

Analysis and forecast of Brazilian's ISO 9001 certifications and its relationship with economic and social indicators

Martins YS¹, da Silva CES, Balestrassi PP, Gaudencio JHD

Abstract With the growing demand for ISO certifications, the number of studies analyzing the evolution of such certifications over time in the most diverse contexts and countries has also increased. There are several forecasting methods, used as a way to verify the behaviour of certain variables in the future, given their history. In this context, this study aims to analyze the number of ISO 9001 certifications from Brazil and their relationship with the variation of the indicators Gross Domestic Product (GDP) and Human Development Index (HDI). The method used was the regression analysis, through which a model was identified and adjusted to 75% of the series data, and its validation through the forecast of the remaining 25%. The model presented adequacy to the data for both variables, characterizing a multiple regression, which quadratic model $Y = -268.3 + 0.000006X_1 + 782X_2 - 553.1X_2^2$ has an adjusted R² of 94.07%. The model validation analyzed from the forecast indicates that it is good for predicting, even with the existing variation between 2017 and 2019. From the forecasting model and the response optimization analysis, it can be concluded that the number of ISO 9001 certifications in Brazil should be at least 25% higher, given its economic, social and demographic conditions, especially if compared to other Latin American countries.

Keywords: Quality Management System, Regression Analysis, Gross Domestic Product, Human Development Index.

1 Introduction

The search for excellence related to quality management practices, leads to an increasing number of companies seeking certification in ISO 9001 (Fonseca, 2015; ISO, 2020). In this context, several authors have developed research on the future of ISO certifications (Ikram, Zhang and Sroufe, 2020; Sampaio, Saraiva and Rodrigues, 2011) considering integrated management systems (Cabecinhas et al, 2018) and also analyzing the effects of certifications on macroeconomic and socio-cultural factors of countries (Ribeiro et al, 2019; Salgado et al, 2015) through forecasting models based on ISO Survey data.

Predicting future events is a critical activity in the most varied planning and decision-making processes, which is why forecasting is so important and is applied in many areas (Montgomery, Jennings and Kulahci, 2008). Forecasting techniques can be both qualitative (Delphi method) and quantitative (regression models, smoothing models and time series analysis).

Therefore, this article has as main objective to analyze the impact of the variation in the GDP and HDI indicators from Brazil, on the number of ISO 9001 certifications. In order to do this, a multiple regression analysis will be performed, with the support of Minitab® 18, identifying the model that better adjusts to the data and then making the forecast for the data series. It is important to highlight that although it is considered a limitation of the research, the number of observations used in the analysis considers all the historical available for the number of ISO 9001 certifications, which begins in 1993. Through a time series

¹Yasmin Silva Martins (e-mail: yasminsm@unifei.edu.br)

Institute of Production Engineering and Management. Federal University of Itajuba. BPS Avenue, 1303, Pinheirinho, Itajuba-MG, Brazil.

plot it is possible to analyze the behaviour of the data observed on the ISO 9001, which presented a growing trend in the number of Brazilian certifications until 2012, with a slight decline since then (see Fig. 1).

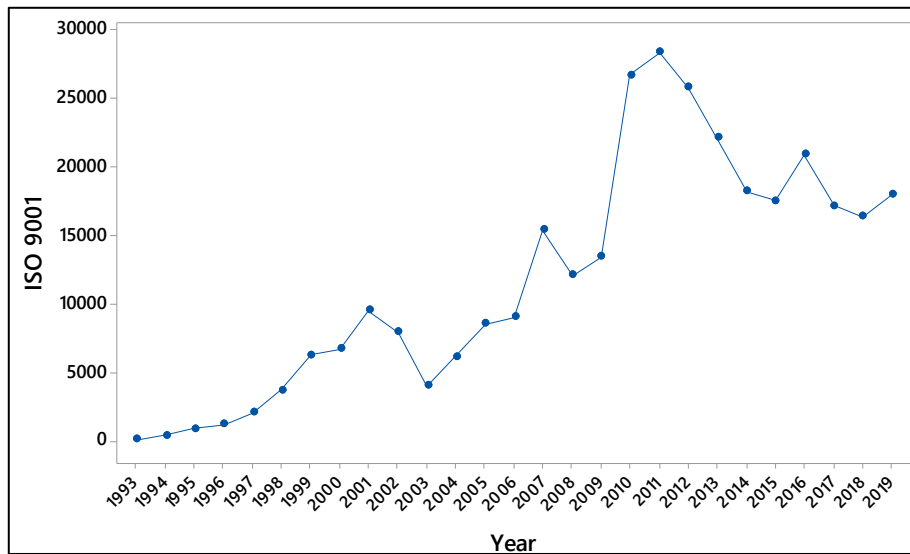


Fig. 1 Times series plot of ISO 9001

2 Research method

2.1 Regression Analysis

According to Chatterjee and Hadi (2006), regression analysis is a method of investigating relationships between response (or dependent) variables and predictors (or explanatory) variables, capable of being applied in the most varied areas of study. It is one of the most widely used statistical tools as it provides a set of simple methods to establish a useful relationship between variables (Montgomery, Jennings and Kulahci 2008; Chatterjee and Hadi, 2006).

Models with a single response variable (Y) and two or more predictor variables (X1, X2, ..., Xn), are called multiple regression models; while models with a single predictor variable are simple regression models; both can be linear or nonlinear. Regression analysis involves the following steps that are considered in this research: statement of the problem, selection of potentially relevant variables and data collection (section 2.2); model specification and fitting (section 2.3); model validation (section 2.4).

2.2 Variables definition and data collection

As mentioned in the introduction, this research problem is based on the following question: "does the variation in economic and social indicators in Brazil impact the number of ISO 9001 certifications from the country?". In order to answer this question and conduct the analysis, the following variables were selected: number of ISO 9001 certifications, as response variable Y [2]; Gross Domestic Product (GDP), as predictor variable X1 (IPEADATA, 2020); and Human Development Index (HDI), as predictor variable X2 (PNUD, 2019; Klafke, 2016). The selection of indicators was based on their relationship with quality, once GDP is related to the country's economy and HDI measures aspects linked to income, education and health, which in turn have an impact on the economy, enterprises and quality. In addition, the availability of historical

data series through reliable sources was also considered. Data were collected between 1993 and 2019 and 21 observations were used to model construction, representing 75% of the data (see Table 1) while 6 observations (25% remaining) were used to forecast and validate the model (described in section 2.4).

Table 1 Historical data of the model construction variables.

Period	Observations	Y: ISO 9001	X_1^1 : GDP	X_2 : HDI
1993	1	113	3,534.9	0.632
1994	2	384	44,210.2	0.640
1995	3	923	66,577.7	0.648
1996	4	1,198	81,667.1	0.655
1997	5	2,068	81,216.8	0.662
1998	6	3,712	83,051.8	0.669
1999	7	6,257	98,772.5	0.676
2000	8	6,719	105,851.6	0.683
2001	9	9,489	113,016.7	0.686
2002	10	7,900	130,241.2	0.690
2003	11	4,012	153,801.4	0.694
2004	12	6,120	178,462.4	0.698
2005	13	8,533	198,480.0	0.702
2006	14	9,014	221,359.3	0.709
2007	15	15,384	242,460.2	0.716
2008	16	12,057	264,404.8	0.723
2009	17	13,452	313,522.8	0.730
2010	18	26,663	355,797.4	0.726
2011	19	28,325	391,595.1	0.718
2012	20	25,791	423,195.9	0.742
2013	21	22,128	473,552.5	0.752

.¹Original values of predictor variable X_1 in millions

2.3 Model specification and fitting

The data were analyzed using the software Minitab® 18, following the sequence of steps previously described (Chatterjee and Hadi, 2006). The relationship established in multiple regression analysis generally uses the method of least squares, that is, to minimize the sum of squares of the errors. A satisfactory estimated model satisfies some assumptions for the variables, the model and the errors: these, specifically, must be normal, independent (not autocorrelated) and stationary (mean zero and constant variance); in addition, the existence of multicollinearity (dependence between the predictor variables) can cause serious effects on the parameter's estimation and the model fitting.

At first, a 3D scatter plot was prepared for the series variables in order to identify the data behaviour (see Fig. 2). Then, a quadratic multiple regression equation was estimated for Y as a function of X1 and X2, obtaining an adjustment of 75.34% for predicted R2 and 87.84% for adjusted R2. However, despite the reasonable fit, the analysis of the predictor variables does not indicate a good adjustment, due to the high p-values obtained. Thus, after attempts to adjust the quadratic model by removing X1*X1 and X1*X2 interactions once they were not significant to the model, and performing a Box Cox transformation, an equation of multiple quadratic regression of Y as a function of X1 and X2 was estimated.

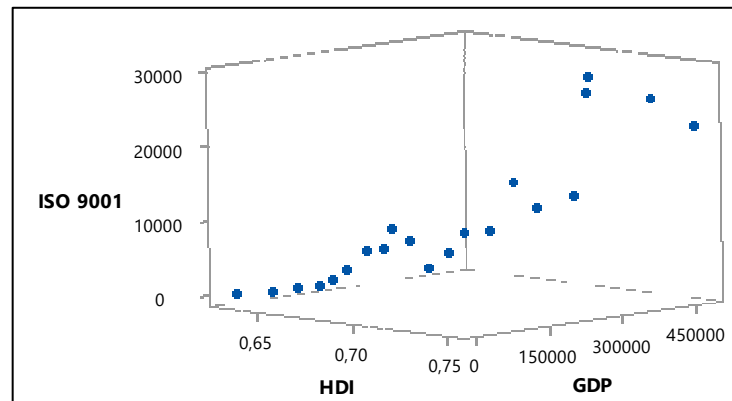


Fig. 2 3D scatter plot for ISO 9001 vs GDP vs HDI

Considering a Box Cox transformation with natural log ($\lambda = 0$), the equation (1) presents a good fit for the quadratic model, with adjusted R2 of 94.07% and predicted R2 of 92.26%. The p-values obtained in ANOVA are less than 0.05, indicating that the model is good for forecasting and that the GDP and HDI variables both explain the model; also, the t and F tests presented significant values. In addition, the normality test was performed for the new residuals, verifying that they follow a normal distribution with p-value 0.870 and mean about equal to zero (see Fig. 3) and the autocorrelation test did not indicate significant autocorrelations for the residuals.

$$Y = -268.3 + 0.000006X_1 + 782X_2 - 553.1X_2 * X_2 \quad (1)$$

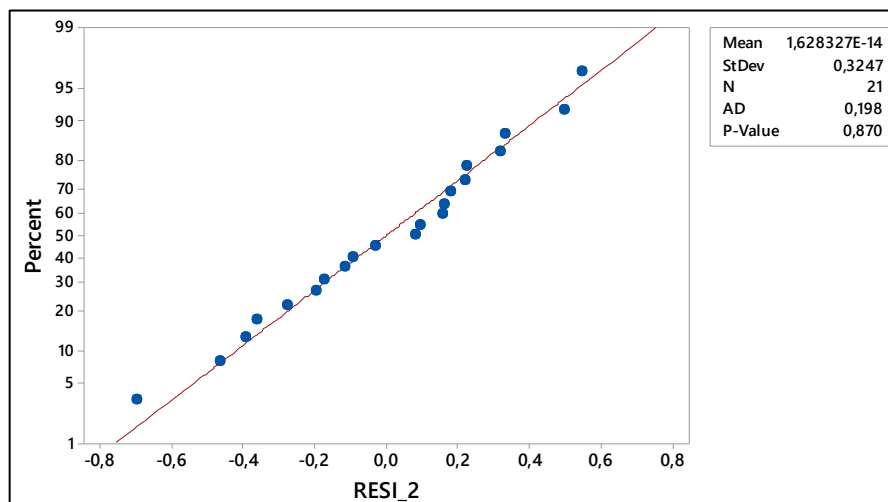


Fig. 3 Probability plot of residuals X1

2.4 Model validation

With the fitted model it is possible to forecast the number of ISO 9001 certifications. In order to validate the model and present the results of its application, 25% of the collected data (observations from 2014-2019) were considered. Using the multiple regression model obtained in the adjustment, the 6 observations were predicted, resulting in the forecast values (PFIT), confidence (CI) and prediction (PI) intervals, presented in Table 2.

Table 2 Prediction values for the estimated regression model

Period	GDP	HDI	Real	<i>PFITs</i>	CI	PI
2014	499,867.7	0.754	18,196	18,583	(10,827.6; 31,893.9)	(7,415.8; 46,567.6)
2015	521,918.7	0.755	17,529	19,982	(11,401.2; 35,020.5)	(7,875.1; 50,700.9)
2016	565,780.5	0.757	20,908	22,995	(12,250.3; 43,163.9)	(8,682.1; 60,903.4)
2017	588,892.8	0.760	17,165	22,099	(11,165.9; 43,737.5)	(8,056.6; 60,617.5)
2018	603,155.2	0.761	16,351	22,585	(11,083.1; 46,023.0)	(8,070.8; 63,200.0)
2019	639,806.5	0.761	17,972	27,805	(12,538.5; 61,658.1)	(9,355.7; 82,634.1)

Through the information from the table above it is possible to analyze data, leading to the identification of some important issues. Considering that the good fit of the model is related to its forecasting ability, satisfactory models tend to generate standard errors closer to zero, approximating the forecasted values to the actual observed ones (Montgomery, Jennings and Kulahci, 2008). Fig. 4 allows to see a comparison between *PFITs* and observed real values, indicating that from 2014 to 2016 there was a small variation, that increased from 2017 on. However, it does not mean that the model is not good for predictions. More variables to be analyzed would make it a more robust model. It is possible to analyze from the figure that the HDI and GDP trend in Brazil would not guarantee a drop in the number of certifications, but external factors such as politics, international economy and other aspects, have provided a drop in the number of certifications in recent years, so this divergence from 2017 onwards.

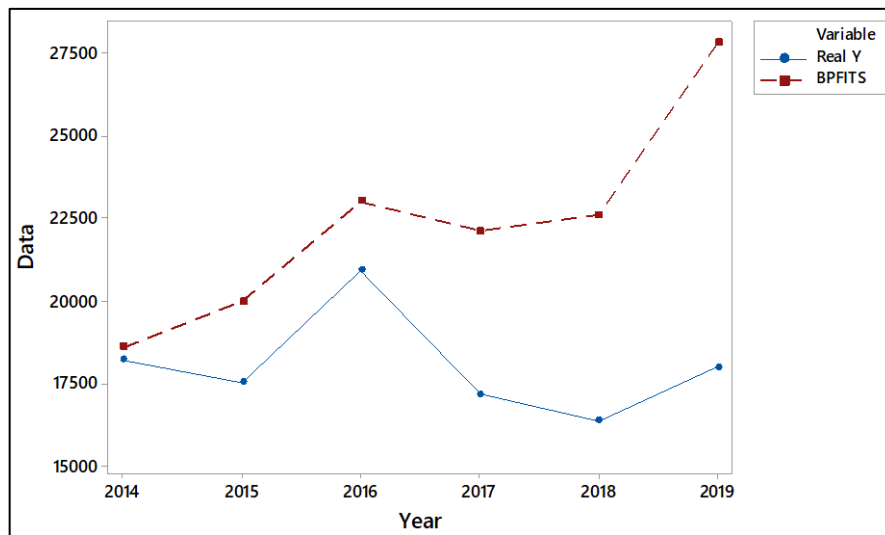


Fig. 4 Time series plot of Real Y (ISO 9001 observations); *PFITs* (predicted values).

In order to compare the results and develop more consistent conclusions under Brazilian’s number of certifications, an optimization was performed as described in the following item.

2.5 Optimal number of certifications in Brazil

Based on the available data from the databases considered in this research, the authors investigated countries whose characteristics, indicators and certification numbers were comparable to those of Brazil.

Brazil is the country with the highest number of ISO 9001 certifications among Latin American countries, where Colombia is ranked in second place (ISO, 2020). With a population four times smaller than Brazil, Colombia has an HDI rating of 0.689 and a per capita GDP of \$6,498 (PNUD, 2019). Still, 87% of all companies throughout the country are ISO 9001 certified, while in Brazil this rate is about 66%. In this context, through a response optimization analysis performed with Minitab® 18, it was verified which

values of GDP and HDI would give a target of 23,836 certifications, representing 87% of the listed companies in Brazil.

The results found in Fig. 5 indicate that with values far below the current indicators (solution: GDP = 473,552.5 and HDI = 0.668), Brazil would reach 87% of its companies certified, as identified in Colombia. With a GDP of 639,806.5 and HDI at 0.761 (2019 results), the optimal number of certifications in Brazil should be about 27,805, forecast obtained through the regression model (see Table 2).

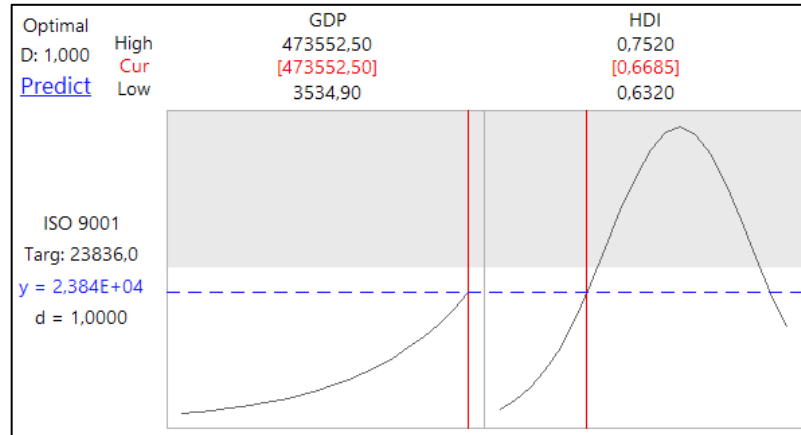


Fig. 5 Optimization plot for a target of 23,836 certifications

3 Conclusion

It was verified through multiple regression analysis and adjustments attempts, that both GDP and HDI have effects on the number of ISO 9001 certifications in Brazil, thus generating a multiple quadratic regression model. With an explanation of approximately 94% the model indicates that the variation of a unit in the GDP tends to increase by 0.000006 the number of ISO 9001 certifications, while a variation in HDI increases by 782 at the same time that its second order interaction (HDI*HDI) decreases by 553.1 the number of certifications. By obtaining an adjusted model it was possible to forecast 6 observations of the data set, through which was identified a variation between the actual and forecasted values of the last three years, that may be caused by external factors such as politics. Also, incorporating more observations into the forecast can strengthen the model.

Due to the Brazilian GDP and HDI of 2019, the analysis suggests that the number of certifications in the ISO 9001 should be at least 25% higher. This shows that, with the conditions and dimension of Brazil, there should be more certified companies in the country, especially when compared to the ratio of certifications by number of companies of Colombia. The authors suggest as future research the analysis of prediction models for certifications related to human and political factors, which also have great influence on the economy, especially where companies operate on the stock market.

4 References

- Cabecinhas M, Domingues P, Sampaio PACA, Bernardo M, Franceschini F, Galetto M, Gianni M, Gotzamani K, Mastrogiacomio L, Hernandez-Vivanco A (2018) Integrated Management Systems Diffusion Models in South European Countries. *International Journal of Quality & Reliability Management*, doi: 10.1108/IJQRM-03-2017-0044
- Chatterjee S, Hadi AS (2006) *Regression Analysis by Example*. Wiley
- Fonseca LM (2015) From Quality Gurus and TQM to ISO 9001:2015: A review of several quality Paths. *International Journal for Quality Research* 9(1), 167–180
- Ikram M, Zhang Q, Sroufe R (2020) Future of quality management system (ISO 9001) certification: novel grey forecasting approach. *Total Quality Management & Business Excellence*, 1-28, doi: 10.1080/14783363.2020.1768062
- IPEADATA Homepage, <http://www.ipeadata.gov.br/Default.aspx>, last accessed 2020/06/10.
- ISO Homepage, <https://www.iso.org/the-iso-survey.html>, last accessed 2020/06/10
- Klafke RV (2016) *Comparação entre os níveis tecnológicos do Brasil e da China em relação ao PIB e o IDH*. Master thesis from Federal Technologic University of Paraná, Brazil, 104p
- Montgomery DC, Jennings CL, Kulahci M (2008) *Introduction to Time Series Analysis and Forecasting*. Wiley
- PNUD (2019) *Human Development Reports*, <http://www.hdr.undp.org/en/2019-report>, last accessed 2020/09/30
- Ribeiro LHMS, Beijo LA, Salgado EG, Nogueira DA (2019) Modelling of ISO 9001 certifications for the American countries: a Bayesian approach. *Total Quality Management & Business Excellence*, 1-27, doi: 10.1080/14783363.2019.1696672
- Salgado EG, Beijo LA, Sampaio PACA, Mello CHP, Saraiva P (2015) ISO 9001 certification in the American Continent: a statistical analysis and modelling. *International Journal of Production Research* 54(18), 5416-5433
- Sampaio PACA, Saraiva P, Rodrigues AG (2011) ISO 9001 certification forecasting models. *International Journal of Quality & Reliability Management* 28(1), 5-26