Mechanics & Industry **25**, 17 (2024)

https://doi.org/10.1051/meca/2024012

Mechanics & Industry

Available online at:
www.mechanics-industry.org

Advanced Approaches in Manufacturing Engineering and Technologies Design
Adinel Gavrus (Guest Editor)

ORIGINAL ARTICLE OPEN ACCESS

Method for holistic optimization of the manufacturing process numerically described as low-dimensional database

Cezarina Chivu[1], Mitica Afteni[1,2], and Gabriel-Radu Frumusanu[1,*]

[1] Dunarea de Jos University of Galati, Domneasca str. 111, 800201 Galati, Romania
[2] Rulmenti S.A., Republicii str. 320, 731108, Barlad, Romania

Received: 18 July 2023 / Accepted: 5 April 2024

Abstract. The management of the production processes in an optimal manner involves the usage of knowledge about past jobs as reference for current decisions. During a manufacturing flow in every process step the process engineers could be in situations that request quick decisions based on comparison of different potential manufacturing paths. The Method for Holistic Optimization was developed in order to be used as support for decisions. The method was validated thru different studies. For the mentioned studies there were used artificial and real instances databases. The approach of the optimal management of the manufacturing processes was developed in the current study in order to estimate the consequences of a decision, are used known methods, such as: NN modeling, big data analysis, statistics, etc. In all these cases, the database size plays an essential role in terms of estimation quality. The main purpose of the study is to analyze and validate that the Method for Holistic Optimization is feasible to be used in case a decision-maker uses a reduced database. This can be a significant advantage compared with other methods. The study it is performed using an instance database which was artificially generated in case of a turning process. The obtained results are consistent and promising.

Keywords: Decision making / method for holistic optimization / instances database / comparative evaluation / turning process

1 Introduction

In the actual global economy, the manufacturing processes plays a key role and quick decisions are requested. To be in the market in the highly competitive global production environment, companies must be able to design, manufacture and deliver products in conformance with customers request related safety, quality and delivery time. The transition from past production (mass and series) to customization production becomes a challenge for the management of the companies. Another important challenge is to obtain a low product costs and shorter product life cycles (shorter lead time).

Each manufacturing stage for a product involve specific activities and actions which must be correlated. The management of the manufacturing processes means to design, implement and control these activities and actions in order to obtain a product in an efficiently and effectively manner of usage of the necessary resources as: raw material, time, energy and personnel requested for manufacturing of a product or a couple of products.

The management of production processes means the organization of the activities/actions by making decisions accordance with a predefined scope. Also, the management means the control of the processes in terms of their expected results. For example, considering a cutting process to compensate the occurred deviations of a parameter or a set of parameters, the reference it is changed with a value equal with the difference between the target values and the predicted values of the deviations. References set-up it is performed during the programming of the production system, so this set-up is also a decisions-making act.

In both management options (process/activity) occurs a need for evaluation and this evaluation is required whenever: (i) an analysis "*what if*" is performed to decide which alternative to be used in manufacturing process case and (ii) the characteristics of the task that have the greatest impact on the effect (objectives) must be determined in order to effectively control the manufacturing process.

* e-mail: gabriel.frumusanu@ugal.ro

Finding a model for the considered process to be used for evaluation is a matter of general interest. However, such a model can be complicated by involving many variables, so finding it becomes a difficult job. Additionally, the model applicability is limited – when the premises on which the model was determined modify the model may become useless or, at least – inaccurate.

Model construction involves two stages: (i) establishment of the model structure, which means, first of all, the selection of the condition-variables by which the result-variable can be evaluated and (ii) model formalization (through the concrete relation linking the result-variable to the condition-variables) – for example, starting from a parametric model, the parametric values are adjusted until the model properly expresses, in a quantitative way, the causal link.

Many techniques for performing the second stage are available in dedicated literature, [1–12]. Papers addressing the first stage can also be highlighted, based on different features selection techniques, e.g. [13].

The optimal management of the manufacturing process and, implicitly, the making of optimal decisions can be done by applying a new optimization method, developed exclusively for application in the case of this type of process, namely the holistic optimization [14]. Holistic optimization generally involves using as process model the history of its operation under similar conditions, in the form of a database.

This paper target is to demonstrate the ability of the method for holistic optimization, in general, and of the causal link identification algorithm (as an essential stage of the method application), in particular, to provide convincing results when using a previous cases database of relatively small size.

The paper is structured as follows: the next section presents the method for holistic optimization and the stages of its application, together with the specific actions. The third section describes the methodology for simulating the application of the causal identification algorithm for a small database. The fourth section is dedicated to results and discussions related to the topic of this study. The last section gives the conclusions.

2 Method for holistic optimization – MHO

In holistic optimization, the optimization request format it is not predefined. In fact, the desiderate formalization is part of the optimization problem solving.

In manufacturing, the managerial policy imposes the desideratum concerning the process. This can be different for different products. Moreover, the desideratum can change over time even for the same product. At the same time, the desideratum reaching can be evaluated according to various criteria, specific objective functions (result-variables) can be assigned for each criterion, and for evaluating such a function different set of arguments (independent condition-variables) can be used. For this reason, the presented method requires this stage for identifying the potential goals, criteria, functions and arguments, among which the most suitable ones will be selected, according to method algorithm presented below [15].

The optimal decisions about the manufacturing process should be made based on process models. A process model generally means the relationship between a considered result-variable and a set of job descriptors (condition-variables). Usually, due to the complexity of the problems, such a model is neither unique nor precisely defined; thus, more or less descriptors may be considered, in different combinations, for the same result-variable.

The proposed MHO consists in successive performing the following loop of actions (see Fig. 1).

In what concerns the choice of the most suitable arguments (job result-variables), this can be done by instance based causal identification of the manufacturing system [14], while the comparative evaluation between two or more typical jobs can be realized after the values of their result-variables, according to the method presented in [16].

The causal identification algorithm can be used to find the most suitable structures for the model of a certain manufacturing process. It aims to identify the sets of variables with potential application in manufacturing process modeling.

The main objective in the development of the algorithm is to allow the selection of the most influential, easy to measure and with as few as possible variables, such as the resulting model has the lowest complexity, according to the required level of estimation accuracy.

The method uses the past instances related to the manufacturing system, registered as a database, to reveal the causal link between the variables that characterize the process ongoing on the considered manufacturing system.

The finality of algorithm application is the elaboration of the causal links graph, which can be considered a Decision Support System (DSS) [17]. The causal identification algorithm works on the base of the existing information, by processing a database associated with the manufacturing process (Instances-based learning, IBL, [18]) and involves going through several successive stages.

The specific actions from these stages are: (i) process identification, (ii) data concatenating, (iii) instances comparing, (iv) variables evaluation and (v) causal models' identification (see Fig. 2).

The specific actions to be performed at each stage are presented below.

1. *Process identification* – in this step the manufacturing process input and output are analyzed in order to identify and select which variables characterize and impact the process results. A set of variables with potential in process modelling are selected it is defined. The variables are classified as condition variables and result-variable.

2. *Data concatenating* – for the selected manufacturing process a data base with previous cases it is generated. The same type of activity can be characterized by more cases using the same condition-variables and result-variables. Three actions are necessary to be performed to concatenate the data namely: clustering, updating and homogenization [14].

3. *Instances comparing* – the main idea to identify the causal models is to search the relations between the variations of condition and result-variables (symbolized with x_i and y_i), instead to see each instance as an event

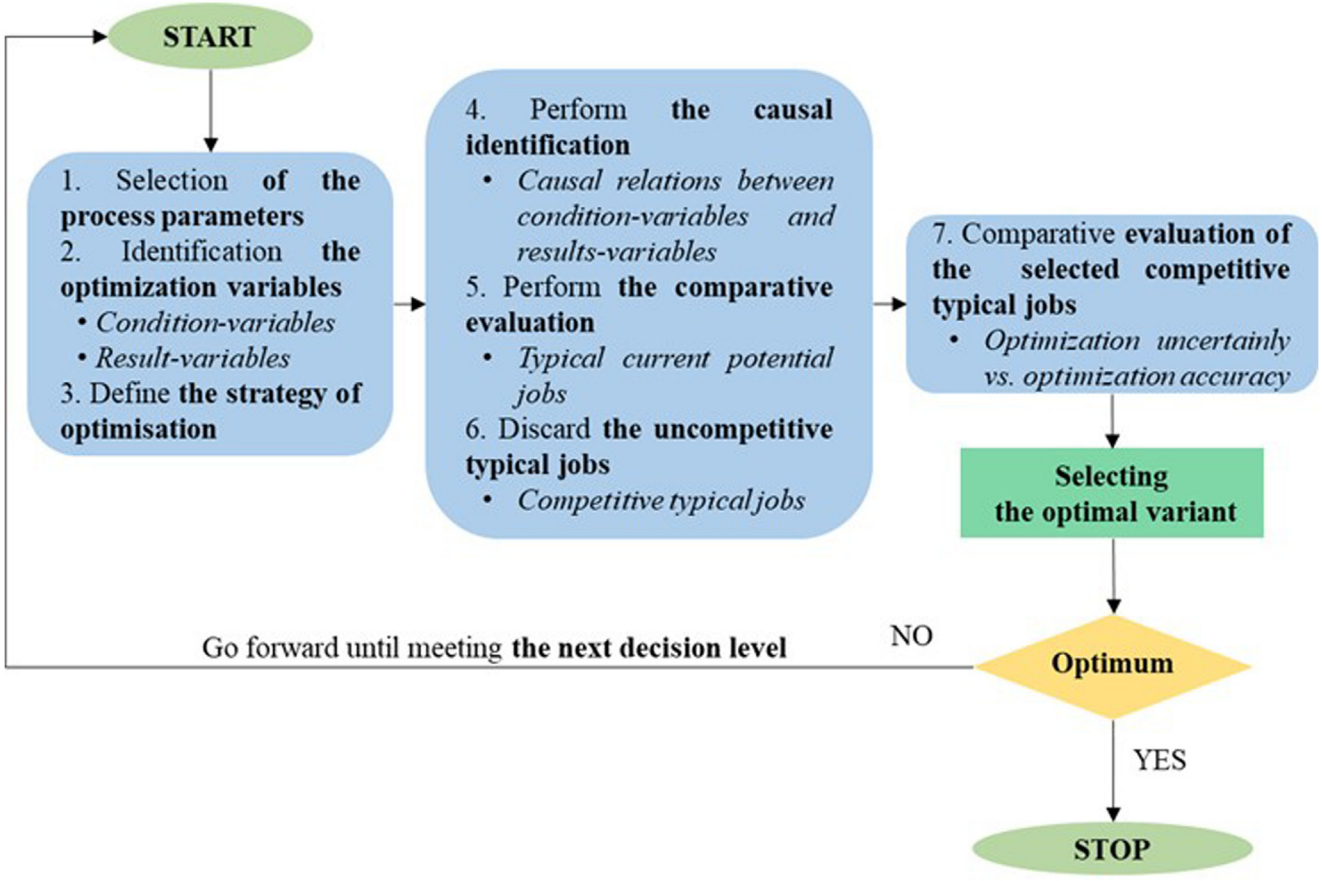


Fig. 1. Flow diagram of the MHO.

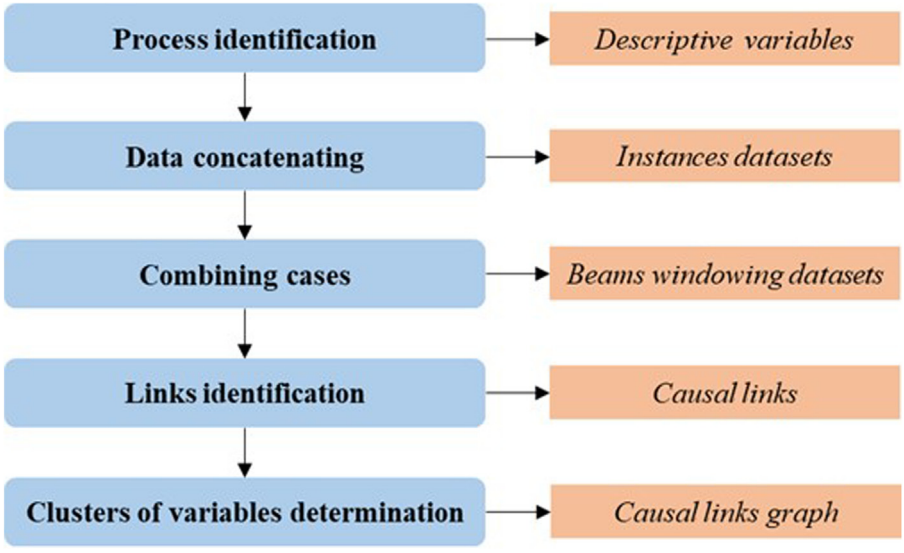


Fig. 2. Causal identification algorithm.

which show the causal relation between variables [14]. The variations of the variables can be obtained through the instances comparison.

The comparison of k^{th} and l^{th} instances from a certain dataset means to compute the differences $\delta x_i(k,l)$ and $\delta y_j(k,l)$ between their corresponding variables:

$$\begin{aligned}\delta x_i(k,l) &= |x_{ik} - x_{il}|, i = 1....n_x \text{ and } \delta y_j(k,l) \\ &= |y_{jk} - y_{jl}|, j = 1...n_y. \end{aligned} \quad (1)$$

In relation (1) n_x is the symbol used for condition-variables and n_y the symbol for result-variables. The comparison result will be further named as $beam(k,l)$. The beams consists in the reunion of the vectors $\delta x_i(k,l)$ and $\delta y_j(k,l)$. Hereby, the $beams(k,l)$ includes beam components more explicit, n_x condition-components and n_y result-components.

The instances and the beams obtained by their comparison have identical dimension and similar structure. Thus, from instances as (x_1, x_2, ..., x_i, x_{nx}, y_1, y_2, ..., y_j, ..., y_{ny}) result beams of the same structure, (δx_1, δx_2, ..., δx_i, δx_{nx}, δy_1, δy_2, ..., δy_j, ..., δy_{ny}). Considering this reason, it will be made a natural correspondence between the condition-variable and condition-component and also between result-variable and result-component. Obviously, each instance from the n composing the dataset can be compared to all other n–1. The beams dataset it is built by ensemble of beams resulting after the comparison of all possible cases.

Because the beam (k,l) and beam (l,k) are identical (with $k,l = 1, ..., n$) only one of them it is registered. Hereby, the beams dataset has $N = C_n^2$ lines.

4. *Variables evaluation* – the scope of this step is to evaluate the dependency relationship between the condition-variables and result-variables.

The evaluation method consists by successive application of two procedures:

– *The procedure for dimensionality reduction* – this is performed in order to eliminate the condition-variables with dependence on other condition-variables;
– *The procedure for evaluation the modeling potential* of each remaining condition-variable.

The results of the first procedure application is the condition-variables maximal cluster. Starting from this point, based on the values of the specific characteristics that characterize the condition-variables in terms of their modelling capacity a sub-cluster of the maximal cluster ca be generated (simply called clusters) [14].

5. *Causal models identification* – the use of the modelling potential characteristics defined in the previous step can be extended to the case of the variable clusters, after the necessary adaptions have been made. The case of a causal model with maximal cluster having y_{mc} condition-variables, v_1, v_2, ..., v_{nmc}. In principle this cluster should have, at least the highest potential of modelling the result-variable y. However, they might encountered situations when the values for one of more of cluster variables are not available, or, as well, it might be useless a complicated model, involving all variables from maximal cluster. In both cases, the solution is to use a causal model defined by fewer condition-variables. This can be realized by successively and repetitively applying a couple of algorithms [19] namely: *i) algorithm for generation of the smaller clusters* and *ii) algorithm for evaluation of the modelling potential of cluster.*

i) The algorithm for generation of the smaller clusters

Let us suppose we must deal with a cluster with x_c condition-variables (which may be in particular the maximal cluster, when $x_c = x_{mc}$). Any of them might be discarded to obtain a cluster with ($x_c - 1$) variables, hence x_c clusters may result. If now, from each smaller cluster we discard another condition-variable, the total number of distinct clusters with ($x_c - 2$) condition-variables that could be obtained is $x_c(x_c\text{-}1)$. Obviously, after only few steps of generating smaller clusters by discarding variables one by one, a very large number of clusters will result, which complicates very much the problem of assessing the potential for all of them. A reasonable solution is to consider only a part of the possible eliminations, more specific – to discard, at each level, only the condition-variables with lower modeling potential. The algorithm applied in this purpose, has three steps:

– Each of the x_c condition-variables is analyzed after a selected criterion for assessing the modeling potential:
  - *The modeling power*; c_1, which shows how much the cause-variable variation is found in the effect-variable variation.
  - *The modeling capacity,* c_2, meaning the measure which the cause-variable is able to describe the effect-variable, itself only and
  - *The modeling unevenness, RMSE,* reflecting the variability of the relation between cause- and effect-variables.
– The number of condition-variables to be discarded is established in concordance to the exigencies of the addressed modeling problem.
– After finding the x_d condition-variables with lowest modeling potential, x_d clusters with ($x_c - 1$) variables are generated by discarding them separately, one by one.

ii) The algorithm for evaluation of the modelling potential of cluster

For evaluating the modeling potential of a variables cluster, a specific algorithm has been developed [19]. The algorithm purpose is to assess the potential of a given cluster of cause-variables for modeling the considered result-variable. The application of criteria defined in previous subsection (I_1, I_2 or I_3) can be extended from assessing condition-variables to assessing clusters of condition-variables, in what concerns their modeling potential, after making the needed adaptations. In the case of a cluster, the values of criteria (denoted by I_1, I_2 or I_3).

6. Causal links graph – in this step the causal links are depicted in graphs. The graphs it is the representation of the causal models concerning the same result-variable. The representation shows the value of a criterion evaluation of the modelling potential for each variables cluster.

The causal links graph is a graph-type representation of the set of result-variables (see Fig. 3), drawn according to the following rules:

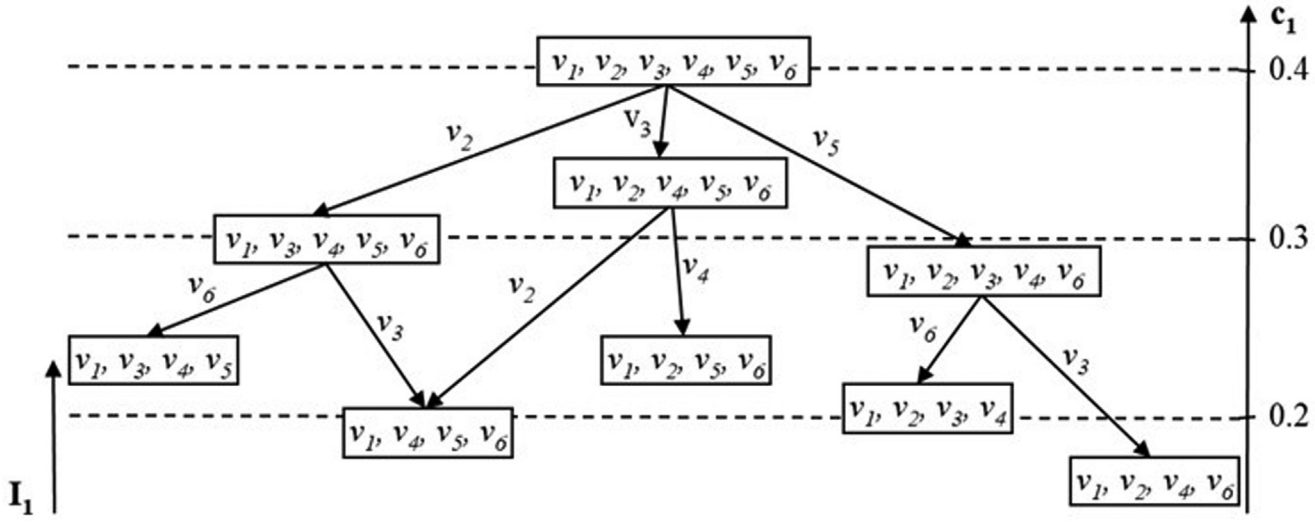


Fig. 3. Causal links graph.

Table 1. Values of Δ_i' images dimension.

Condition-variables	Successive steps of dimensionality reduction (Database for 150 cases)				
	Step 1	Step 2	Step 3	Step 4	Step 5
v_1	0.9333	0.9333	0.9333	0.9333	0.9333
v_2	0.7889	0.7889	0.7889	0.7889	0.9056
v_3	0.2036	0.8409	0.8409	0.8409	0.8409
v_4	0.7611	0.7611	0.7611	0.7611	0.7611
v_5	0.2078	0.2078	0.3278	0.3278	0.3600
v_6	0.4084	0.4084	0.4084	0.4084	0.4084
v_7	***0.2036***	–	–	–	–
v_8	0.2043	***0.2043***	–	–	–
v_9	0.3842	0.3842	0.3851	0.3851	0.3851
v_{10}	0.2050	0.2050	***0.2050***	–	–
v_{11}	0.2385	0.2385	0.2385	***0.2385***	–

Condition-variables	Successive steps of dimensionality reduction (Database for 50 cases)				
	Step 1	Step 2	Step 3	Step 4	Step 5
v_1	0.8037	0.8444	0.8444	0.8444	0.8444
v_2	0.7167	0.9056	0.9056	0.9056	0.9056
v_3	0.3192	0.3192	0.3192	0.7197	0.7197
v_4	0.7800	0.7800	0.9189	0.9189	0.9189
v_5	0.3278	0.3278	0.3278	0.3278	0.6001
v_6	0.4535	0.4535	0.4535	0.4535	0.4535
v_7	0.3192	0.3192	***0.3192***	–	–
v_8	0.3240	0.3240	0.3240	***0.3240***	–
v_9	0.5849	0.5849	0.5849	0.5849	0.5849
v_{10}	0.2661	***0.2661***	–	–	–
v_{11}	***0.2066***	–	–	–	–

Table 2. The values of c_1, c_2 and RMSE for 150 and 50 cases.

Clusters variables	Database for 50 cases			Database for 150 cases		
	c_1	c_2	RMSE	c_1	c_2	RMSE
$[v_1, v_2, v_3, v_4, v_5, v_6, v_9]$	0.2995	0.0167	0.0049	0.5051	0.0163	0.0040
$[v_1, v_2, v_4, v_5, v_6, v_9]$	0.2717	0.0243	0.0029	0.5110	0.0217	0.0086
$[v_1, v_2, v_3, v_5, v_6, v_9]$	0.2596	0.0291	0.0035	0.4800	0.0227	0.0032
$[v_1, v_2, v_3, v_4, v_6, v_9]$	0.3072	0.0159	0.0068	0.4923	0.0116	0.0032
$[v_1, v_2, v_5, v_6, v_9]$	0.2576	0.0299	0.0030	0.4510	0.0382	0.0046
$[v_1, v_2, v_4, v_6, v_9]$	0.2841	0.0232	0.0053	0.4954	0.0166	0.0076
$[v_1, v_2, v_3, v_6, v_9]$	0.2801	0.0251	0.0038	0.4657	0.0180	0.0031
$[v_1, v_2, v_3, v_4, v_6]$	0.2818	0.0237	0.0050	0.4769	0.0235	0.0035
$[v_1, v_2, v_6, v_9]$	0.2969	0.0217	0.0042	0.3654	0.0641	0.0045
$[v_1, v_2, v_5, v_9]$	0.2332	0.0310	0.0019	0.3879	0.0472	0.0011
$[v_1, v_2, v_4, v_6]$	0.2543	0.0334	0.0052	0.3114	0.0730	0.0054
$[v_1, v_2, v_3, v_9]$	0.2436	0.0301	0.0040	0.3818	0.0453	0.0033
$[v_1, v_2, v_3, v_6]$	0.2304	0.0400	0.0012	0.4275	0.0197	0.0031

– The cluster of each causal model is represented as rectangle, inside which its condition-variables are mentioned.
– *The maximal cluster* represents the starting point. The arrow drawn between two rectangles shows that the second cluster results from the first one by discarding the variable whose symbol is mentioned near the arrow.
– The level (height) of representing a certain cluster shows the values of selected criterion (I_1, I_2, I_3, or a weighted combination of them).

3 Methodology for simulating the application of the causal identification algorithm for a low-dimensional database

Within this chapter, the applicability of the causal links identification algorithm among variables that describe the turning process of a cylindrical part, using a smaller artificially generated instances database is being evaluated.

The study was performed by comparing the results of the algorithm application in two cases: a database with 150 instances, and another with 50 instances. The causal identification in the case of the database with 150 lines has already been done and the results presented in the paper [14]. For the application of the algorithm in the addressed case (database with 50 lines) the steps presented in Figure 2 were followed.

The following set of 11 condition-variables was considered:

v_1–turned part length L [mm] and v_2–diameter D [mm], v_3–required level of part accuracy A [mm], v_4–machinability of part material M [mm], v_5–rigidity R [mm], v_6–cutting speed v [m/min], v_7–feed s [mm/rot], v_8–cutting depth t [mm], v_9–main cutting force F [daN], v_{10}–power absorbed by lathe P [kW], v_{11}–removed chips volume V [cm^3] and 3 result-variables: v_{12}–machining cost C[EURO], v_{13}–machining timespan TS [min] and v_{14}–consumed energy E [kWh]. The values for the first 2 variables were chosen in the range of variation [30, 300] and [20, 200], the next 3 variables take conventional values in the range 1 to 10. Starting from here, the values for the other 6 condition-variables were calculated with:

$$t = \frac{5.1 \times R - 0.1 \times A}{10}\,[\text{mm}], \tag{2}$$

$$s = \frac{4.4 - 0.4 \cdot A}{10}\,[\text{mm/rot}], \tag{3}$$

$$v = \frac{C_v}{s^{0.3} \cdot t^{0.2} \cdot T^m}\left(\frac{10}{M} \cdot x_v + \frac{R}{10} \cdot y_v\right)[\text{m/min}], \tag{4}$$

$$F = C_F \cdot s^{0.8} \cdot t\left(x_F + \frac{M}{10} \cdot y_F\right)[\text{daN}] \tag{5}$$

$$P = \frac{F \cdot v}{6000} \cdot \frac{1}{\eta}\,[\text{kW}], \tag{6}$$

$$V = \frac{\pi \cdot D \cdot L \cdot t}{10^3}\,[\text{cm}^3]. \tag{7}$$

In relations (4) and (5) C_v, x_v and y_v, respective C_F, x_F and y_F means constants to which the values are given. Based on these, values were calculated for C, TS and E:

$$C = \frac{V}{v \cdot s \cdot t}\left[\left(1 + k + \frac{\tau_{sr}}{T}\right)c_\tau + \frac{\tau_{sr} \cdot c_\tau + c_s}{T} + \frac{P \cdot c_e}{60}\right][\text{Euro}], \tag{8}$$

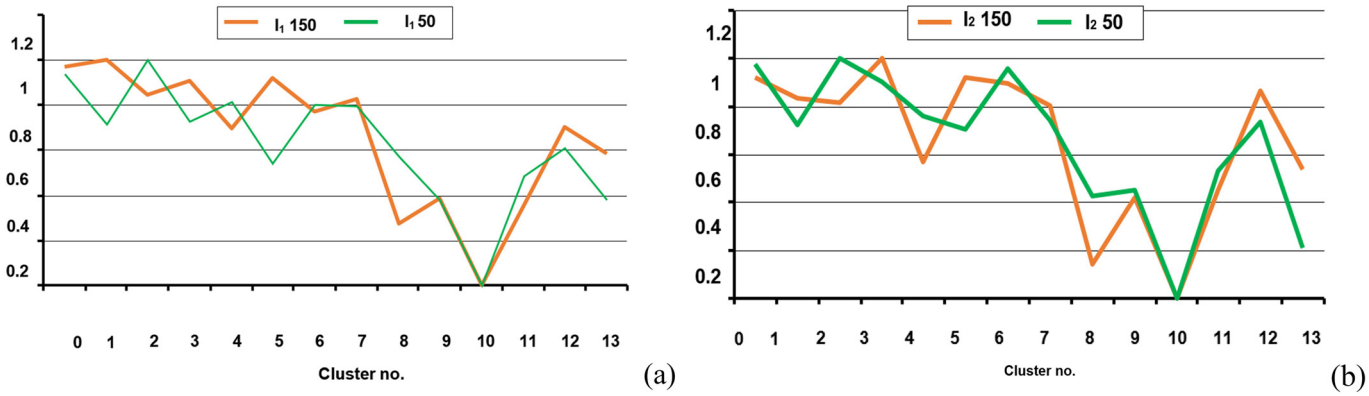


Fig. 4. Comparison between the values of the indicators obtained for the data sets with 150 and 50 cases respectively: (a) Modeling power (I_1) and (b) Modeling capability (I_2), [20].

$$TS = \frac{V}{v \cdot s \cdot t}\left(1 + k + \frac{\tau_{sr}}{T}\right)[min], \tag{9}$$

$$E = \frac{P \cdot V}{v \cdot s \cdot t} \cdot \frac{1}{60}[\text{kWh}]. \tag{10}$$

In the relations from above, η means the energy efficiency of the lathe, k – the ratio between the auxiliary time and the machining time, τ_{sr} – the time for worn tool changing [min], T – tool durability [min], c_τ – the wage specific cost [Euro/min], c_s – the tool expenditure [Euro], c_e – the energy price [Euro/kWh].

Using the above formulas, the database of 150 cases was artificially generated at first [14]. Here, 50 of them were randomly selected. The comparative study of the causal identification algorithm application in both cases was performed to reveal the effect of using a smaller database on the algorithm performance.

4 Results and discussion

The steps of the causal identification algorithm were followed (see Fig. 2). The values in each of the 14 columns (assigned to the 11 condition-variables, and 3 result-variables) were generated and scaled separately in the range [0, 1]. In the case of the database with 50 lines a combination of $N' = C_{50}^2 = 1225$ lines (beams) results.

In the case of the database with 50 lines for the causal link identification stage, the reference threshold was set at $h_{ref} = h_5 = 0.3277$. The values obtained for Δ_i' using the same MatLab application that was used for the entire database, are shown in Table 1. As it can be seen, $\Delta_{min} = 0.2066$ corresponds to the variable V, therefore it can be eliminated. At step 2, the action from previous step is repeated for the remaining ten condition-variables and another one is discarded, namely P, and so on. After step 5, $\Delta_{min} = 0.4535 > h_5$, so the seven condition-variables remaining until here can be considered relative independent and the maximal cluster is [v_1, v_2, v_3, v_4, v_5, v_6, v_9], the same maximal cluster as when using the entire database.

One can notice that the actually independent condition-variables (the first five from Tab. 1) retrieve themselves all in the maximal cluster, which confirms what it was known from the very beginning (when the artificial instances database has been built) and proves the reliability of the proposed method. Another important remark is that only 7/11 condition-variables remained for modelling the result-variables, which means a significant ease of the modelling problem.

Table 2 shows the results obtained after the causal links identification stage, where the lines containing sets other than those resulting from the data set with 150 cases [14] are shaded in grey. The same clusters were evaluated by implementing the specific algorithm based on both data sets.

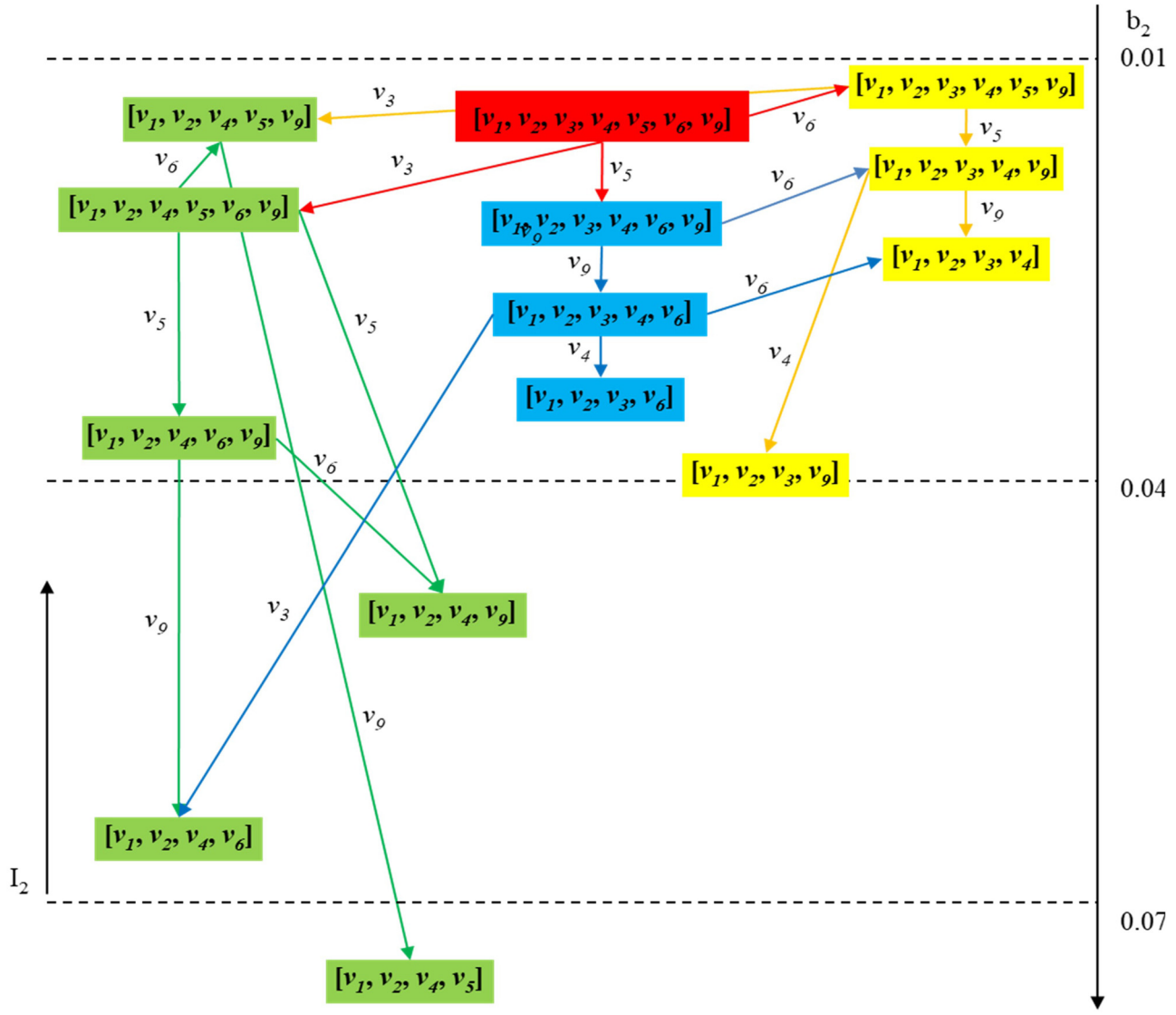
The modelling potential of a condition-variables belonging to a given cluster is evaluated with one of the criteria: (i) *the modelling power indicator,*I_1, which shows how much the condition-variable variation is found in the result-variable variation and (ii) *the modelling capacity indicator,* I_2, meaning the measure on which the condition-variable can describe the result-variable. The resulting values for I_1 and I_2 are shown in Figure 4.

Based on the above results the causal links graph (see Fig. 5) was drawn-up. The causal links graph shows the causal models concerning the same result-variable.

The graphs identified for the causal links are depicted in Figure 5. The graphs for the database with 150 cases are showed in Figure 5a and the graphs related to the database with 50 cases in Figure 5b. After analysis of the two causal links graphs showed in Figure 5, it can be concluded that the hierarchy of sets is identical or similar in both studied cases.

Following the algorithm application of the 50 cases dataset, the following observations can be made:

– The same maximum cluster results after the application of the dimensionality reduction algorithm [v_1, v_2, v_3, v_4, v_5, v_6, v_9].
– Most sets of variables (about 2/3) have the same composition in both cases.



(a)

Fig. 5. Causal links graph drawn after indicator I_2 for 150 cases (a) and 50 cases (b).

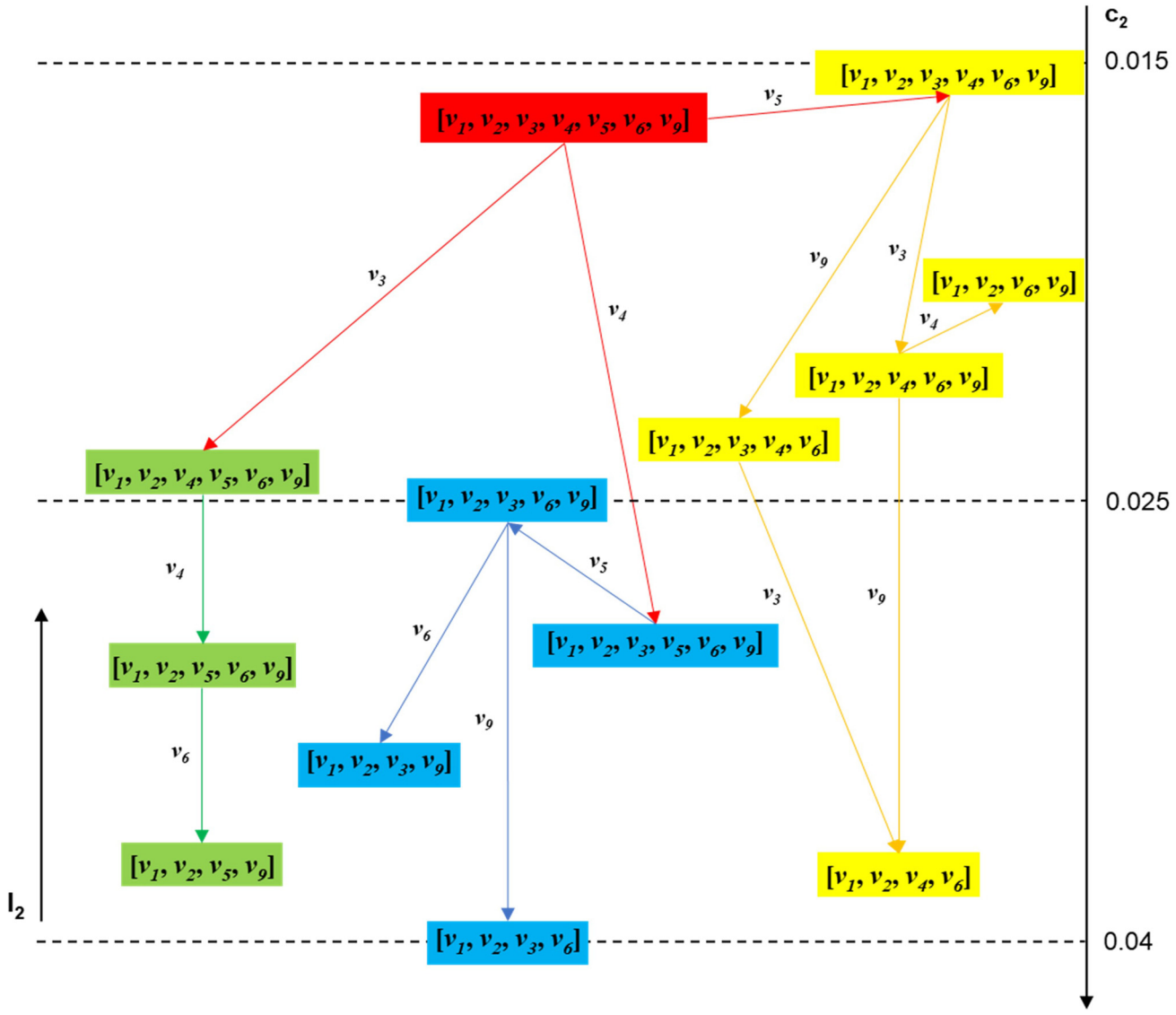
– The values of the criteria for evaluation the ability to model clusters are different in some cases from those obtained from the extended database, but the monotony of the poles of the lines in Figure 4 is the same, as are the clusters with extreme behavior.

Despite the low number of cases, one can conclude that the MHO method works with satisfactory results even when the information on the manufacturing process to be modeled is not (very) consistent.

5 Conclusions

Considering the results of the presented work the following conclusions were identified:

– The results obtained in implementing the MHO in the addressed case are showing reliability, efficiency in application and a high potential for solving diverse practical problems in manufacturing field optimization.
– The causal identification algorithm also works with satisfactory results when using a database with a smaller number of cases (50 versus 150).
– In both cases studied (50 and 150 data respectively), following the causal identification algorithm application, the same maximum cluster is obtained [v_1, v_2, v_3, v_4, v_5, v_6, v_9].
– MHO is proving to be a viable alternative to causal modeling for NN modeling methods, which have the disadvantage that their operation is problematic when a small amount of information is available.



(b)

Fig. 5. (Continued).

– The MHO application accuracy improves on its own with each new case added to the database, as its size increases.
– The MHO can be used in companies in case of some analysis to identify the manufacturing feasibility of a product or in case of a homologated process to identify the impact and causal relation of process parameters in manufacturing costs.
– Also, an application in case of auxiliary variables (such as maintenance cost, down time, quality level, process capability) which influence the process results in terms of cost could introduced and tested.

Funding

The publication of this research has been supported by the French Association of Mechanics (AFM).

Conflicts of interest

The authors declare no conflict of interest.

Data availability statement

The research data associated with this article are included within the article.

Author contribution statement

Conceptualization, C.A., G.F. and M.A.; Methodology, C.A. and G.F; Software, C.A.; Validation, C.A., G.F. and M.A.; Formal analysis, G.F.; Investigation, M.A.; Resources, M.A.; Data curation, M.A.; Writing – Original Draft Preparation, C.A., G.F. and M.A.; Writing – Review & Editing, C.A., G.F. and M.A..

References

[1] E. Assidjo, B. Yao, K. Kisselmina, D. Amané, Modeling of an industrial drying process by artificial neural networks, Braz. J. Chem. Eng. **25**, 515–522 (2008)

[2] E. Tafazzoli, M. Saif, Application of combined support vector machines in process fault diagnosis, Proc. Am. Control Conf., 3429–3433, Publisher: IEEE, St. Louis, MO, USA (2009)

[3] M. Deja, M. Siemiatkowski, Machining process sequencing and machine assignment in generative feature-based CAPP for mill-turn parts, J. Manuf. Syst. **48**, 49–62 (2018)

[4] N. Rehman, Data mining techniques methods algorithms and tools, Int. J. Comput. Sci. Mob. Comput. **6**, 227–231 (2017)

[5] P. Denno, C. Dickerson, J.A. Harding, Dynamic production system identification for smart manufacturing systems, J. Manuf. Syst. **48**, 1–11 (2018)

[6] R. Corne, C. Nath, M. Mansori, T. Kurfess, Study of spindle power data with neural network for predicting real-time tool wear/breakage during inconel drilling, J. Manuf. Syst. **43**, 287–295 (2017)

[7] S.B. Kotsiantis, Supervised machine learning, a review of classification techniques, Informatica **31**, 249–268 (2007)

[8] W. Su, M. Bo, Ant colony optimization for manufacturing resource scheduling problem, IFIP Int. Federat. Inf. Process. **207**, 863–868 (2006)

[9] Y. Song, J. Huang, D. Zhou, H. Zha, C.L. Giles, Informative K-nearest neighbor pattern classification, knowledge discovery in databases: PKDD 2007, in *Lecture Notes in Computer Science*, vol. **4702** (2007). pp. 248–264

[10] Z.M. Bi, L. Wang, Optimization of machining processes from the perspective of energy consumption: a case study, J. Manuf. Syst. **31**, 420–428 (2012)

[11] M. Rogalewicz, M. Piłacinska, A. Kujawinska, Selection of data mining method for multidimensional evaluation of the manufacturing process state, Manag. Prod. Eng. Rev. **3**, 27–35 (2012)

[12] R. Sika, Z. Ignaszak, Data acquisition in modeling using neural networks and decision trees, Arch. Foundry Eng. **11**, 113–121 (2011)

[13] Feature selection, https://en.wikipedia.org/wiki/Main_Page last accessed 2022 /05/10

[14] G.R. Frumusanu, C. Afteni, A. Epureanu, Data-driven causal modelling of the manufacturing system, Trans. Famena. **45**, 43–62 (2021)

[15] C. Afteni, G.R. Frumusanu, A. Epureanu, Method for holistic optimization of the manufacturing process, int. J. Model. Optim. 9, 265–270 (2019)

[16] C. Afteni, G.R. Frumusanu, A. Epureanu, Instance-based comparative assessment with application in manu-facturing, IOP Conf. Ser.: Mater. Sci. Eng. **400**, 1–8 (2018)

[17] Decision support system, https://en.wikipedia.org/wiki/Main_Page, last accessed 2022 /05/10

[18] Instance-based learning, https://en.wikipedia.org/wiki/Main_Page, last accessed 2022 /05/10

[19] C. Afteni, Holistic optimization of manufacturing process, PhD Thesis, 'Dunarea de Jos' University of Galati, Series I 4: Industrial Engineering (2020)

[20] C. Afteni, M. Afteni, G.R. Frumusanu, Study on the application of the holistic optimization method of the manufacturing process in the case of a reduced instances database, MATEC Web Conf. **368**, 1–10 (2022)

Cite this article as: C. Chivu, M. Afteni, G.-R. Frumusanu, Method for holistic optimization of the manufacturing process numerically described as low-dimensional database, Mechanics & Industry **25**, 17 (2024)