Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 55-10 (2022) 484–489

Prescriptive Cost Analysis in Manufacturing Systems

Sérgio Silva*. Vishad Vyas. Paulo Afonso**. Bret Boris***

**Bosch Car Multimedia, Braga, Portugal (e-mail: sergio.silva@pt.bosch.com; boris.bret@pt.bosch.pt)*

***University of Minho, Braga, Portugal (e-mail: b12192@algoritmi.uminho.pt; psafonso@dps.uminho.pt)*

Abstract: In many industries there is a high probability that production lines might produce extreme quantities, far away from what would be expected and/or desired considering the production and market conditions. This impacts considerably in product costs and respective margins. Thus, to support planning and decision-making in manufacturing systems, new analytical approaches must emerge to enable the transformation and representation of the available data. Namely, mathematical modeling can be used to improve cost analysis and optimization. This paper presents and discusses the use of prediction models, based on supervised machine learning algorithms and, particularly, linear regression in the context of manufacturing systems. Multiple linear regression models can be trained for different time periods allowing a better control of costs. They can be also used as decision-making tools for subsequent periods. Data from September 2020 to June 2021 of a first-tier supplier of the automotive industry was used considering two different training periods, in two groups of production lines of a specific product. Results of R^2 and MAE validate and show the relevance of the proposed models. The accuracy of these models depends on the artificial intelligence techniques and the training periods.

Keywords: Cost Analysis, Manufacturing Systems, Machine Learning, Regression Analysis, Prescriptive Analysis, Prediction Models.

1. INTRODUCTION

Companies are very interested in forecasting production costs. To help companies in planning and decision-making, new analytical tools have been emerging to enable the transformation and representation of the data, namely mathematical modeling. Good forecasting models will, above all, allow a better planning and control of the production. Thus, prescriptive analysis emerges as an opportunity, as it learns from the past to suggest options that can be used in decision making and risk reduction. Predictive analysis allows estimating and predicting the value of a variable, such as quantity and cost. It involves operations, mathematical and statistical techniques that allow studying and discovering trends, predicting failures (or anomalous situations) and variations that tend to exist in production systems. Thus, for a better understanding of behaviors over time, mathematical models that can describe present data and help in future prediction have emerged and been implemented in several companies. The idea is to assist on decision making, preventing, and being prepared to deal with serious problems that are predictable and may have a high probability of occurring. Nowadays, companies face higher levels of volatility, uncertainty, complexity and ambiguity (VUCA). Thus, important decision making in the industry such as introducing new products into the market, continuation, or termination of the existing products, must be supported by reliable information based on industrial cost and profitability. No less important is the contribution of this type of analysis to a consequent risk assessment associated with the behavior and trend of events. Prescription together with Artificial Intelligence (AI) techniques helps in the analysis of variations and risks associated with them, namely, in situations where forecasting is complex due to the uncertainty and instability that surrounds the company. The work presented here aims to contribute to the design of forecast models, namely, to predict the quantities to be produced, as well as the consequent costs. Multiple regression models allow very good approximations to the real data, even under high degrees of unpredictability. In this work, a set of forecasting models were developed and implemented, using linear regression techniques, namely, multiple linear regression. These are models that, given a set of variables as input, calculate a value for the output variable, which is the variable to be predicted. Data from September 2020 to June 2021 of a first-tier supplier of the automotive industry was used considering two different training periods, in two groups of production lines of a specific product.

For training the models, two distinct time periods were considered, corresponding to the average values of machines cycle times in twelve consecutive weeks. They were tested for the subsequent periods. The assumptions underlying the creation of a multiple linear regression model were analysed. In addition, measures of goodness of fit of the model were also analysed, namely, R^2, and Mean Absolute Error (MAE). To test the robustness of the models, the percentage errors of the observed values in relation to the predicted ones were analysed. On the other hand, based on the 95% confidence intervals for the model coefficients, the risk associated with the respective component was evaluated. Next section presents some studies on industrial contexts using from traditional statistical techniques to artificial intelligence approaches.

2. LITERATURE REVIEW

10.1016/j.ifacol.2022.10.223

In general terms, among the more traditional techniques found in the literature are those related to autoregressive models (AR), moving average (MA), simple exponential smoothing (SES), and Autoregressive Integrated Moving Average (ARIMA). These are techniques that are based on time series and use the independent variable masking time to predict other variables. Autoregressive is related with autocorrelation, that is, correlation of certain past periods with the current period. Autoregressive time series are commonly seen as an approximation to univariate time series modelling. The moving average, in turn, is an average calculated for a certain period. Regression errors result from a linear combination of error terms that occurred at past times. However, other approaches have emerged with potential for the development of forecasting techniques. These arise to, together with specialists, contribute to the identification of weaknesses and behavioural deviations in the business. Thus, allowing to take, in advance, decisions adapted to the changes and circumstances. Among others, they allow estimating production costs and volumes, forecasting sales and profit.

Galli et al. (2021) use machine learning in conjunction with stochastic optimization to minimize drug stocks. They present a set of scenarios and over these a stochastic optimization approach to provide good decisions, both in quantity and quality.

Sengewald et al. (2022) use the Bayesian technique to learn a stochastically optimal sourcing strategy directly from quote data. They did not focus on making great predictions but on making optimal decisions. They present a significant improvement in the costs of the purchase and acquisition process. The model turns out to be also more robust to forecast errors.

Bertsimas et al. (2020) combine machine learning, operational research and management science for decision making. They focused on a conditional stochastic optimization problem, with data related to costs/revenues and on other associated data. They end up showing that the developed methods are applicable and computationally tractable in several situations, even when the data are not independent and identically distributed. They extended the technique to cases of some decision variables, such as price, that may affect the uncertainty of the future and whose causes are unknown. They also developed a prescriptive coefficient P to measure the prescriptive content of the data and operational effectiveness. In the manipulated data, the authors found an improvement in the order of 88%, according to the P coefficient.

Koops et al. (2020) presents a prescribing solution based on the probabilistic approach to cost-benefit analysis and the definition of relevant metrics. They combine a Wiener model, to repair imperfections, and simulations by the Monte Carlo method, to support the decision. The probabilistic approach helps to detect the best decision options for greater profit, in addition to the associated risk and potential cost. In addition, they allow to reduce the uncertainty of the results. According to the authors, this approach can give the industry guidelines to identify and optimize business cases for Prescriptive Maintenance, pointing to the sources of valuable data or information and, therefore, justify investments.

Poornima et al. (2020) presents a technique to improve decision-making support in big data, prescriptive analytics, which offers an improved advance in predicting consequences and their outcome. The best result is achieved using optimization techniques in prescriptive analysis that identify uncertainties in making better decisions. Since optimization improves the effectiveness of prescriptive analytics with varied applications.

Devriendt et al. (2018) provides a structured and detailed literature search on elevation modelling, identifying, and contrasting various groups of approaches. In addition, evaluation metrics for evaluating the performance of elevation models are reviewed. Random forests are among the best performing techniques in terms of Qini and Gini measures, although considerable variability in performance is observed between the various datasets of the experiments. In addition, elevation models are often observed to be unstable and exhibit strong variability in terms of performance. Furthermore, it appears that the available evaluation metrics do not provide an intuitively understandable indication of the actual use and performance of a model. Specifically, existing evaluation metrics do not make it easy to compare elevation models and predictive models and evaluate performance at an arbitrary cut or across the full spectrum of potential cuts. In conclusion, we highlight the instability of uplift models and the need for an application-oriented approach to evaluating uplift models as key topics for future research.

Bertoni et al. (2018) approached prescription by integrating value and sustainability models in a machine learning (ML) environment. This allows to operate with qualitative and quantitative variables and a faster exploration. Berk et al. (2019) used real data simulation, also to create a planning model for human resources. They applied robust optimization and maximum flexibility which, according to the authors, leads to increased profit. Seyeden et al. (2020) analysed algorithms related to K-neighbors, neural networks, regression, and vector support machines. They mention that, although neural networks and regression analysis are the most common approaches used, they can be improved by optimization or simulations. Chen et al. (2021) forecast the costs and duration of a project through the added value management theory, through the application of the Monte Carlo method, to simulate large amounts of data, together with statistics. Punia et al. (2020) propose a machine learning model, based on quantile regression, to determine order quantities for NewsVendor. They suggest a distribution-free and Z-constraint-oriented solution approach. Something important to mention, and according to the author, despite all the techniques that can be applied, external variables play a significant role in the accuracy of a model and in the estimates and, these, in turn, have a strong impact on the cost of inventory management. The authors state that quantile regression together with machine learning is better than the empirical solution. In addition, Machine or Deep Learning techniques surpass the traditional ones, ARIMA or ARIMAX, normally used in time series.

3. MODEL DEVELOPMENT

A simplified linear regression model is of the form:

$$Y = a_1X1 + a_2X2 + \cdots + a_{i-1}X_{i-1} + a_{i+1}X_{i+1} + a_nX_n + C,$$

Y is the variable to be predicted, that is, the dependent variable. The variables $X_i, i = 1, \dots, n$ are the explanatory variables, that is, the independent variables. The values $a_i, i = 1, \dots, n$, are the coefficients associated with each of the explanatory variables. C is a real constant. From this, we obtain one estimative, designed by $\hat{Y}$. The values of the coefficients and the constant C are such as to minimize the error, for each observation, between the actual values and the estimated value, $\varepsilon_i = Y_i - \hat{Y}_i$. For this, it is applied the least squares method which an estimate for the model coefficients is obtained so that the sum of squared errors is minimal, that is, so that the value of $\sum_{i=1}^{n} \varepsilon_i^2$ be minimal.

One of the most common measures in evaluating a linear regression model is the value of R^2. This indicates the amount of variation in the data is explained by the covariate and the closer to 1, the better the model's goodness of fit. Case $R^2 = 1$, this is a perfect model in the sense that the values predicted by the model coincide with the observed values. Case $R^2 = 0$, we are facing an inadequate model for the data. However, in other situations, the model may still be considered acceptable. If $R^2 > 0.9$, the model is also considered to have a good fit and if $R^2 > 0.5$ the model is also acceptable from the point of view of adjustment to the data (Marôco, 2018).

For a linear regression model to be considered valid, it is necessary to ensure the validation of the underlying assumptions. Firstly, the errors should follow a normal distribution of constant and null value and are independent. That is, $\varepsilon_i \sim N(0, \sigma^2)$, and they are independent. Secondly, homoscedasticity in which a constant variance is assumed, that is, an equal variance for all observations. That is, $var(\mu_i) = \sigma^2, i = 1,2, \dots, n$.

There are statistical tests such as the *Shapiro-Wilk* and *T-test*, that are suitable for verifying these assumptions. The first one allows testing the normality of a sample of data and the second if the mean of a sample is equal to a certain value (Marôco, 2018). However, the literature suggests alternatives, if it is not possible to validate the model's assumptions and obtain an adjusted model. Namely, the standardization and the removal residuals outliers. Standardization is a linear transformation that results in a new distribution of null mean and unit standard deviation. In addition, it does not change the shape of the initial distribution of the data. The removal of outliers consists in eliminating the observations for which the condition is verified. $|\varepsilon_i^*| \geq 2$ (Marôco, 2018).

For the design of a forecast model and its validation, it was extracted production data for the period between September 2020 to June 2021. The data under analysis comes from four production lines of a first-tier supplier of the automotive industry.

They were developed two multiple linear regression models for two distinct time periods, a first between September and November 2020 and a second period between January and March 2021. The models were tested for subsequent months. The respective models were tested to understand the behavior of the quantities observed in relation to the forecast. The test periods were, for the first model, between December 2020 and March 2021 and, in the case of the second model, the period between March and June 2021.

A week is a period of seven consecutive days. Thus, the "Week" variable assumes the values of 1, 2, 3, or 4, referring to the periods between the 1st and 7th, 8th and 14th, 15th and 21st and 22nd and 28th day of each month, respectively. For training and creating the models, the average values of the quantities (Qty) of 12 consecutive weeks of a period are considered. The model was then tested for the subsequent weeks.

"Line" is an independent variable in the developed models. Line types assume the values of "A" and "C", if they refer to Small Lines and the values of "B" and "D", if they refer to Big Lines. A model based on the pair of lines "A" and "C" and another one based on the pair of lines "B" and "D" were developed. The models were implemented with the assumption that the variables "A", "B", "C" and "D" are independent and binary (0 or 1). That is, if one variable assumes the value of 1, all the others assume the value of 0. Let PQty be the quantity forecast function, and k1, k2, k3 and k4 the number of each line type "A", "B", "C" e "D", respectively. Then:

$$\begin{aligned} PQty(k1\text{"A"}, k2\text{"B"}, k3\text{ "C"}, k4\text{ "D"}) \\ = k1 * PQty(\text{"A"}, 0,0,0) + k2 \\ * PQty(0, \text{"B"}, 0,0) + k3 \\ * PQty(0,0, \text{"C"}, 0) + k4 \\ * PQty(0,0,0, \text{"D"}), k1, k2, k3, k4 \in N \quad (1) \end{aligned}$$

The expression corresponding to PQty, considering the pair of lines ("A", "C") is obtained by multiple linear regression, and is defined by:

$$PQty(\text{"A"}, \text{B"}, \text{"C"}, \text{"D"}) = k1 * \text{"A"} + k2 * \text{"C"} + C \quad (2)$$

Analogously, considering the pair of lines ("B", "D"), the corresponding expression of PQty is defined by:

$$PQty(\text{"A"}, \text{B"}, \text{"C"}, \text{"D"}) = k1 * \text{"B"} + k2 * \text{"D"} + C \quad (3)$$

For this work, the study of the model coefficients, the constant and the statistics associated with them deserve our attention. In addition, the values of *standard error* and the test result, with a significance of 5%, to its nullity, P>|t|, were made.

A 95% confidence interval, *I*, was defined for the value of each coefficient. Adding the value of C, we obtain the set of predicted and expected values *IP*, for the respective row associated with the coefficient.

An "Evaluation" function was defined which, given an observed value (OV), shows the value of 1, if OV belongs to the IP range. Otherwise, presents the value 0.

$$Evaluation(OV) = \begin{cases} 1, & if\ OV \in IP \\ 0, & if\ OV \notin IP \end{cases}$$

To quantify the deviant behaviour of the predicted values in relation to the observed values, the variable "DP" was defined, called the percentage differential. It represents the percentage of differential of the Qty value in relation to the PQty value and is given by:

$$DP = \frac{Qty - PQty}{PQty} \times 100 \quad (4)$$

The "Evaluation" function assists in the risk analysis associated with each line. It allows to assess whether the average quantity produced by a line in each week is within the predicted values. The value of "DP" shows how high is the deviation of the observed value from the predicted value. The higher its value, the greater the discrepancy between the observed and predicted values.

Within each model, two sub models were trained and tested. The first one is based on the small lines and the second on the big lines.

4. ANALYSIS OF RESULTS

4.1 Model I: training period: September - November 2020

The period from 01-09-2020 to 31-11-2020 was considered for training purposes.

4.1.1 Model I.1: based on the small production lines

Table 1. Coefficients of Model 1.1

Variable	Coefficients	Std. Error	t-value	P>|t|
A	-10125.86	358.54	-28.24	0.0
C	-10809.61	358.54	-30.15	0.0
Intercept	12446.86	213.00	58.44	0.0

The forecast model is defined by the function:

$$PQty(\text{A"}, \text{B"}, \text{"}C\text{"}, \text{"}D\text{"}) = -10125.86 * \text{"A"} - 10809.61 * \text{"}C\text{"} + 12446.86 \quad (5)$$

The model presented values of R^2 and MAE of 0.97 and 759.05, respectively.

Table 2. Prediction Values for Model 1.1

Variable	Lower prediction	Upper prediction
A	1749.46	2892.54
C	1065.71	2208.79
B	12233.86	12659.86
D	12233.86	12659.86

Model I.1 was tested for the following months, December 2020, January, February and March 2021. Table 3, shows the results of the "Evaluation" function and the DP values for each month of the test period.

Table 3. Test Model I.1

Line	Month	Qty	PQty	Evaluation	DP
A	Dec	2017.25	2321	1	-13,09
A	Jan	2684.75	2321	1	15,67
A	Feb	2195.5	2321	1	-5,41
A	Mar	2862	2321	1	23,31
B	Dec	10548	12446.86	0	-15,26
B	Jan	13628.25	12446.86	0	9,49
B	Feb	10417.75	12446.86	0	-16,3
B	Mar	12484.25	12446.86	1	0,3
C	Jan	2320	1637.25	0	41,7
C	Feb	2026.25	1637.25	1	23,76
C	Mar	2332.75	1637.25	0	42,48
D	Dec	10108.25	12446.86	0	-18,79
D	Jan	13175.25	12446.86	0	5,85
D	Feb	10114.25	12446.86	0	-18,74
D	Mar	12308.25	12446.86	1	-1,11

For each observation, the absolute values of DP, Abs(DP), were computed and analysed. In addition, small line C is the one with the greatest risk of producing quantities outside the forecast, contrary to small line A, which is more regular, that is, values closer to the forecast. In monthly terms, the month of January is more likely to present large variations in production. Big lines B and D are more sensitive in the months of January and February. We can observe that, when we are facing an Abs(DP) greater than 40%, it is likely that the observed value is outside of what was expected. When Abs(DP) assumes values lower than 40%, we may have situations in which the recorded value is or is not within what is expected.

4.1.2 Model I.2: based on the big production lines

Table 4. Coefficients of Model I.2

Variable	Coefficients	Std. Error	t-value	P>|t|
B	10645.88	372.59	28.57	0,0
D	10289.60	372.59	27.62	0,0
Intercept	1979.13	208.88	9.47	0,0

The forecast model is defined by the function:

$$PQty(\text{"A"}, \text{"}B\text{"}, \text{"}C\text{"}, \text{"}D\text{"}) = 10645.88 * B + 10289.60 * \text{"}D\text{"} + 1979.13. \quad (6)$$

This model presented values of R^2 and MAE of 0.97 and 816.92, respectively.

Table 5. Prediction Values for Model I.1

Variable	Lower prediction	Upper prediction
A	1770.25	2188.01
C	1770.25	2188.01
B	12043.54	13206.48
D	11687.26	12850.20

The model was tested for the following months, December 2020, January, February and March 2021. Table 6 shows the results of the "Evaluation" function and the DP values for each month of the test period.

Table 6. Test Model I.2

Lineid	Month	Qty	PQty	Evaluation	DP
A	Dec	2017.25	1979.13	1	1,93
A	Jan	2684.75	1979.13	0	35,65
A	Feb	2195.5	1979.13	0	10,93
A	Mar	2862	1979.13	0	44,61
B	Dec	10548	12625	0	-16,45
B	Jan	13628.25	12625	0	7,95
B	Feb	10417.75	12625	0	-17,48
B	Mar	12484.25	12625	1	-1,11
C	Dec	2673.5	1979.13	0	35,08
C	Jan	2320	1979.13	0	17,22
C	Feb	2026.25	1979.13	1	2,38
C	Mar	2332.75	1979.13	0	17,87
D	Dec	10108.25	12268.73	0	-17,61
D	Jan	13175.25	12268.73	0	7,39
D	Feb	10114.25	12268.73	0	-17,56
D	Mar	12308.25	12268.73	1	0,32

In this case, it is only when Abs(DP) are around 2% that the observed values are within what would be predicted. In general, all lines present a high risk of producing quantities that are far from what would be expected. This asks for a better production planning.

4.2 Model II: Training period: January – March 2021

In this model, it was considered for training purposes the period between 01-01-2021 to 31-03-2021.

4.2.1 Model II.1: based on the small production lines

The quantity forecast model is given by:

$$PQty(\text{"A"},\text{"}B\text{"},\text{"}C\text{"},\text{"}D\text{"}) = -10450.39 * \text{"A"} - 10804.80 * \text{"}C\text{"} + 13031.14 \quad (7)$$

The model presents values of R^2 and MAE of 0.88 and 1487,21, respectively. The model was tested for the following months, April, May and June 2021. Table 7 shows the results of the "Evaluation" function and the DP values for each month of the test period.

Table 7. Test Model II.1

Lineid	Month	Qty	PQty	Evaluation	DP
A	Apr	2797.5	2580.75	1	8,4
A	May	3401.25	2580.75	1	31,79
A	Jun	2831	2580.75	1	9,7
B	Apr	7185.5	13031.14	0	-44,86
B	May	14257.25	13031.14	0	9,41
B	Jun	12687.75	13031.14	1	-2,64
C	Apr	1995.75	2226.33	1	-10,36
C	May	2824.25	2226.33	1	26,86
C	Jun	1864	2226.33	1	-16,27
D	Apr	6605.5	13031.14	0	-49,31
D	May	14087	13031.14	0	8,1
D	Jun	12826.5	13031.14	1	-1,57

Analysing the relationship between the values of Evaluation and Abs(DP), for model II.1, it is possible to observe that, only the production line A keeps the values within the predicted ones. Months 4 and 5 need to be better planned for lines B and D.

4.2.2 Model II.2: based on the big production lines

The quantity forecast model is given by:

$$PQty = 10814.82 * \text{"}B\text{"} + 10440.37 * \text{"}D\text{"} + 2403.54 \quad (8)$$

The model presented values of R^2 and MAE of 0.88 and 1493,88, respetivamente. The model was tested for the following months, April, May and June 2021. Table 8 shows the results of the "Evaluation" function and the DP values for each month of the test period.

Table 8. Test Model II.2

Lineid	Month	Qty	PQty	Evaluation	DP
A	Apr	2797.5	2403.54	1	16,39
A	May	3401.25	2403.54	0	41,51
A	Jun	2831	2403.54	0	17,78
B	Apr	7185.5	13218.36	0	-45,64
B	May	14257.25	13218.36	1	7,86
B	Jun	12687.75	13218.36	1	-4,01
C	Apr	1995.75	2403.54	0	-16,97
C	May	2824.25	2403.54	0	17,5
C	Jun	1864	2403.54	0	-22,45
D	Apr	6605.5	12843.91	0	-48,57
D	May	14087	12843.91	0	9,68
D	Jun	12826.5	12843.91	1	-0,14

Analysing the relationship between Evaluation and Abs(DP) for model II.2, we also conclude that only production line A maintains the values of quantities produced within the predicted values.

5. DISCUSSION AND CONCLUSIONS

The developed models allow, in addition to forecasts of production quantities, to predict deviant behaviour of each specific production line. From both models I and II, the line A is the one that presents production levels within the predicted values. In general, we can see that the production by line A tends to remain regular, that is, the prediction and estimation, with an interval 95% confidence demonstrates values close to the real production. Analysing the period from September to November 2020, it was expected that the absolute value of DP will not exceed 20%. If the variability is higher, there is a high risk that the respective production line is producing unwanted quantities, above or below the expected value. We can also observe that absolute DP values are higher in model I than in model II. That is, there is a tendency for a greater variation of quantities in relation to the forecast. The absolute DP values present different variations in the different production lines. These variations may be associated with external factors that occurred during the period under analysis. Lay-offs, temporary stoppages in production, lack of people, material, or among others. The work developed allowed us to call attention to the care that we must have in the control and planning of production, namely, in the first months of the year, January and February. In these months, practically all lines are more sensitive to large variations in production. With greater control of productivity, consequently, costs will also be more controlled and not subject to large variations. In models based on small lines, a variation in production around or above 40% leads us to a certainty of productivity beyond the expected. In models based on big lines, a variation of 5% is enough for this to happen. However, in this case, if production variations are less than 5%, we can say that productivity is controlled. The same does not happen in the model based on small lines, since a variation in production below 40% does not guarantee productivity control. In this case it was found that there are lines that, even with a productivity variation of less than 40%, still present values outside the expected.

Thus, the approach proposed here help in the detection of deviant behaviour in quantities produced and, therefore, in the analysis of the associated product costs and the risk of such costs. This risk is related to the probability that a given production line will produce extreme amounts, far away from what would be expected and/or desired considering the production and market conditions. This impacts considerably in product costs and respective margins, asking for a frequent revision of standard costs using real data on the production and market conditions. Multiple Linear regression is a good technique for creating predictive models, even when we are in the presence of non-numerical variables. Furthermore, the linear regression model allows to define a confidence interval for the respective predicted values, both in quantities and costs. The analysis of different periods helps in the understanding of the variation in production over time. They reveal different behaviours, patterns, and trends of the production process. At the same time, the model allows detecting and predicting anomalous situations during the production process (breaks, layoff, lack of material, etc.). Using confidence intervals for the expected values, it is possible to define the limits for which the variation is acceptable. In other words, the model allows the assessment of the risk of each of the lines produce outside the expected limits.

There is a need of better models for predictive and prescriptive analysis in production systems and several opportunities for further research can be explored from a conceptual and practical perspective namely, using different artificial intelligence techniques and applying these models to different companies and industries.

ACKNOWLEDGMENTS

This work has been supported by FCT – Fundação para a Ciência e Tecnologia within the R&D Units Project Scope: UIDB/00319/2020.

REFERENCES

Berk, L., Bertsimas, D., Weinstein, A. M., & Yan, J. (2019). Prescriptive analytics for human resource planning in the professional services industry. *European Journal of Operational Research*, *272*(2), 636-641.

Bertoni, A., Dasari, S. K., Hallstedt, S. I., & Andersson, P. (2018). Model-based decision support for value and sustainability assessment: Applying machine learning in aerospace product development. In *DS 92: Proceedings of the DESIGN 2018 15th International Design Conference* (pp. 2585-2596).

Bertsimas, D., & Kallus, N. (2020). From predictive to prescriptive analytics. *Management Science*, *66*(3), 1025-1044.

Chen, X., Cheng, L., Deng, G., Guan, S., & Hu, L. (2021, July). Project duration-cost-quality prediction model based on Monte Carlo simulation. In *Journal of Physics: Conference Series* (Vol. 1978, No. 1, p. 012048). IOP Publishing.

Devriendt, F., Moldovan, D., & Verbeke, W. (2018). A literature survey and experimental evaluation of the state-of-the-art in uplift modeling: A stepping stone toward the development of prescriptive analytics. *Big data*, *6*(1), 13-41.

Galli, L., Levato, T., Schoen, F., & Tigli, L. (2021). Prescriptive analytics for inventory management in health care. *Journal of the Operational Research Society*, *72*(10), 2211-2224.

Koops, L. G. (2020). Optimized Maintenance Decision-Making–A Simulation-supported Prescriptive Analytics Approach based on Probabilistic Cost-Benefit Analysis. In *PHM Society European Conference* (Vol. 5, No. 1, pp. 14-14).

Marôco, J. (2018). Statistical Analysis with SPSS StatisticsAnálise Estatística com o SPSS Statistics.: 7ª ed. ReportNumber.

Poornima, S., & Pushpalatha, M. (2020). A survey on various applications of prescriptive analytics. *International Journal of Intelligent Networks*, *1*, 76-84.

Punia, S., Singh, S. P., & Madaan, J. K. (2020). From predictive to prescriptive analytics: A data-driven multi-item newsvendor model. *Decision Support Systems*, *136*, 113340.

Sengewald, J., & Lackes, R. (2022). Prescriptive Analytics in Procurement: Reducing Process Costs.

Seyedan, M., & Mafakheri, F. (2020). Predictive big data analytics for supply chain demand forecasting: methods, applications, and research opportunities. *Journal of Big Data*, *7*(1), 1-22.