Impact Analysis of KPI Scenarios, Automated Best Practices Identification, and Deviations on Manufacturing Processes

Maria J. Lopes
Bosch Termotecnologia S.A.
Aveiro, Portugal
mariajoao.lopes@pt.bosch.com

Eugénio M. Rocha
Center for Research & Development in Mathematics and Applications
Department of Mathematics, University of Aveiro
Aveiro, Portugal
eugenio@ua.pt

***Abstract*—Data-driven applications are becoming more and more ubiquitous throughout the manufacturing industry. The decision of which projects to start often comes from the input of process experts who identify a concrete potential for improvement in a certain area. However, a different approach may be taken when it is not entirely clear where to start looking for patterns or potential information which can trigger continuous improvement activities. This is specially relevant in manufacturing processes with a high level of maturity and stability. In this article, the authors propose a generic approach for conducting impact analysis with a use case which aims to deconstruct the Overall Equipment Effectiveness (OEE), a quite known Key Performance Indicator (KPI), in a manufacturing production line from a Bosch plant located in Portugal. This methodology is focused on identifying the best and worst scenarios by creating a ranking and subsequently pinpointing possible causes, identifying best practices and devising strategies to deal with these non-optimal scenarios. The methodology can be seen as an alternative to complex statistical hypothesis testing by relying on measuring several distribution differences.**

***Index Terms*—Impact metrics, continuous improvement, industry 4.0, process deviations, root cause analysis**

I. Introduction

The quest for optimization has been a long one indeed in the manufacturing industry. In order to make manufacturing more efficient and cost effective, different methodologies have been devised to find the next area where to deploy improvement measures. In the literature, one can find countless examples of such strategies (e.g. Six Sigma) with very well defined stages and implementations, see [1]. These methods usually combine the expertise of dedicated professionals and some data. They also represent a significant overhead due to all the data collection involved, see [2]. Therefore, it would be relevant to reflect on the hypothesis of whether a mainly data-driven approach can be used to identify potential areas for improvement, see [3]. Interesting approaches, albeit different from this one, have been developed in [4] and [5].

The idea is thus that this approach can be applied by someone who is not in any way an expert on the subject. process experts will still intervene yet will do so in very well-defined points in time. This allows for them to be involved only when necessary and grants them the ability to focus on improvement activities where their expertise is paramount. Looking for what to improve next can thus be delegated to someone with data analysis skills and can be automated afterwards. This approach relies on the definition of scenarios, which are simply aggregations of one or more factors which are believed to have an impact on the process being analysed. After both the best and the most problematic scenarios are identified, a positive deviance approach (see [6]) will be used to understand what distinguishes the top ranked scenarios from the bottom ranked ones. By the end of this work, a complete approach for the analysis of the impact of the various scenarios will have been explored and conclusions drawn.

II. Methodology and the OEE Use Case

The main goal behind this work is to devise a strategy which allows to look for improvement opportunities within an manufacturing process measured by a specific management measure. This approach is actually a component of a much broader pipeline as depicted in Figure 1. The method described in this article is represented as the Root Cause 1 (RC1) component and complements the approach RC2, see [7]. The basis is the DMAIC framework (Define, Measure, Analyse, Improve, and Control), see [8] and [9].

In order to test the proposed approach, a data set containing real production data of a Bosch Manufacturing Production Line (MPL), located in Portugal, with the OEE data for shifts over the period of 9 months was used. This data has been stored and outputted through a Manufacturing Execution System (MES). The OEE was the chosen KPI for this use case since it is generally acknowledged as being the industry standard for measuring performance in MPLs. However, other KPI, metric or value could be used. Besides this, data on the type of shift plus the Target Cycle Time (TCT) were used as scenario variables. The TCT is a good candidate for a scenario since it condenses the information of both the product type (e.g. which product family is being produced) with the production scenario being carried out (e.g. the number of workers allocated to the line in that shift). Later on, we can add additional (process) variables which will provide

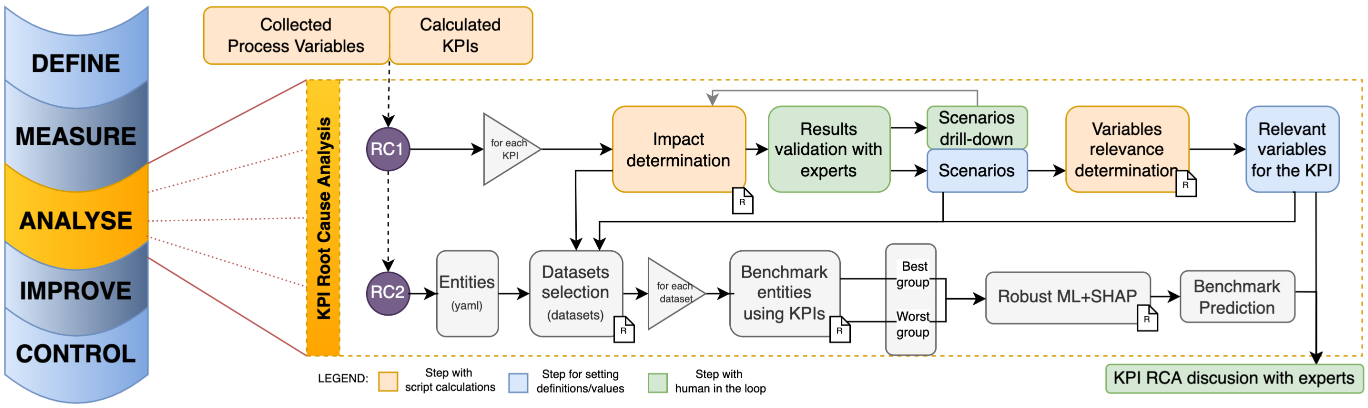

Fig. 1: Scheme representing the integration of the approach (RC1) in an automated data-driven DMAIC methodology.

more information in an effort to unveil the causes behind efficient/non-efficient behavior.

Data sets were split by month (time periods) before being inputted into the plugin being used for this data processing. This allows for the creation of an array of scenarios to be tested, as explained in subsection A.

The following subsections describe in more detail the methodology presented in Figure 2, where the gray squares account for the subsection identifier where it is discussed as a convenience to the reader.

A. *Impact determination*

This method was designed with the intent of being as generic as possible and with that in mind, the starting point became the Probability Density Function (PDF) f_{KV} of a certain metric, KPI or factor variable; what from now on we call generically a key value (KV). This abstraction allows for a purely statistical interpretation of the data at hand and enables all sorts of scenarios to be devised and tried out.

With this being said, the first step in the calculation is to list all the possible scenarios $\{S_1, \ldots, S_n\}$, for $n \in \mathbb{N}$, that the process experts would like to test. This can thus be interpreted as a highly exploratory approach. These scenarios will be applied as a KV filter made on the KV values by a combinations of n indicator functions of scenario variables $\{V_1, \ldots, V_m\}$, with values previously converted into categorical labels. In detail, each scenario variable V_j has a set of unique values $\{C_1^j, \ldots, C_{k_j}^j\}$, which generate the scenarios

$$S_i = (C_{k_1}^1, C_{k_2}^2, \ldots, C_{k_m}^m) \tag{1}$$

as a sequence of values that the (ordered) scenario variables $\{V_1, \ldots, V_m\}$ have at the current observation. Then, for each time period T, a scenario S_i induces a data filter on the KV data, generating a new KV variable that has the PDF denoted by $f_{KV_i}^T$. This establishes a generic baseline for comparison of the performance of the overall data set versus the one of the scenario being studied.

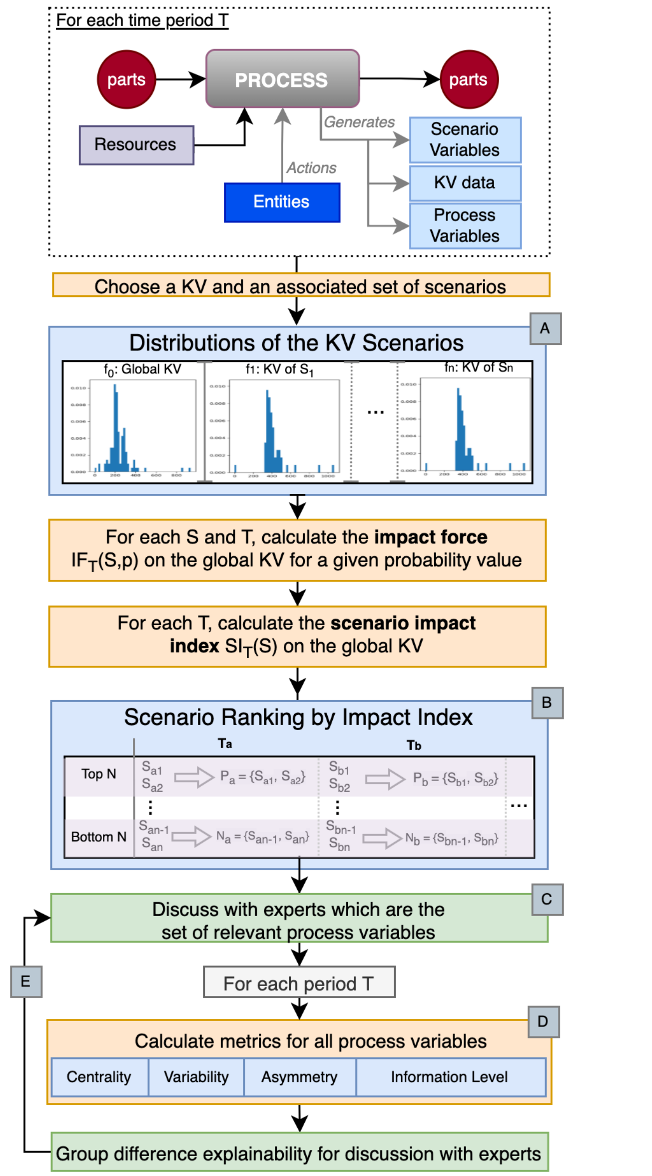

Fig. 2: Methodology general scheme (colors follow the legend from Figure 1).

Mathematically, we assume that variables KV_i are real-value random variables that are absolutely continuous or discrete and, in this case, the probability density function is defined by

$$f(y) = \sum_{x \in A} p(x)\delta(y - x), \tag{2}$$

where A is a countable subset of the probability measurable space, $p : A \to [0, 1]$, and δ is the Dirac measure. We do not deal with singular distributions. From here, we may calculate the values

$$Q_i^T(p) = \inf \left\{ x \in \mathbb{R} \;:\; p \le \int_{-\infty}^{x} f_{KV_i}^T(y)\, dy \right\}, \tag{3}$$

where $p \in [0, 1]$. Applying the same idea to the KV variable, without any scenario filtering, we also have

$$Q_0^T(p) = \inf \left\{ x \in \mathbb{R} \;:\; p \le \int_{-\infty}^{x} f_{KV}^T(y)\, dy \right\}. \tag{4}$$

The definition (2) allows to deal with data sets (e.g., data collected by a MES) and definitions (3) and (4) behave as generalizations of percentile functions, where the corresponding cumulative distribution function may be relaxed to be just non-decreasing (i.e., may be piecewise constant).

Since, for each time period T, a scenario S_i behaves as a data filter of the KV data, we can associate a weight w_i as the percentage of observations that result from applying such filter. Fixing a small $\epsilon > 0$, for each scenario S_i and time period T, the **impact force** of the scenario on the KV value for a given probability p level is defined by

$$\mathrm{IF}_T(S_i, p) = \frac{w_i \int_{p-\epsilon}^{p} Q_0^T(y) - Q_i^T(y)\, dy}{\int_{p-\epsilon}^{p} Q_0^T(y)\, dy}. \tag{5}$$

The multiplication by w_i accounts for the fact that different scenarios can stem from largely different amounts of data.

Finally, we choose a finite set of probability levels $\mathcal{P} = \{p_1, p_2, p_3, \dots\}$ and define the **scenario impact index** $SI_T(S)$ on the values of KV by

$$\mathrm{SI}_T(S) = \frac{\sum_{p_j \in \mathcal{P}} \mathrm{IF}_T(S, p_j)}{|\mathcal{P}|} \tag{6}$$

where $|\mathcal{P}|$ denotes the cardinality of $\mathcal{P}$.

For a better understanding of the reader, the calculations (2)-(6) for an example data set would shed some light on the purposed method. However, such is not viable since this approach requires a significant number of observations since it deals with filtrations of distributions that need to be also valid distributions.

B. Scenarios Ranking by Impact Factor

The approach described before generates a impact factor ranking, allowing the extraction of the top N scenarios and the worst N scenarios. The results for the use case were condensed into the heat map shown in Fig. 3.

The Y Axis depicts the scenarios being tested. Scenarios A, B and C refer to the type of shift - morning, afternoon and night. However, this association has been scrambled for anonymity reasons. Scenarios with three digits refer to the TCT whereas scenarios combining both the shift type and TCT follow the notation TCT_ShiftType. TCTs have been anonymised through an offset. Months are depicted in the X Axis. The colour legend depicts variations within the +15% to the -15% range. Scenarios were also sorted in a way that allows for an easier heatmap readability. A heatmap is an adequate way to display these sort of results because it allows for both pattern identification and a higher sensibility to the ranges of the values being interpreted.

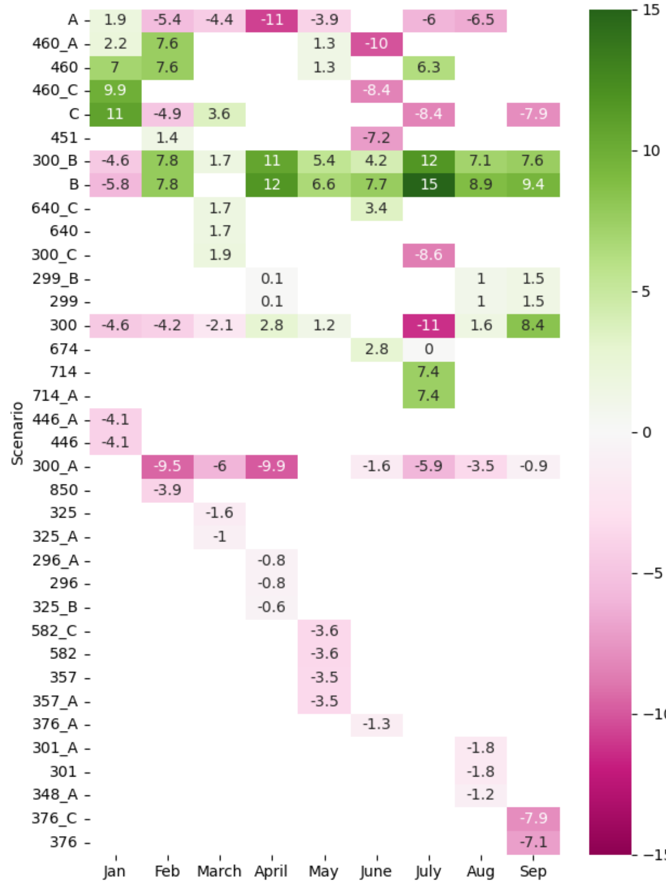

Fig. 3: Heatmap of the scenario impact index. Top performers are shown in green while bottom ones are shown in pink. Blank cells refer to scenarios which do not feature in neither the top or bottom ranking having been removed for clarity purposes.

Based on the results displayed in Figure 3, one can conclude that:

- Reducing the amount of available scenarios would probably lead to a much higher stability in the MPL i.e. less performance fluctuations;
- Shift A started off the year well but then displayed subpar performance for most of the remaining part of the year;
- Shift B is a systematic top performer throughout the year;
- Shift C displays the higher performance variability.
- Regarding the TCT, e.g. 376 is perceived as a *difficult*

TCT and TCT 460 as a *good* TCT in production.
- It is interesting to notice that the pair 300_A compensates the impact of the pair 300_B.

C. Results validation with experts

These results have been presented to those responsible for production planning in this MPL and they have been perceived as a correct depiction of what has happened throughout this time period. The benefits of the approach were readily understood since a purely numerical method mirrors the experience that these professionals have. Process experts have specifically highlighted the potential of the usage of this tool in close loop with PDCA (Plan-Do-Check-Act) improvement cycles. Since the scenarios to be improved have been identified, one can move onwards to the next step where the explainability of these results is attempted by relating them with process variables.

D. Variables relevance determination

After the ranking of the scenarios presented in the section above, the focus is now on enriching the initial data set with (experts determined) process variables, which will allow for some best practices to be identified. In this particular case, 3 variables were selected: a variant of the First Pass Yield (vFPY), the Quality Factor (QF) and the amount of changeovers (CH). The variable vFPY is calculated as

$$vFPY = \frac{\text{Nº OKs on first try}}{\text{All tests done}}. \tag{7}$$

The reason why this variable amounts for a variant is that the denominator should include the number of unique parts being produced, not the overall number of tests done. This translates into values which are actually lower than the actual FPY.

Up next, we have the QF which is one of the OEE factors. The QF is higher than the vFPY or equal to it since it takes into account all OKs and not just the ones on the first try. It is defined by:

$$QF = \frac{\text{Nº OKs}}{\text{All tests done}}. \tag{8}$$

The CH variable acts as a measurement of the number of different references produced throughout the shift in order to assess whether the diversity of references influences the overall performance. This metric is rather different in the sense that it is simply an integer, not a ratio. In order for it to be in the same range, it was normalized.

In the next step, a set of functions where chosen to determine: centrality (μ for the mean), variability (σ for the standard deviation), asymmetry (s for the skewness), and information level (E for the entropy).

1) Metric Calculations: From the scenario ranking, we have determined the top N scenarios $\mathcal{T}$ and the bottom N scenarios $\mathcal{B}$. Those imply the generation of two data filters over the process variables generating two sets of new variables

$$X_{\mathcal{T}} \in \{\text{vFPY}_{\mathcal{T}}, \text{QF}_{\mathcal{T}}, \text{CH}_{\mathcal{T}}\}, \tag{9}$$

$$X_{\mathcal{B}} \in \{\text{vFPY}_{\mathcal{B}}, \text{QF}_{\mathcal{B}}, \text{CH}_{\mathcal{B}}\}. \tag{10}$$

Then, the metrics associated with the chosen functions can be computed as

$$\delta_h(X) = \frac{h(X_{\mathcal{T}}) - h(X_{\mathcal{B}})}{h(X_{\mathcal{T}})} \quad \text{for } h \in \{\mu, \sigma, s, E\}. \tag{11}$$

Table I resumes the metrics values obtained for the use case.

TABLE I: Comparison metrics across variables

Variable	δ_μ	δ_σ	δ_s	δ_E
vFPY	2.47%	-69.13%	43.48%	-66.38%
QF	-0.42%	6.98%	-38.89%	-23.03%
CH	-2.38%	5.87%	13.54%	12.33%

2) Metric Calculations Cut-Off: From the comparison metrics computed on the previous step, a cut-off procedure is established based on a parameter $\epsilon \in [0, 1]$ that is selected by the user. The procedure is done in two steps: (1) find the relevant metrics; and (2) within each relevant metric, choose the relevant variables.

In mode detail, the first step consists in calculating cut-off values ζ_h, for $h \in \{\mu, \sigma, s, E\}$, satisfying

$$\zeta_h = (|\max(\delta_h)| - |\min(\delta_h)|)\,|\max(\delta_h)|, \tag{12}$$

and identifying the set of relevant metrics h defined by

$$\mathcal{R}_\epsilon = \{h \in \{\mu, \sigma, s, E\} \,:\, \zeta_h > \epsilon\}\,. \tag{13}$$

In this case, the parameters ϵ = 0.1 and ϵ = 0.3 were used, just for comparison reasons. These cut-off values produces Table II. Hence, we have

$$\mathcal{R}_{0.1} = \{\sigma, s, E\} \quad \text{and} \quad \mathcal{R}_{0.3} = \{\sigma, E\}. \tag{14}$$

Having chosen which metrics h are relevant, the next step follows.

TABLE II: Cut-off values, where blue values are relevant for ϵ = 0.3 and ϵ = 0.1; and cyan values are relevant just for ϵ = 0.1.

parameter	ζ_μ	ζ_σ	ζ_s	ζ_E
$\epsilon = 0.1$	0.0005	0.4375	0.1302	0.3588
$\epsilon = 0.3$	0.0005	0.4375	0.1302	0.3588

The second step consists in normalizing the absolute values from Table I in respect to the maximum values of each column since we want to identify the maximum discrepancies, by defining

$$\text{Rel.}\,\delta_h(X) = \frac{|\delta_h(X)|}{|\max(\delta_h)|}. \tag{15}$$

The results of such calculations are presented in Table III.

Now relevant variables are the higher ones. To choose which ones to mark, we use the condition

$$\text{Rel.}\,\delta_h(X) > 0.75\,(1 - \epsilon). \tag{16}$$

Equation 16 when valid, is used to assess whether a variable X is labelled as relevant or non-relevant. Values of $\delta_h(X)$ that do not satisfy the equation are ignored. This is an heuristic cut that allows us to ensure that values that are (at least) 25% below the maximum are always selected. In particular, for ϵ = 0.3, we have Rel. $\delta_h(X) > 0.525$; and, for ϵ = 0.1, we have Rel. $\delta_h(X) > 0.675$. Notice that the lower the ϵ, the fewer variables are considered since the variables appearing with smaller parameters also are included in higher parameters.

TABLE III: Relative delta values per variable and metric, where blue values are relevant for ϵ = 0.3 and ϵ = 0.1; and cyan values are relevant just for ϵ = 0.1.

Variable	Rel. δ_μ	Rel. δ_σ	Rel. δ_s	Rel. δ_H
vFPY	1.0000	1.0000	1.0000	1.0000
QF	0.1681	0.1010	0.8944	0.3470
CH	0.9626	0.0849	0.3114	0.1857

Using the input from Table II, Table III can be used to acknowledge that the vFPY is the main contributor to asymmetry and information level for the overarching KPI in this study, the OEE. Whether this contribution is positive or negative, this will be assessed using the sign of the data from Table I. Although there are also values above the cut-off value, these must be ignored since only the variability and the information level were considered relevant through this method. The relative delta for the mean of the vFPY is one such example having been ignored due to not passing the cut-off metric - roughly meaning that the difference between the mean of the top N observations compared with the mean of the bottom N observations is not significant.

For the vFPY variability and information level, Table I displays for both a negative sign. In practical terms, this means that within the bottom performers group, the vFPY displays the most variability and less information level (entropy). This is another way of saying that the local (from shift to shift) and global variability of the vFPY has a significant negative impact in the worst performers, composed of scenarios 300_B, C, 460_1, 460_C, 460. In order to improve the OEE, measures should be taken in order to reduce the vFPY variability. The results of this analysis are summarised in Table IV.

TABLE IV: Results between KPI and positive/negative impact (ϵ = 0.1 and ϵ = 0.3)

Impact	Centrality	Variability	Asymmetry	Inf. Level
Positive	-	-	vFPY	-
Negative	-	vFPY	QF	vFPY

It is important for one to realize that this analysis is dependent on the choice of ϵ. Table IV shows the impact of choosing different ϵ. Even though ϵ = 0.3 seems to convey more information since it compounds the values from ϵ = 0.1, it highlights variables whose components have a significant lower impact.

E. Scenarios Drill-down

Even though it is now known that the vFPY is one of the process variables with the most impact on the negative performers group, it is still not completely clear where this happens or which workstations from the MPL are contributing the most to this phenomenon. Therefore, in this section, this question will be addressed.

A new dataset was prepared using data for each workstation with the number of successful (OK) vs unsuccessful (NOK) processes in relation to the target part count. The scenarios which include the top and bottom performers are exactly the same having originated from the first step of the methodology. Table V displays the ratios cut-off metrics, indicating that all statistics are relevant. Table VI features the variables representing the NOK ratio per stations. The stations are uniquely identified by their variable code for anonymization purposes. An ϵ of 0.3 was used.

TABLE V: Cut-off metric for minimizing variables

Cut-Off	Centrality	Variability	Asymmetry	Inf. Level
ϵ = 0.3	3.938	40.069	34.514	1.6170

TABLE VI: Relative delta values for the minimizing group

Variable	Rel. δ_μ	Rel. δ_σ	Rel. δ_s	Rel. δ_H
v07_NOK	0.026	0.005	0.000	0.000
v15_NOK	0.576	0.300	0.390	0.838
v17_NOK	0.214	0.053	0.045	0.524
v19_NOK	0.177	0.021	0.268	0.030
v21_NOK	0.155	0.018	0.323	0.798
v24_NOK	0.103	0.021	0.033	0.492
v26_NOK	0.000	0.000	0.140	0.610
v28_NOK	1.000	1.000	1.000	1.000
v31_NOK	0.249	0.186	0.168	0.415
v33_NOK	0.143	0.082	0.069	0.174

For the variables which are to be maximized, the variability, asymmetry and information levels are relevant criteria for a cut-off value of 0.525 as seen in Table VII.

TABLE VII: Cut-off metric for maximizing variables

Cut-Off	Centrality	Variability	Asymmetry	Inf. Level
ϵ = 0.3	0.025	0.779	3.064	0.458

In this group, the relative delta values are more scattered.

F. Drill down variables relevance determination

From these, we can convey the following results:

- Station represented by variable v28 displays values implying it should be looked into due to topping all the relevant delta values.
- The entropy of the distributions is a focus topic on almost all stations indicating high local variability in the OK percentages.
- Stations represented by variables v30 and v32 have the same manufacturing process - however, variable v32

TABLE VIII: Relative delta values for the maximizing group

Variable	Rel. δ_μ	Rel. δ_σ	Rel. δ_s	Rel. δ_H
v01_OK	0.276	0.121	0.087	0.836
v02_OK	0.235	0.287	0.043	0.753
v03_OK	0.000	0.484	0.131	0.000
v04_OK	0.438	0.480	0.064	0.445
v05_OK	0.478	0.458	0.066	0.442
v06_OK	0.478	0.458	0.066	0.442
v08_OK	0.387	0.000	0.186	0.638
v13_OK	0.172	0.017	0.191	0.654
v14_OK	0.411	0.042	0.147	0.870
v16_OK	0.510	0.012	0.127	0.780
v18_OK	0.440	0.030	0.138	0.760
v20_OK	0.411	0.0570	0.142	0.762
v22_OK	0.429	0.065	0.169	0.744
v23_OK	0.502	0.026	0.162	0.796
v25_OK	0.443	0.037	0.149	0.751
v27_OK	1.000	0.544	0.000	0.546
v29_OK	0.450	0.130	0.102	0.773
v30_OK	0.053	0.204	0.557	0.029
v32_OK	0.707	1.000	0.351	1.000

displays a higher level of global and local variability in relation to its counterpart.

- The station represented by variable v27 (whose NOK value also was identified above) also possesses high local and global variability.
- The station represented by variable v01 also displays a relatively high entropy value. This can be due to the fact that this is the first station of the line and the rhythm oscillates at a higher level in first stations, see [11].

III. Conclusion

In this article the foundations for a best practices identification method as well as a root cause analysis one were established. Using statistical methods, it was possible to show its usefulness by process experts and ease of use by non-experts. Important conclusions have been achieved throughout the study and the authors view this as a compelling approach which deems further research necessary. In particular, an extended study of the MPL can be done by enriching the set of process variables in order to capture even more relevant situations at the workstation level, beyond the use of only OKs and NOKs. Another matter of improvement is how to translate the interpretation of above values to a non-expert language in a simple and unified way.

With that being said, future research directions include the ability to provide specific recommendations to the user based on fuzzy inference (to account for uncertainty) and also the joining of methods RC1 (this one) with RC2 (described in [7]) as previously mentioned.

Acknowledgment

The present study was developed in the scope of the Project Augmented Humanity (PAH) [POCI-01-0247-FEDER-046103], financed by Portugal 2020, under the Competitiveness and Internationalization Operational Program, the Lisbon Regional Operational Program, and by the European Regional Development Fund. The first author was supported by Bosch Termotecnologia. The second author was partially supported by the Center for Research and Development in Mathematics and Applications (CIDMA), through the Portuguese Foundation for Science and Technology, reference UIDB/04106/2022.

Data availability statement: The data sets used in this work are confidential information of Bosch, so they are not publicly available.

References

[1] F. M. G. S. Gonçalves, C. I. M. Bezerra, A. D. Belchior, C. C. Coelho and C. G. S. Pires, “A strategy for identifying, classifying and prioritizing improvement and innovation actions: A CMMI Level 5 and Six Sigma approach,“ 19th Australian Conference on Software Engineering (aswec 2008), 2008, pp. 104-111.

[2] T. Toroi, A. Raninen and L. Väätäinen, “Identifying process improvement targets in test processes: A case study,“ 2013 IEEE International Conference on Software Maintenance, 2013, pp. 11-19.

[3] M. B. Schrettenbrunnner, “Artificial-Intelligence-Driven Management,“ in IEEE Engineering Management Review, vol. 48, no. 2, pp. 15-19.

[4] J. Li, Y. Jia, B. Xu, F. Chen, Z. Yang and X. Li, “An improvement scheme for the overall line effectiveness of a production line: A case study,“ 2018 5th International Conference on Industrial Engineering and Applications (ICIEA), 2018, pp. 304-309.

[5] Y. Qian, J. Arinez, G. Xiao, Q. Chang, “Improved production performance through manufacturing system learning,“ 15th IEEE International Conference on Automation Science and Engineering (CASE), 2019, pp. 517-522.

[6] A. Chakraborty, “Process improvement — A positive deviance approach,“ 2013 IEEE International Conference on Industrial Engineering and Engineering Management, 2013, pp. 1117-1121.

[7] E.M.Rocha, A.F. Brochado, B. Rato, J. Meneses. “Benchmarking and Prediction of Entities Performance on Manufacturing Processes through MEA, Robust XGBoost and SHAP Analysis.“ (submitted to ETFA 2022).

[8] C. Yu, T. Huang, K. Chen, T. Huang, “Construct Six Sigma DMAIC Improvement Model for Manufacturing Process Quality of Multi-Characteristic Products,“ Mathematics, 2022, vol. 10.

[9] K. Kanyinda, I. Lazarus and O. Olanrewaju, “Influence of Six Sigma DMAIC to Reduce Time Wasting of Line Supervisor in Production Manufacturing,“ 2020 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), 2020, pp. 1296-1300.

[10] Malega P., Daneshjo N., Rudy V., Drábik P, “PDCA Cycle - Tool for Improvement of the Business Processes – Case Study,“ TEM Journal, 2021, vol. 10, pp. 1336 - 1347

[11] E. M. Rocha, M. J. Lopes, “Bottleneck prediction and data-driven discrete-event simulation for a balanced manufacturing line,“ Procedia Computer Science, 2022, vol. 200, pp. 1145-1154.