo Simulation is used to assess the risk of the primal Budget of a company and not a specific project. This application helps fill the gap existent in the literature as in those few studies that apply Monte Carlo to the primal Budget; the Total Expense distribution ends up being very spread due to the various subjective assumptions (Hager, Yadavalli, and Webber-Youngman 2015). For managers, this approach also has advantages, namely its practicality and not necessity to subjectively make assumptions, eventually making the justifications and Budgeting Process faster. In addition, those responsible for allocating the Budget to the teams can previously know the risk of a cost overrun of a certain allocation based on the historical and simulated data, eventually contributing towards a fairer organization.

The remainder of this paper is organized as follows. Section 2 provides an overview of simulation-related work with a particular focus on its application to Budgeting. We characterize the dataset used in this paper in Section 3, and in Section 4, the methodology followed is presented. Section 5 presents and discusses the results, and, lastly, Section 6 summarizes the main findings and outlines opportunities for future research.

2. Theoretical Background

This section refers to relevant papers related to Simulation and its usage in estimating budget risk. Moreover, it presents some relevant Bootstrapping methods used for Time Series and some Importance Measures applied to rank the influence each input simulation variable has on the outcome uncertainty, even when correlations are present.

2.1. Budgeting under Uncertainty and Application of Monte Carlo Simulation

Collier and Berry (2002) define the budgeting process as a “formal method by which plans are established for future time periods”. The authors consider that uncertainty should be considered while budgeting, though it may be set aside in the budget (i.e., the final document). Several authors agree with this statement and even state that the main critics referred to budgets (being easy to manipulate or difficult to prepare) are founded on the fact that the process is deterministic and only considers point estimates (R.J. Lord 1977; Rubin and Patel 2017; Hager, Yadavalli, and Webber-Youngman 2015; Crum and Rayhorn 2019). Thus, the traditional plans tend to ignore uncertainty or significantly reduce the potential risk the organization faces (R.J. Lord 1979).

Scott (1998) referred that risk can be included in budgeting through three ways: excluding risk, comparing risk or modelling risk. The first does not really include risk in the budgeting process; the second uses simple analysis such as sensitivity analysis and probability; and the last has an explicit formal use of probability models.

Modelling risk in budgets is also known as Probabilistic Budgeting, and it can be done with a Monte Carlo Simulation that allows for a robust analysis of the potential risk associated with each input variable of the budget. Stochastic Simulation gives useful insights about the main drivers of the outcome of analysis (e.g., profit, costs, Net Present Value etc.). This technique is widely present in the literature as a tool to help create budgets in the field of Project Management applied to construction projects (Urgilés, Claver, and Sebastián 2019) or as a capital budgeting tool (Platon and Constantinescu 2014).

Nevertheless, the literature about the application of Monte Carlo Simulation in modelling the primal financial budget of a company is scarce (Hager, Yadavalli, and Webber-Youngman 2015). Hager, Yadavalli, and Webber-Youngman (2015) pointed out the large widespread obtained for the outcome distribution (i.e., total profit or expense) as a result of considering a vast number of variables and the presence of interactions in the model as the main reason for the lack of application of stochastic modelling to the Budgeting Process of firms. Moreover, the subjective uncertain inputs which are usually included in the models are also a cause of great variability. This criticism traces back to R.J. Lord (1977). Additionally, Shim and Siegel (2005) refer to the difficulty of creating the model as another possible reason.

To overcome these problems, Hager, Yadavalli, and Webber-Youngman (2015) recommends considering the model only as “accurate enough” and solely include primary drivers of the output variable. For this matter, Moro Visconti, Montesi, and Papiro (2018) outlined the importance of big data on the predictive ability of stochastic modelling, especially for short term budgets. **Table 1** summarizes what has been done and its relevance and limitations.

Study	Relevance	Limitations
R.J. Lord (1977)	It was an initial study about the topic. Introduced Probabilistic Budgeting and applied Monte Carlo Simulation to obtain the profit distribution of a small company.	Did not consider correlations in the study, and the input distributions and ranges were defined subjectively, resulting in widespread outcome distribution.
Platon and Constantinescu (2014)	Capital Budgeting study that shows how Monte Carlo can be used to obtain the total cost of a specific project/investment.	Assumes specific distributions.
Urgilés, Claver, and Sebastián (2019)	Typical Project Management study of assessing Budget Risk of a specific project.	Assumes specific distributions.
Hager, Yadavalli, and Webber-Youngman (2015)	Concludes that literature is scarce about the application of Monte Carlo to the primal budget of a company. One of the main reasons is due to subjective assumptions that lead to widespread outcome distribution.	Only initial exploratory study without validation of recommendations.

Table 1: Research Works About the Application of Monte Carlo Simulation to Budgets

2.2. Sensitivity Analysis and Importance Measures

Saltelli (2002) defines Sensitivity Analysis (SA) as the study of how uncertainty in the output is allocated to different sources of input uncertainty. Sensitivity analysis helps managers understand the developed models, being a bridge between the analysts and decision-makers (Emanuele Borgonovo and Plischke 2016). Importance Measures are the measures given by Sensitivity Analysis that allow ranking inputs based on their influence on the uncertainty of the model output.

There are many classifications of Importance Measures, being the most relevant Variance-based measures and Density-based ones.

Variance-based sensitivity measures are used to assess the contribution that each input of the Simulation has on the variance of the output's distribution. Xu and Gertner (2008) proposed using regression to decompose the total variance explained by each input (S_i) in its uncorrelated and correlated parts. Hao et al. (2012) generalized this analysis for nonlinear additive models using Artificial Neural Networks (ANNs). Other Importance Measures introduced are Density-based methods, such as the δ_i of E. Borgonovo (2007), that are based on the difference between probability distributions of the output and of the conditional output known the value of a certain input. This class of methods have the advantage of being moment free, meaning that they look at the entire output distribution without referring to any of its moments. This is particularly relevant when there are correlations between inputs, as in this case, Variance-based methods fail to determine the most influential set of inputs, making Density-Based the preferred methods (E. Borgonovo 2006).

2.3. Bootstrap Methods for Dependent Data (Time Series)

The Bootstrap is a technique proposed by Efron (1979) as an extension of the Jackknife method. It consists of sampling with replacement from a sample (resampling) so as to obtain an empirical distribution that by maintaining the original relation between the "population" and the sample will approximate an unknown theoretical population distribution, as the number of resamples gets bigger (the law of large numbers).

The Independent and identically distributed (IID) bootstrap (Efron 1979) is inadequate for time series data as it completely ignores its dependence structure (Lahiri 2003). Many techniques were developed to overcome this limitation, being the Block Bootstrap methods the basis of many of them. However, solely applying the Block Bootstrap method is not enough if the series is not stationary. For this matter, model-based bootstrapping methods were introduced that first apply a statistical model to fit the data and only bootstrap the resultant residuals. The generated bootstrapped series is then the sum of the fitted values derived from the statistical model with the bootstrapped residuals. We also refer to the study of Bergmeir, Hyndman, and Benítez (2016), who developed a novel bootstrapping procedure involving a Box-Cox transformation, Seasonal and Trend decomposition using Loess (STL) and the Moving Block Bootstrap (MBB). This method has the advantage of not being model-dependent and not assuming that the residuals of the STL decomposition are IID. Furthermore, for particularly noisy time series, Laurinec et al. (2019) proposed the K-Means based Bootstrap (KM), which does not create bootstrap series with a model, but samples from similar points of the original time series. **Table 2** shows the evolution of the Bootstrapping techniques and their application to time series.

Bootstrap Method	Study	Relevance
IID Bootstrap	Efron (1979)	The first study on the bootstrapping technique. Insufficient when dealing with time series data, as it does not capture its dependence structure.
Moving Block Bootstrap (MBB)	Liu and Singh (1992)	Can capture the dependence of data only if the series is stationary.
Model Based Bootstrap methods	Cordeiro and Neves (2009)	For particular seasonal series, it captures the dependence structure and its evolution in time.
BLD Bootstrap	Bergmeir, Hyndman, and Benítez (2016)	Bootstrap procedure for dependent data that does not rely on a specific statistical model.
K-Means Based Bootstrap (KM)	Laurinec et al. (2019)	Can adapt better when series are noisy.

Table 2: Evolution of the Bootstrapping Techniques for Time Series data

3. Data Characterization

The data used in this project consists of the monthly expenses from 2016-2020 for the teams of a major Portuguese retailer's division. For confidentiality reasons, the data was normalized. For each of its six teams (see **Table 3**), the two most relevant Expense components were analyzed: W12 Personnel Expenses and W13 Operating Expenses, making a total of 12 Expense components subject of analysis.

The preliminary analysis verified that the components of class W12 Personnel Expenses had annual seasonality and corresponded to approximately 90% of the Total Expenses. The W13 Operating Expenses class, on the other hand, had higher percentage deviations from the budget, which can be related to its more random nature (i.e., the not existence of seasonality) and less time dependence existent.

All the analyses were performed using the R programming language, and the conclusions for the trend and seasonality studies were derived from the application of the Cox-Stuart and the Friedman tests, respectively, that can be obtained by the *seasplot()* function from the R package *tsutils*. **Table 3** shows the results of these tests.

		Trend		Yearly Seasonality	
Component	**Team**	**p-value[a]**	**Evidence of Trend**	**p-value[b]**	**Evidence of Seasonality[c]**
W12 Personnel Expenses	Accounting PT	0.181	FALSE	≈0	TRUE
	Accounts Payable	≈0	TRUE	≈0	TRUE
	Accounts Receivable	≈0	TRUE	0.002	TRUE
	Consolidation	≈0	TRUE	≈0	TRUE
	Document Solutions	≈0	TRUE	0.009	TRUE
	BPI_Innovation	0.5	FALSE	0.038	TRUE
W13 Operating Expenses	Accounting PT	≈0	TRUE	0.073	FALSE
	Accounts Payable	0.001	TRUE	0.515	FALSE
	Accounts Receivable	0.292	FALSE	0.270	FALSE
	Consolidation	≈0	TRUE	0.137	FALSE
	Document Solutions	≈0	TRUE	0.454	FALSE
	BPI_Innovation	≈0	TRUE	0.015	TRUE

[a]p-value to test evidence of trend, obtained according to the Cox-Stuart test
[b]p-value to test evidence of yearly seasonality, obtained according to the Friedman test
[c]The FALSE evidence conclusion means that the null hypothesis of the tests (i.e., series does not have trend/seasonality) failed to be rejected

Table 3: Summary of Trend and Seasonality Study

4. Methodology

Throughout this paper, the risk of a certain budget is characterized by the probability and consequence of the expenses exceeding the budget. The probability of a cost overrun quantifies the uncertainty that the performance will deviate from a certain budgeted level. The consequence is expressed by the expected and maximum deviation of the Total Expenses from the budgeted level if the budget is surpassed.

For this purpose, Simulation is used because it enables evaluating input variability's impact on Total Expenses, avoiding the complexity of employing mathematical operations with random variables (Bakhshi and Touran 2014). Besides, in the simulation model implemented, historical data will serve as a basis creating the distribution of each component's Expense. This way, many relationships between its inner variables are already regarded, and more realistic results can be obtained. This approach makes it faster to evaluate the risk of a budget and not consider many necessary inputs to be given by the decision-maker, which can be a source of bias in many risk analyses (Paté-Cornell and Dillon 2006).

The approach followed can be divided in five steps (A-E), as shown in **Figure 1**.

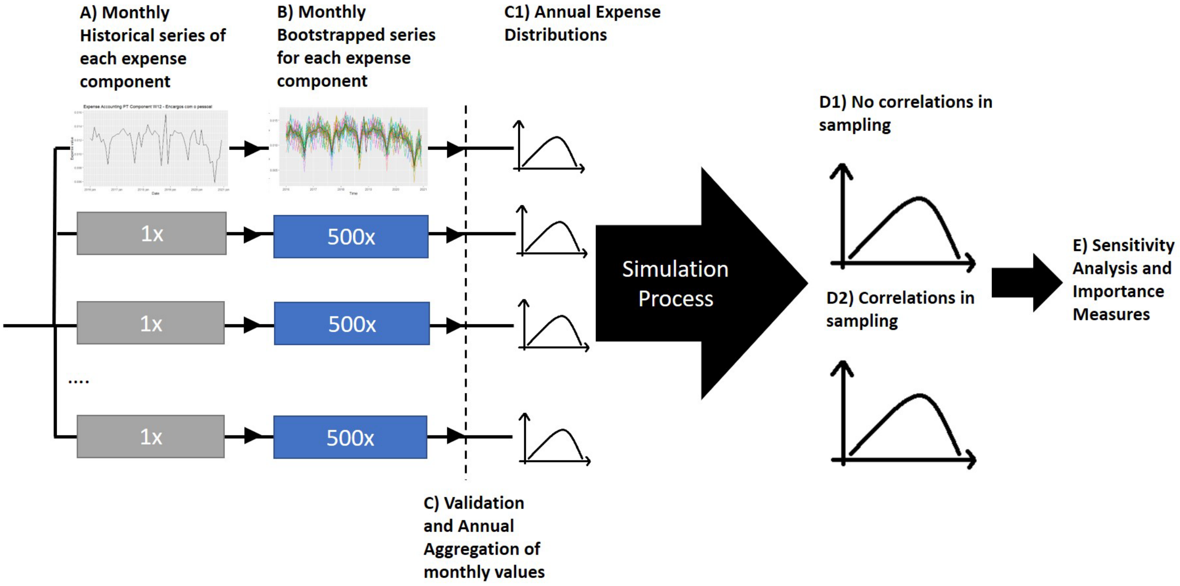


Figure 1: Scheme of the Steps Implemented

These five steps are the following:

a) The starting point was an analysis of the monthly historical series of each of the 12 Expense components (see Section 3).

b) From these series and in order to expand the data necessary to conduct the Simulation, several possible series were created from the original ones through time series bootstrapping techniques. This enabled to introduce variability in the input data and simulate different plausible expenses occurring for each component. The techniques implemented are the Moving Block Bootstrap (MBB), Model-Based techniques (based on Exponential Smoothing, SARIMA and TBATS), the Box-Cox and Loess-based decomposition (BLD) Bootstrap and the recent K-Means Based Bootstrap (KM). For each Expense component, 500 alternative series were created by each of these bootstrapping techniques.

c) The next step was to do a validation process, in which, for each Expense component, the technique that generated the most similar plausible series was selected. This choice was based on the lowest median Normalized Root Mean Square (NRMSE) criteria of the 500 series created, as it penalizes values very different from the original expense series and maintains the time dependence of the original series. Moreover, to ensure that the best selected generated series had values that could be considered to belong to the same distribution as the original ones, a Kolmogorov-Smirnov test (KS) was also performed on each of the 500 best series selected. The minimum p-value was evaluated and compared to the threshold of 5%. If above, the null hypothesis that the two series follow the same distribution failed to be rejected, and so the method successfully created similar series. Afterwards, and as the expense budget to be analyzed is assessed annually, an annual aggregation of the monthly values of each best-selected series was performed, resulting in the annual distribution for each Expense component (C1).

d) Then with these input distributions, a Monte Carlo Simulation was conducted, for which two approaches were adopted related to the sampling procedure, or in other words, the way the realizations were sampled from each of these distributions. The first one (D1) was the standard procedure in which each sampling was assumed to be independent, meaning that there was total freedom to choose a value from the distribution. The second (D2) was considering the correlations existent between each of the annual expenses. This second approach is motivated because, by reality, correlations between variables may exist and consider those will yield more accurate results and a lower forecast confidence interval if negative correlations prevail (Mun 2012). To generate correlated random expense realizations, the procedure of Neine and Curran (2021) was used.

e) Finally, to assess the importance and contribution that each input expense variable has on the Total Expense distribution generated, a sensitivity analysis with the variance explained by each expense component ($\boldsymbol{S_i}$) according to Hao et al. (2012) and the moment independent index $\boldsymbol{\delta_i}$ were implemented.

5. Results

5.1. Performance of Time Series Bootstrapping Techniques

As previously explained in Section 4, several Time Series Bootstrapping techniques were implemented to create similar monthly expense series for each Expense component. For the techniques that rely on the definition of the block length (MBB, Model-Based techniques and BLD), different lengths were tested to evaluate the impact of this parameter on the results. The lengths of the blocks tested were 1; 3; 6; 12; and 24 months. The choice of length 1 can be justified if the residuals to be bootstrapped for the model are independent and do not have any remaining time dependence. For the case of the MBB, it corresponds to the normal Bootstrap of Efron (1979). The other values were chosen due to the general fact that the expenses can be analyzed in quarters, so there can be seasonality of multiple of this time period. The maximum length of 24 derives from the recommendation of Bergmeir, Hyndman, and Benítez (2016) to ensure that any remaining seasonality is captured. As for the K-Means Based bootstrap, the maximum number of clusters allowed to be formed was 6 to allow some variability in the data but at the same time isolate more distinct values. In all the methods, five hundred series were created.

The performance of these techniques was evaluated with the median NRMSE for each Expense component. **Table 4** shows the results of the Expense component *Accounting PT W12 Personnel Expenses*. The technique that delivered the best results was the BLD with a block length of 3 and a corresponding NRMSE of 0.1142. The five hundred generated series with the BLD bootstrap for this expense component and the original monthly expense series can be seen in **Error! Reference source not found.** in colors and in black, respectively.

Method	**Block Based Methods**					**K-Means**	
	Block Length					Number of clusters	NRMSE
	1	3	6	12	24		
MBB	0.2540	0.2537	0.2522	0.2423	0.2488	-	-
ETS	0.1444	0.1425	0.1453	0.1483	0.1587	-	-
SARIMA	0.1599	0.1589	0.1607	0.1640	0.1604	-	-
TBATS	0.1372	0.1338	0.1337	0.1362	0.1396	-	-
BLD	0.1176	0.1142	0.1152	0.1168	0.1205	-	-
KMEANS	-	-	-	-	-	2	0.1321

Table 4: NRMSE Values of the Different Time Series Bootstrapping Techniques for *Accounting PT W12 Personnel Expenses*

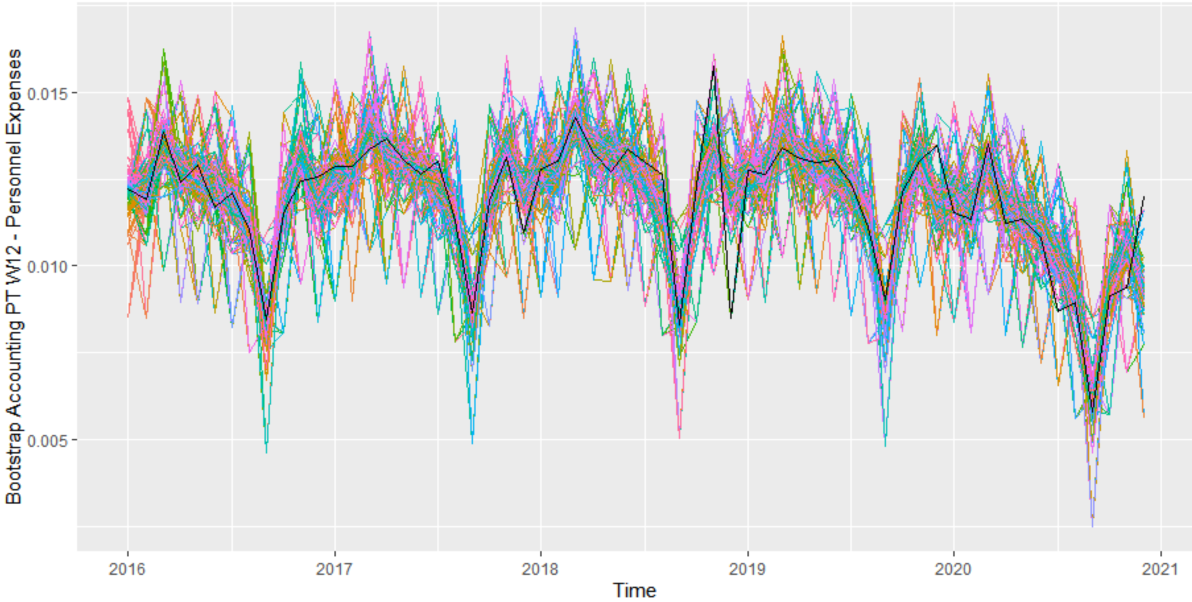


Figure 2: Bootstrap Generated Series with BLD for Accounting PT W12 Personnel

A summary of the results concerning the best techniques and their parameter values is presented in **Error! Reference source not found.**. Additionally, it can be seen the minimum p-value obtained from doing the KS test to each of the five hundred generated series as described in Section 4.

Expense Component	Time Series Method	Block Length or Number of clusters	NRMSE	Minimum p-value of KS test
Accounting PT W12	BLD	3	0.114	0.120
Accounting PT W13	KM	4	0.095	0.375
Accounts Payable W12	KM	6	0.057	0.809
Accounts Payable W13	KM	5	0.045	0.181
Accounts Receivable W12	KM	5	0.072	0.375
Accounts Receivable W13	KM	3	0.094	0.181
Consolidation W12	KM	4	0.069	0.509
Consolidation W13	KM	6	0.041	0.509
Document Solutions W12	KM	3	0.106	0.267
Document Solutions W13	BLD	3	0.176	0.120
BPI_Innovation W12	KM	5	0.062	0.375
BPI_Innovation W13	KM	6	0.057	0.660

Table 5: NRMSE of the Best Selected Time Series Bootstrapping Technique for Each Expense Component

The results demonstrate that, generally, the MBB technique yields the worst results. This can be justified as the majority of the expense series in the analysis is not stationary. Moreover, as of the model-based techniques implemented, TBATS performs better, but BLD can create even more similar series. This may be because BLD can also account for non-seasonality with the Loess Decomposition, contrarily to TBATS that works better for seasonal components.

Another relevant insight is that the choice of the model is more important and has more impact on the results than the block length that only changes the NRSME values slightly. This suggests that the residuals have the same magnitude and do not present much time dependence between them. This result goes in accordance with Radovanov and Marcikić (2014).

The technique that generally had the best results was the K-Means Based Bootstrap. This has to do with its ability to isolate more extreme values in separate clusters and randomly choose from the rest of the points. The end result is that generated series are more similar to the original one, as suggested by the higher p-values from the KS test.

Moreover, it can be concluded that the components that had the highest NRMSE were the *Document Solutions W13*, *Accounting PT W12* and *Document Solutions W12*. It can also be verified that mostly there is a higher gap of NRMSE values between block length-based techniques and the K-Means for the W13 Operating Expenses components. This fact can be

justified by its more random nature, as explained in Section 3, which results in increased modelling difficulty with statistical models. This conclusion can be corroborated by the fact that more components belonging to the W13 Operating Expenses class needed the highest number of clusters allowed to be formed.

To finalize, note that for all the optimal series created, none rejected the KS test, which gives assurance that they can be considered similar to the original series. In other words, they could be a plausible expense alternative series for the period 2016-2020.

5.2. Simulation Process

After the generation of the additional expense series, the next step was to perform an annual aggregation of the monthly values of each series component. The resulting distribution is shown in **Error! Reference source not found.** for *Accounting PT W12 – Personnel Expenses*. As it can be observed, the distribution is bimodal since the year 2020 was very atypical in terms of personal costs that decreased considerably.

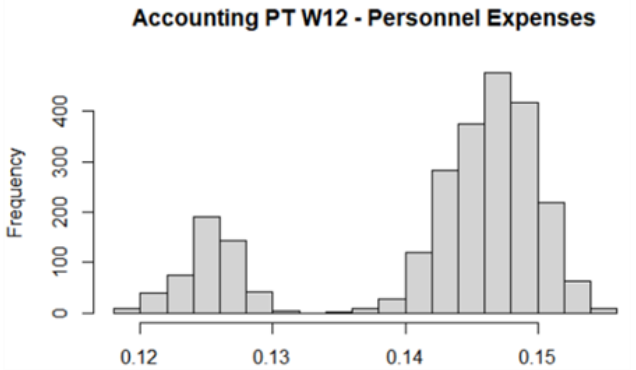
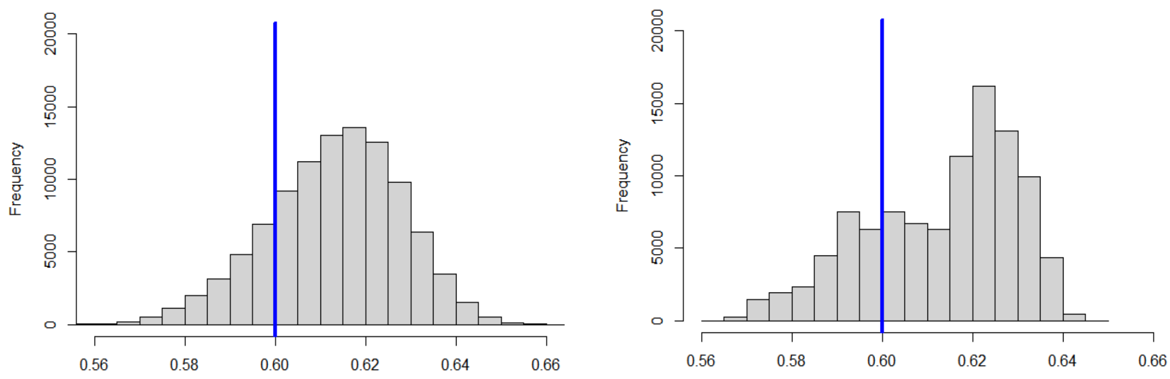


Figure 3: Annual Distribution obtained for Accounting PT W12 Personnel Expenses

From these distributions, a Monte Carlo Simulation with one hundred thousand iterations was performed, resulting in the distribution of the Total annual Expense shown in

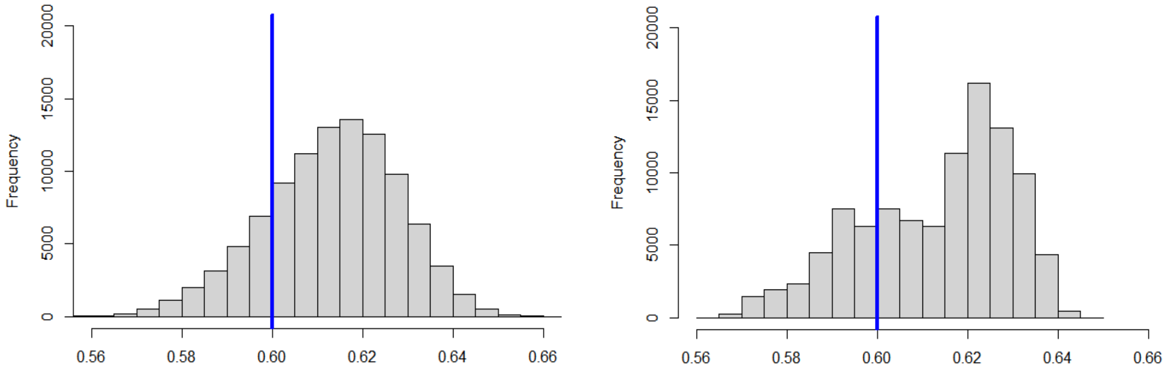


a) Independent Radom Sampling

b) Correlated Random Sampling

Figure 2. The one from the left (a) corresponds to when the independent random sampling between the components is considered, whilst the right (b) refers to the correlated one. The

blue vertical line depicts the total budgeted value that, for this example, is considered to be 0.60. **Table 5** summarizes the differences between the two distributions and, at the bottom, presents the risk metrics for the budget level of 0.60. **Figure 3** depicts the historical annual Spearman's rank correlations between the components.



a) Independent Radom Sampling b) Correlated Random Sampling

Figure 2: Total Expense Annual Distributions

Statistics	Random Sampling: Independent	Random Sampling: Correlated	Change (%)	
Mean	0.6131	0.6131	-0.0010	
Median	0.6142	0.6177	0.5762	
Mode	0.6080	0.6298	3.5969	
Standard Deviation	0.0147	0.0165	11.9565	
Skewness	-0.3221	-0.5656	75.5930	
Kurtosis	2.9086	2.3941	-17.6896	
Range = Max - Min	0.1132	0.0831	-26.6220	
95% prediction interval	[0.5818; 0.6393]	[0.5770; 0.6369]	-	
Value at Risk (5%)	0.6357	0.6348	-0.1323	
Cost of uncertainty	0.6183	0.6207	0.3967	
% Probability of expense exceed budget	81.2410	75.7410	-5.5000	pp
% Expected Deviation if budget exceeded	3.0472	3.4560	0.4088	pp
% Maximum Deviation if budget exceeded	11.1254	7.5280	-3.5974	pp

Table 5: Comparison of Total Expense Distribution Statistics with Independent and Correlated

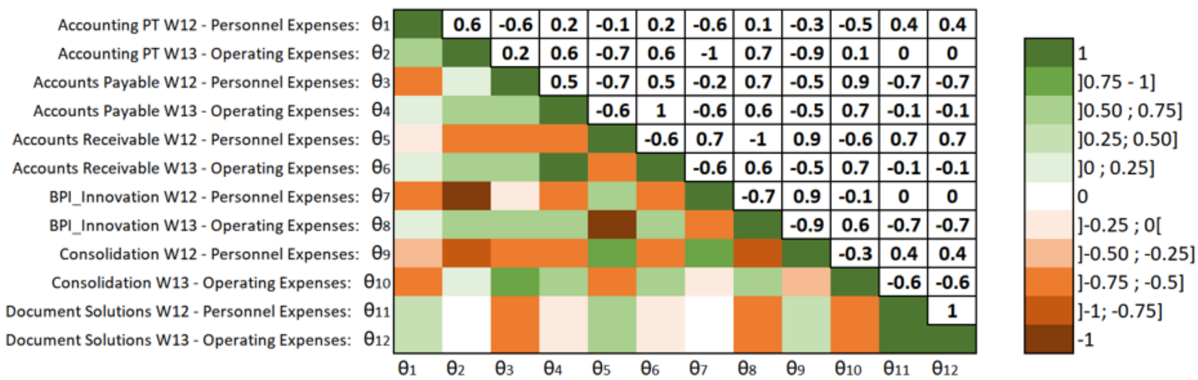


Figure 3: Spearman's Rank Correlation Coefficients

As it can be observed, the presence of correlations in the sampling does little change to the expected value of the Expense. Nonetheless, it affects considerably the spread of the distribution (characterized by the range) that decreased by 26%. This reduction in the tail values can also be seen because of a lower kurtosis value in the correlated distribution indicating the presence of less extreme values compared to its mean. The decrease in the spread can be justified by the presence of negative correlations between some components, corroborating the conclusion of Mun (2012) concerning the effect of correlations. Besides, as the components influence others, the Total Expense distribution does not center nicely in one value, presenting more uncertainty and variability when correlations are considered (higher standard deviation). Another impact of the correlations, in this case, is the shift of the expenses to the right, which is suggested by the skewness difference between the two distributions.

Note that by including correlations in the simulation (and mainly because of some negative coefficients), the probability of the Total Expense exceeding the budget decreases, which means that the same budgeted level would apparently have less risk of being surpassed. However, the caused shift of the distribution to the right led to an increase in the expected and maximum deviation from the budgeted level, making it necessary to consider the complete risk profile and not just probability to assess the risk of a budget.

5.3. Sensitivity Importance Measures

Now that the distribution of the Total Expense is fully characterized, it is useful to know what input Expense components influence this output distribution the most. For that, the Sensitivity Importance Measures total variance explained estimation ratio (S_i) and the moment independent index δ_i were implemented as described in Section 4. **Figure 4** shows the values obtained for each Expense component by the different measures tested, ordered by the rankings in the correlated random sampling version.

a) Si values b) δi values

Figure 4: Importance Measures Obtained in the Independent (in dark grey) and Correlated Random Sampling (in light grey) for the Different Expense Components, Ordered According to the Rankings in the Correlated Version

From the results, it can be observed that when independent random sampling is considered, the rankings given by the measures to the inputs converge in defining the most and least important input expense variables. Besides, as the rankings coincide with the δ_i, variance can be used to describe the effect on the outcome distribution. This happens because it directly reflects the model function decomposition and structure (E. Borgonovo 2006). Variance-based measures also provide guidance in determining the best components that can reduce total variance and, thus, are particularly useful as they single-handedly characterize variance and distribution effect when independence between inputs can be assumed. The three most influential components belong to the W12 Personnel Expenses class and concern the teams Accounting PT, Document Solutions and Accounts Receivable. Contrarily, the three least important ones belong to the W13 Operating Expenses class and refer to the teams' Accounts Payable, Consolidation and Accounts Receivable.

In comparison, when correlations are considered (and as they differ considerably from zero), the values from the computed measures are distinct. The variance explained ratios (S_i) increase because each input can now influence more of the output by changing other input variables. In fact, the explained variance of each input is mainly due to the presence of correlations. As far as the δ_i is concerned, its values are increased in the correlated analysis being in tune with the correlation study of E. Borgonovo and Tarantola (2008). This fact means that each input causes a bigger expected shift by influencing other variables.

The three most important components given by the δ_i are the Document Solutions team expenses and the *Consolidation W12 Personnel Expenses*. The ones with the least importance are the *Accounting PT W12 Personnel Expenses*, *Accounts Payable W13 Operating Expenses* and *Accounts Receivable W13 Operating Expenses*.

Note from **Figure 4** that the model inputs that influence variance the most are not necessarily the ones that impact more the distribution. However, the two most and least important variables are the same in both measures.

From these results, it means that addressing ways to control the expenses more, especially in the Document Solutions team, may potentially reduce the Total Expense variance and positively influence the Total Expense distribution. Therefore, a joint utilization of variance-based techniques and the δ_i allows one to better characterize input importance and comprehend its uncertainty propagation.

6. Conclusions and Future Work

In order to create accurate budgets, companies need to consider uncertainties when budgeting. This means that they have to acknowledge that the budget compliance will be subject to uncertain factors that may influence the company's performance. For this matter, Simulation is a valuable tool that enables managers to assess budget risk by comparing a created distribution with its budgeted value. This enables managers to consider the contingency applied to the budget and the risk of a cost overrun before its execution.

For creating similar expense series through time series bootstrapping techniques, we concluded that there is no universal best technique that creates the most similar series. However, the K-Means Based Bootstrap is very competent for the majority of the cases. We also verified that techniques relying on bootstrapping the residuals adapt better to only bootstrapping blocks of data (as it happens in the MBB). Additionally, our results show that the choice of the bootstrap method is more important than its parameterization (e.g., changing block length).

Furthermore, we concluded that the order of importance of the input Expense components provides valuable information for the manager because by controlling the most important Expense components, he can reduce uncertainty's impact on Total Expenses. For that, δ_i is the preferred measure to identify the most important input variables since it considers how the output distribution shifts when fixing the input variable (E. Borgonovo 2007). When the goal is to identify the most relevant Expense components in reducing Total Expense variability, variance-based measures (such as S_i) should be used.

For future work, the presented methodology should be extended to overcome some of its current limitations. A relevant one is that this current risk analysis is very dependent on historical data, which may be restrictive when the future strategy of the division is considerably different from its past. For example, if the division is not in a balanced expense plateau, registering exponential growth every year, this risk analysis performed would not be very insightful as it would always indicate a high risk for the new budget. Therefore, one can obtain the Expense components' distribution more adapted to the division's future strategy. For this, the usage of Machine Learning algorithms with explanatory variables could be explored to create the alternative series relative to the future period when the budget is being assessed. This would substitute the Bootstrapping procedure of this methodology, but the other steps would be the same. Another relevant aspect that should be considered is the reduced amount of data used to estimate the correlation coefficients (i.e., only used five years). To overcome this issue and be more confident that the real relationships between the components were captured, one could combine it with interviews with some key people of the organization or use more data if available.

References

Bakhshi, Payam, and Ali Touran. 2014. 'An Overview of Budget Contingency Calculation Methods in Construction Industry'. Selected Papers from Creative Construction Conference 2014 85 (January): 52–60. https://doi.org/10.1016/j.proeng.2014.10.528.

Bergmeir, Christoph, Rob J. Hyndman, and José M. Benítez. 2016. 'Bagging Exponential Smoothing Methods Using STL Decomposition and Box–Cox Transformation'. International Journal of Forecasting 32 (2): 303–12. https://doi.org/10.1016/j.ijforecast.2015.07.002.

Borgonovo, E. 2007. 'A New Uncertainty Importance Measure'. Reliability Engineering & System Safety 92 (6): 771–84. https://doi.org/10.1016/j.ress.2006.04.015.

Borgonovo, E., and S. Tarantola. 2008. 'Moment Independent and Variance-Based Sensitivity Analysis with Correlations: An Application to the Stability of a Chemical Reactor'. International Journal of Chemical Kinetics 40 (11): 687–98. https://doi.org/10.1002/kin.20368.

Borgonovo, Emanuele. 2006. 'Measuring Uncertainty Importance: Investigation and Comparison of Alternative Approaches'. Risk Analysis 26 (5): 1349–61. https://doi.org/10.1111/j.1539-6924.2006.00806.x.

Borgonovo, Emanuele, and Elmar Plischke. 2016. 'Sensitivity Analysis: A Review of Recent Advances'. European Journal of Operational Research 248 (3): 869–87. https://doi.org/10.1016/j.ejor.2015.06.032.

Collier, Paul M., and Anthony J. Berry. 2002. 'Risk in the Process of Budgeting'. Management Accounting Research 13 (3): 273–97. https://doi.org/10.1006/mare.2002.0190.

Cordeiro, Clara, and M. Manuela Neves. 2009. 'Forecasting Time Series with Boot.EXPOS Procedure'. REVSTAT-Statistical Journal, July, 135-149 pages. https://doi.org/10.57805/REVSTAT.V7I2.78.

Eaton, Graham. 2005. Management Accounting Official Terminology. 2nd ed. CIMA Official Terminology. Oxford: CIMA Publishing. ISBN 978-0-7506-6827-9.

Efron, B. 1979. 'Bootstrap Methods: Another Look at the Jackknife'. The Annals of Statistics 7 (1). https://doi.org/10.1214/aos/1176344552.

Elmassri, Moataz, and Elaine Harris. 2011. 'Rethinking Budgetary Slack as Budget Risk Management'. Edited by Julia Mundy. Journal of Applied Accounting Research 12 (3): 278–93. https://doi.org/10.1108/09675421111187700.

Hager, J., V.S.S. Yadavalli, and R.C.W. Webber-Youngman. 2015. 'Stochastic Simulation for Budget Prediction for Large Surface Mines in the South African Mining Industry'. Journal of the Southern African Institute of Mining and Metallurgy 115 (June): 531–39.

Hao, Wenrui, Zhenzhou Lu, Pengfei Wei, Jun Feng, and Bintuan Wang. 2012. 'A New Method on ANN for Variance Based Importance Measure Analysis of Correlated Input Variables'. Structural Safety 38 (September): 56–63. https://doi.org/10.1016/j.strusafe.2012.02.003.

Lahiri, S. N. 2003. Resampling Methods for Dependent Data. Springer Series in Statistics. New York, NY: Springer New York. https://doi.org/10.1007/978-1-4757-3803-2.

Laurinec, Peter, Marek Lóderer, Mária Lucká, and Viera Rozinajová. 2019. 'Density-Based Unsupervised Ensemble Learning Methods for Time Series Forecasting of Aggregated or Clustered Electricity Consumption'. Journal of Intelligent Information Systems 53 (2): 219–39. https://doi.org/10.1007/s10844-019-00550-3.

LePage, Raoul, L. Billard, eds. 1992. Exploring the Limits of Bootstrap. New York: Wiley. ISBN 978-0-471-53631-4.

Libby, Theresa, and R. Murray Lindsay. 2010. 'Beyond Budgeting or Budgeting Reconsidered? A Survey of North-American Budgeting Practice'. Management Accounting Research 21 (1): 56–75. https://doi.org/10.1016/j.mar.2009.10.003.

Lidia, Tănase Gabriela. 2014. 'Difficulties of the Budgeting Process and Factors Leading to the Decision to Implement This Management Tool'. Procedia Economics and Finance 15: 466–73. https://doi.org/10.1016/S2212-5671(14)00484-5.

Lord, Robert J. 1977. 'Probabilistic Budgeting: One Practical Experience'. In Proceedings of the 9th Conference on Winter Simulation - Volume 2, 608–16. WSC '77. Winter Simulation Conference.

———. 1979. 'Coping with Future Uncertainties through Probabilistic Budgeting'. Developments in Business Simulation and Experiential Learning: Proceedings of the Annual ABSEL Conference 6 (March). https://absel-ojs-ttu.tdl.org/absel/article/view/2581.

Moro Visconti, Roberto, Giuseppe Montesi, and Giovanni Papiro. 2018. 'Big Data-Driven Stochastic Business Planning and Corporate Valuation'. Corporate Ownership and Control 15 (3–1): 189–204. https://doi.org/10.22495/cocv15i3c1p4.

Mun, Johnathan, ed. 2012. Modeling Risk: Applying Monte Carlo Risk Simulation, Strategic Real Options, Stochastic Forecasting, and Portfolio Optimization. Hoboken, NJ, USA: John Wiley & Sons, Inc. https://doi.org/10.1002/9781118366332.

Neine, Mohamed, and Desmond Curran. 2021. 'An Algorithm to Generate Correlated Input-Parameters to Be Used in Probabilistic Sensitivity Analyses'. Journal of Market Access & Health Policy 9 (1): 1857052. https://doi.org/10.1080/20016689.2020.1857052.

Paté-Cornell, M. Elisabeth, and Robin L. Dillon. 2006. 'The Respective Roles of Risk and Decision Analyses in Decision Support'. Decision Analysis 3 (4): 220–32. https://doi.org/10.1287/deca.1060.0077.

Platon, Victor, and Andreea Constantinescu. 2014. 'Monte Carlo Method in Risk Analysis for Investment Projects'. Procedia Economics and Finance 15: 393–400. https://doi.org/10.1016/S2212-5671(14)00463-8.

Radovanov, Boris, and Aleksandra Marcikić. 2014. 'A Comparison of Four Different Block Bootstrap Methods'. Croatian Operational Research Review 5 (2): 189–202. https://doi.org/10.17535/crorr.2014.0007.

Robinson, Anis Chariri, and Tri Jatmiko Wahyu Prabowo. 2018. 'The Role of Risk Management on Budget Quality and Slack'. Academy of Accounting and Financial Studies Journal 22 (4): 1–12.

Rubin, Geoffrey D., and Bhavik N. Patel. 2017. 'Financial Forecasting and Stochastic Modeling: Predicting the Impact of Business Decisions'. Radiology 283 (2): 342–58. https://doi.org/10.1148/radiol.2017161800.

Saltelli, Andrea. 2002. 'Sensitivity Analysis for Importance Assessment'. Risk Analysis 22 (3): 579–90. https://doi.org/10.1111/0272-4332.00040.

Scott, W. Richard. 1998. Organizations: Rational, Natural, and Open Systems. 4th ed. Upper Saddle River, N.J: Prentice Hall.

Shim, Jae K., and Joel G. Siegel. 2005. Budgeting Basics and Beyond. 2nd ed. Hoboken, N.J: J. Wiley & Sons.

Urgilés, Paúl, Juan Claver, and Miguel A. Sebastián. 2019. 'Methods for Quantitative Risks Analysis of Cost and Deadline Overruns in Complex Projects'. Procedia Manufacturing 41: 658–65. https://doi.org/10.1016/j.promfg.2019.09.055.

'Using Monte Carlo Simulation for Pro Forma Financial Statements'. 2019. Journal of Accounting and Finance 19 (5). https://doi.org/10.33423/jaf.v19i5.2248.

Xu, Chonggang, and George Zdzislaw Gertner. 2008. 'Uncertainty and Sensitivity Analysis for Models with Correlated Parameters'. Reliability Engineering & System Safety 93 (10): 1563–73. https://doi.org/10.1016/j.ress.2007.06.003.

Acknowledgments

We would like to thank the retailer Sonae MC for providing the data for this study and the anonymous reviewers that helped turn this paper more insightful and organized.