Computers & Industrial Engineering 112 (2017) 71–83

Contents lists available at ScienceDirect

Computers & Industrial Engineering

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

Use of control charts with regression analysis for autocorrelated data in the context of logistic financial budgeting

Jorge Pérez-Rave [a,*], Leandro Muñoz-Giraldo [a], Juan Carlos Correa-Morales [b]

[a] Grupo de investigación IDINNOV, Medellín, Colombia
[b] Universidad Nacional de Colombia, Escuela de Estadística, Medellín, Colombia

ARTICLE INFO

Article history:
Received 8 April 2017
Received in revised form 10 August 2017
Accepted 12 August 2017
Available online 14 August 2017

Keywords:
Control charts
Financial variables
Regression analysis
Residual control charts
Logistics
Autocorrelation

ABSTRACT

The aim of this paper is to explore whether the use of control charts with regression analysis is an effective way to evaluate financial budget requests (autocorrelated data) in the transport logistics sector. First, the variables are selected. Second, a regression analysis is performed to model the financial variables. Third, three types of traditional control charts are tested (individuals, CUSUM and EWMA), using simulation to monitor the regression scaled residuals. The results show that the individual control chart of 2.7-sigma offers an appropriate performance for the context of this study. This paper provides new evidence regarding a type of variable and context not reported in the literature. In addition, it proposes a control chart approach of scaled regression residuals, with two differentiators: (1) residuals offer better practical interpretation and (2) regressions do not incorporate the time variable, as traditionally occurs, but a missionary process variable (loading units) and a control one.

1. Introduction

The financial areas of companies are responsible for planning and controlling budget. During financial control, a specific area of a given company could requests amounts allocated them to be greater than or less than the actual budget. This discrepancy has tended to be solved on the basis of experience alone, as revealed by the historical mean of financial accounts. However, unless process variability is considered, it is objectively impossible to differentiate between common and "special" budget requests. Therefore, it is necessary to monitor, understand and control the variability of financial processes, in order not to affect the normal operation of companies.

Understanding and controlling the variability of financial processes could be achieved using control charts, which have been widely and successfully used in the manufacturing sector (Nair, 2006). In fact, though still in its infancy, recent studies have shown the application of control charts to financial variables to be successful (Aliverdi, Moslemi Naeni, & Salehipour, 2013; Dimaki, Psarakis, Virras, & Angelis, 2011; Dull & Tegarden, 2004; Kovářík & Sarga, 2014; Kovářík, Sarga, & Klímek, 2015; Krehbiel, Havelka, & Scharfenort, 2007; Marrero, Fonseca, Falcon, & Ramirez-De-Arellano, 2014).

Financial variables tend to emulate how non-financial processes change through time (Dull & Tegarden, 2004) and present the following properties: autocorrelation, non-normality, data generation delay (monthly, quarterly, etc.). Autocorrelation reduces the performance of control charts (Kovářík & Sarga, 2014). Non-normality can introduce inadequate control limits (false alarms most frequently), especially in the case of the individual control chart (Montgomery, 2013). Data generation delay increases the distance between causes and effects. Consequently, timely detection of changes in process becomes difficult.

One approach to overcome autocorrelation is the use of time series (Kovářík & Sarga, 2014; Kovářík et al., 2015; Montgomery, Jennings, & Kulahci, 2008; Woodall & practice in basic statistical process monitoriPlease update reference 'Woodall, 2017). However, the time variable is a "black box" in that it precludes exploring the reasons for the changes in the financial variable, through the use of non-financial process variables as predictors. These reasons may be useful during a cause analysis stage. Hence, developing models that establish a link between specific financial variables and variables from the operational processes is desired in this context. A variant approach is that of regression analysis, which can yield predictors, other than the variable time, that correlate significantly with the response variable and help predict its behaviour. However, this approach has been used less because of the

* Corresponding author at: Cll. 55 # 46 – 14, Of. 706, Medellín, Colombia.
E-mail addresses: investigacion@idinnov.com (J. Pérez-Rave), investigacion@idinnov.com (L. Muñoz-Giraldo), jccorrea@unal.edu.co (J.C. Correa-Morales).

http://dx.doi.org/10.1016/j.cie.2017.08.015
0360-8352/

complexities and the cost of discovering such predictors (deepening other process metrics, collecting new data, solving possible problems of collinearity, etc.), especially in the financial sphere. Additionally, under the time series or regression approach, in the financial context, it is customary to monitor residuals (estimated prediction errors) to mitigate, among others, non-normality and autocorrelation problems. Nevertheless, the way in which the residuals are expressed, generally lacks of practical interpretational value for practitioners (Mastrangelo & Montgomery, 1995). This increases the gap between theory and practice (Jarrett & Pan, 2009; Woodall, 2017). On the other hand, in order to address data generation delay (e.g., weekly, monthly, etc.), which is intrinsic to certain processes, like the financial kind, it is useful to use individual charts, CUSUM and EWMA (Kovářík & Sarga, 2014; Kovářík et al., 2015; Mohamed, 2010) to alert about possible special causes.

Currently there is not enough data in terms of the usage of control charts with financial variables to fully understand its complexities, particularities, and variations. However, previous studies suggest that financial variables be considered in novel contexts (countries, industries and processes). According to Frisén (2011), industrial applications dominated the development of process control (surveillance), but in recent years, the need for applications in finance (a complex context) has increased. In short, new strategies to address such financial variable properties as autocorrelation, non-normality, and data generation delay, as well as the emulation of changes from non-financials processes, are necessary. In fact, there are no previous studies that address financial budget variables in the transport logistics sector from the perspective of control charts.

The aim of this study is to explore whether the use of control charts with regression analysis is effective in evaluating financial budget requests in the transport logistics sector. Data from the financial variables (cost of fuel refill, vehicle repair and auto part replacement, for local routes), and the loading units were provided by a private logistics transport company (courier and parcel) in Colombia. Unlike the traditional time-series approach in the study of financial metrics, this paper adopts a look at regression analysis with control variables and a predictor of the mission-logistic operation (loading unit), which is vital for decision-making in the logistics sector. This type of approach is requested in the literature, in order to understand the "black box" that underlies the time-series perspective and, thus, to facilitate the way to identify possible special causes. It is about taking advantage of the information of auxiliary variables (Ex: unit of load) to improve the explanation of the variable of interest. These auxiliary variables are sought to be related to the variable to be monitored, but doing it in such a way that an eventual shift in the latter does not affect the former (Ahmad, Abbasi, Riaz, & Abbas, 2014). Likewise, residuals will be monitored by means of control charts, which, unlike the traditional approach, add practical interpretational value for professionals.

This work becomes relevant, as it fits in with the customary classification of the research on control charts according to their applications (Doroudyan, Owlia, Sadeghi, & Amiri, 2017; Frisén, 2011; Kovářík et al., 2015). Also, in the financial scenario, most control chart studies have addressed macroeconomic metrics and little attention has been given to internal corporate performance metrics, and even less to the financial-logistical context. However, each variable, within a given context, presents particularities that make it unique. That is, the design and development of a control chart is not part of a universal approach. Therefore, it is necessary to understand the properties of the variable and to fit appropriate models from an initial data set, before monitoring the residuals, which in itself is recognized as a complex task in financial data analysis (Doroudyan et al., 2017; Frisén, 2011; Kovárík & Klímek, 2012; Kovářík et al., 2015). Another inherent complexity in monitoring residuals is choosing the appropriate control chart for each specific case (Kovárík & Klímek, 2012; Žmuk, 2016), hence the invitation to design control charts for individual observations in different applications (Kovárík & Klímek, 2012).

This study, in general, demonstrates that it is feasible to use control charts with regression analysis to monitor and control internal financial variables of the logistic context, which has not been proven in previous studies. Specifically, this work makes three contributions: First, it provides new evidence from both a variable and a context not addressed in previous studies: financial budget requests in the transport logistics sector (properties of the data, behaviour of three types of traditional charts using logistics financial data, and type and parameters of the most appropriate chart). Secondly, this work describes a procedure based on scaled residuals, from two regression models elaborated from an initial data set. Each model establishes links between financial variables and both the loading unit (variable of the logistics operation) and a control variable. This procedure departs from the traditional ones, based on time series. This becomes relevant to the logistics context since it uses the loading unit as one of the predictors. Moreover, the literature points out the value of newly models, yet warns about the complexity of arriving at them in the financial arena. Thirdly, this study uses scaled residuals, which carries practical meaning for practitioners, as these are expressed as the percentage difference between the real value and the predicted value with respect to the predicted value. This differs from the traditional approach, in which the residuals lack practical interpretational value, and where their use and maintenance is difficult.

This paper is divided into five sections. Section 1 presents the introduction to this study. Section 2 summarizes its main concepts, findings, lessons, and future research, as extracted from previous studies. Section 3 describes the procedure, which consists of three stages: choice of variables, regression analysis and application of control charts. Section 4 presents the results, after the procedure. Section 5 synthesizes the discussion engendered by these results.

2. Literature review

A thorough discussion regarding aspects of differentiation between Statistical Process Control theory and practice can be found in Woodall (2017). Here, it is worth mentioning four of these aspects: First, control charts are increasingly useful to address problems in such fields as health and networks, among others. However, more collaboration is needed between practitioners and researchers, in order to reduce the inherent gap in interaction. Second, the autocorrelation effect on the performance of control charts is a topic of currency, considering that now data are less time-spaced. Third, process quality can be better measured using a function that links one dependent variable to its predictors. And fourth, the identification of the root causes of a "special event" is a challenge from the standpoint of Statistical Process Control (Woodall, 2017).

While exploring the available literature on the use of control charts for financial variables, several queries were carried out on Scopus (one of the largest scientific databases around the world). Initially, the following query is performed (27/01/2017): [*TITLE ("control charts" OR "statistical process control" OR "statistical control" OR "control limits")*]. This query reports 5015 documents. Then, some terms about financial variables are added *[TITLE(Budget OR Fund OR Quote OR Finance OR Financial OR Funding OR cost)]*. However, the documents reported are very few: only 50 documents (1%). After verifying these documents, only 13 of them use control charts for financial variables (0.26% of 5015 documents). The rest of the documents study theoretical aspects or computational costs of control charts. The above queries are repeated on