Gabriel-Radu Frumusanu ✉
Cezarina Afteni
Alexandru Epureanu

https://doi.org/10.21278/TOF.451020920
ISSN 1333-1124
eISSN 1849-1391

DATA-DRIVEN CAUSAL MODELLING OF THE MANUFACTURING SYSTEM

Summary

In manufacturing system management, the decisions are currently made on the base of 'what if' analysis. Here, the suitability of the model structure based on which a model of the activity will be built is crucial and it refers to multiple conditionality imposed in practice. Starting from this, finding the most suitable model structure is critical and represents a notable challenge. The paper deals with the building of suitable structures for a manufacturing system model by data-driven causal modelling. For this purpose, the manufacturing system is described by nominal jobs that it could involve and is identified by an original algorithm for processing the dataset of previous instances. The proposed causal modelling is applied in two case studies, whereby the first case study uses a dataset of artificial instances and the second case study uses a dataset of industrial instances. The causal modelling results prove its good potential for implementation in the industrial environment, with a very wide range of possible applications, while the obtained performance has been found to be good.

Key words: *manufacturing system, causal modelling, 'what-if' analysis, instance-based learning*

1. Introduction

The decisions concerning the manufacturing system (MS) management require appropriate models. The model construction involves, in general, two stages. The first stage is dedicated to the determination of a structure based on which the future model will be built (which supposes, at first, the selection of model variables). The second stage consists of model formalization (which can be done, for example, by adjusting the parameter values for a parametric model). While a large number of techniques for accomplishing the second stage are available in literature, e.g. [1]–[5], the first stage of the model construction has been also addressed by some researchers. For example, [6] suggests a meta-heuristic method for simultaneous identification of the model structure and associated parameters for linear systems, which is achieved via a constrained multidimensional particle swarm optimization mechanism. The present paper deals with the first stage of model construction.

As a consequence of recent developments in the manufacturing field, the data related to MSs is growing in size and in complexity. This presents new challenges for real-time monitoring, diagnostics, and prognostics and it means that there is a lot of work to be done in order to gain insights from high-dimensional data of varying types. For example, [7] presents a data-driven holistic approach to fault prognostics in cyclic manufacturing processes concerning data generation, acquisition, storage, processing, and prognostics, applied in the case of a plastic injection moulding process. In [8], it has been noticed that vast amounts of time series data generated by distributed sensors have to be analysed in real-time and converted to actionable knowledge that can be fed to process monitoring, control, and optimization algorithms. An original approach to the integration of information flow in the production process is presented in [9].

Some studies have been dedicated to the selection of features of interest ('key-features') inside available databases, depending on the considered purposes. Related to this selection, a novel gradient-enhanced Kriging modelling method was developed [10]. Within the framework of this method, a balance between model accuracy and modelling efficiency can be achieved. More specifically, the influence of each input ('condition') variable on the output is estimated and ranked by applying the feature selection technique. In the manufacturing field, there are specific approaches concerning feature selection. Thus, [11] deals with finding the machining features needed to build relationships between geometry, processing technology and monitoring signals in order to provide a basis for cutting tool condition identification. In [12], nine types of time-domain features are extracted from the milling force signals, after which the sensitive features that can accurately indicate the tool wear states are selected by using the correlation analysis. The study [13], addressing the assembly process optimization, develops the causal network method, considering multi-source information to assist in identifying root causes of the dimension variation. Based on this method, the diagnosis ability is evaluated, and then the minimal feature number and optimal measurement features are selected by using a sensor optimization algorithm. A method for selecting an optimum feature subset, with a modified wrapper-based multi-criteria approach using genetic algorithms is presented in [14]. The best compound features are used to construct fault prediction models. Outside the manufacturing domain, [15] proposes a computational method for identifying informative features for predicting pro-inflammatory potentials of engine exhausts and applies a principal component regression algorithm for developing prediction models.

In close connection with the individual selection of key features, the action of finding clusters of features to underlie the control models of MSs is also of critical importance. Many studies from the data mining area present diverse clustering algorithms [16]. Among these, being closer to the subject of this paper, [17] introduces a density-based algorithm for discovering clusters in large spatial databases with noise, while [18] refers to the dynamic data assigning assessment clustering. The authors of [19] propose that the identification of key features for accurate prediction of MS outputs be accomplished by using the topological data analysis. More specifically, the Mapper algorithm [20] has been applied to two benchmark datasets for the chemical process yield prediction and semiconductor wafer fault detection, in order to capture intrinsic clusters and connections among the clusters present in the datasets, which are difficult to detect using traditional methods.

Despite the results of the above presented studies, the selection of MS features used for its modelling is still a challenge because of the following specificities:

- MS complexity is permanently increasing due to both evolution of products' features and general technological development. The number of MS features is often impressive, as instead of a single feature characterizing the outcome, there are several to look for at once (e.g. time span, cost, productivity, consumed energy etc.).
- MS content changes rapidly and frequently due to market evolution. This is further leading to changes in model variable weighting and, sometimes, the number of instances available for MS features selection is small.

- MS performance is evaluated in a wide range of situations. The evaluation accuracy varies from rough to very sharp approximation. The requested speed in evaluation may be low but also it may be very high when real-time information is expected. The number of condition variables that are available at a given moment for model construction is between few and dozens.

This challenge can be answered by overcoming the following shortcomings of actual approaches:

- Only a unique set of features is selected for MS modelling.
- The model accuracy level is not correlated to the actually required level. Moreover, a variation of this level is not at all taken into account.
- The difficulties in evaluating MS features or the possible unavailability for some of them are often ignored.
- The features selection based on small past casuistry is not always feasible.

The present paper proposes that the model structure for MS be found by causal modelling (CM). The selection of MS features suitable for its modelling is performed according to an original algorithm, which eliminates or, at least, diminishes these shortcomings. The MS is approached to as causality and this allows that on the base of the causal relationships analysis more sets of features are found, that are seen as causal models for each effect variable of interest. Thus, the user can select the causal model that is best fitting to the addressed modelling problem. In existing literature, causal models are defined as mathematical models representing causal relationships within an individual system or population [21]. According to [22], a causal model is an ordered triple < U, V, E >, where *U* is a set of exogenous variables whose values are determined by factors outside the model, *V* is a set of endogenous variables whose values are determined by factors within the model, and *E* is a set of structural equations that express the value of each endogenous variable as a function of the values of other variables in *U* and *V*.

The main benefit of the proposed MS causal modelling is that it supports the required trade-off between model accuracy and modelling efficiency, leading to the selection of the most influent, easiest to measure and smallest set of MS job features, so that the resulting model has the lowest complexity with respect to the required accuracy of evaluation. Last but not least, CM works well enough even if applied in a case when the available past casuistry is small.

The paper is structured as follows: the next section introduces the background and the algorithm based on which the MS causal modelling is performed. The third section is dedicated to validation and assessment of MS causal modelling based on two case studies addressing turning processes. The last section gives the conclusion.

2. Causal modelling of the manufacturing system

2.1 Background

The MS is described by *nominal jobs* that it can accomplish. Such a job should be defined as accomplishment of *a nominal manufacturing task* according to *a given procedure* and using *a given MS component.* For each such job, the variables describing the task, the performance, and the accomplishment procedure are defined. They are of two types: condition variables (C-variables) that are independent, and effect variables (E-variables) that are dependent.

In our approach, each job accomplishment represents *an instance*. The instances are recorded in collections, which describe MS past casuistry.

Job causal modelling consists of the selection of a number of *clusters of C-variables* for each E-variable by means of which the E-variable can be evaluated. Such a cluster defines a causal model corresponding to *the given E-variable* and represents, in fact, its structural equation. For this reason, CI output is the *causal model tree*, i.e. a representation of all causal models concerning the same E-variable.

MS causal identification refers to the finding of causal models for all nominal jobs performed by the addressed MS, as well as *the assessment of the obtained causal models*. The here proposed MS causal identification consists of applying an algorithm that supposes the following steps: data concatenation, comparison of instances, assessment of variables, and identification of models, [23].

2.2 Data concatenation

For a given nominal job, three actions are necessary in order to do data concatenation, namely, clustering, updating, and homogenization.

Clustering deals with selecting the instances referring to a given job from a collection of instances. These instances form the *instance dataset*.

Updating of variable values is necessary because of inherent changes occurring since the instances have been recorded. For example, energy and material prices may vary and the object metamorphosis costs at different times may be calculated in different terms. Thereby, the finding of causal models concerning the manufacturing costs, as E-variable, cannot be made without updating all costs to the current terms.

Homogenization aims to make the instances comparable by scaling the values of each variable to values between 0 and 1.

2.3 Comparison of instances

The core idea in finding causal models is to look for relations between the variations of C- and E-variables (denoted by p_i and $q_{j,}$ respectively), instead of viewing each instance as event that illustrates the causal relation between these variables.

Variable variations can be revealed by instance comparison. Comparing kth and lth instances from a certain dataset means to calculate the differences $\delta p_i(k,l)$ and $\delta q_j(k,l)$ between their corresponding variables:

$$\delta p_i(k,l) = |p_{ik} - p_{il}|,\ i = 1 \ldots n_p \quad \text{and} \quad \delta q_j(k,l) = |q_{jk} - q_{jl}|,\ j = 1 \ldots n_q . \tag{1}$$

In relation (1), n_p and n_q represent the number of C- and E-variables, respectively. The result of such a comparison will be further referred to as beam(k, l). It consists of the reunion of vectors $\delta p_i(k,l)$ and $\delta q_j(k,l)$. Hereby, beam(k, l) includes *beam components* (more specifically, n_p *C-components* and n_q *E-components*).

The instances as well as the beams resulting from their comparison have identical dimension and similar structure. Thus, from the instances such as $\left(p_1, p_2, \ldots p_i, \ldots p_{n_p}, q_1, q_2, \ldots, q_j, \ldots q_{n_q}\right)$ beams of the same structure $\left(\delta p_1, \delta p_2, \ldots \delta p_i, \ldots \delta p_{n_p}, \delta q_1, \delta q_2, \ldots \delta q_j, \ldots \delta q_{n_q}\right)$ are obtained. For this reason, it will be hereinafter made a natural correspondence between C-variable p_i and C-component δp_i as well as between the E-variable q_j and E-component δq_j.

Obviously, each instance from n instances composing the dataset can be compared to all other n–1. The ensemble of beams resulting from all possible comparisons represent the *beam dataset*. Because beam(k, l) and beam(l, k) are identical (with k, l = 1...n), only one of them is recorded. Hereby, the beam dataset has $N = C_n^2$ lines.

2.4 Assessment of variables

This step aims to assess the C-variables in order to find the ones having potential for evaluating the given E-variable. This is done by *beam windowing,* which is an original technique, developed for this purpose.

The elementary action in beam windowing is the *windowing sequence*, which refers to imposing restrained domains (*'windows'*) to one or more of the beam components corresponding to the analysed variables, while the domains (*'images'*), that consequently result for each of the other beam components, are measured. The ensemble of beams concomitantly passing through all considered windows will be further called *strand of beams.*

Assessment of variables is performed by applying two algorithms, namely, i) the algorithm for dimensionality reduction, and ii) the algorithm for assessing the modelling potential of variables. The output of the first algorithm is the *maximal cluster* of variables, while the second algorithm delivers the values of specific criteria characterizing the C-variables with respect to their modelling potential.

The algorithms are further introduced by presenting their application in the case of the *basic problem* when one has to deal with a set of m C-variables corresponding to a single E-variable.

Algorithm for dimensionality reduction

Let us consider a beam dataset as defined above. Obviously, all beam components have values in the interval [0, 1]. Here, a predefined series of thresholds, h_k, is considered (Fig. 1) as:

$$h_k = 0.8^k, k = 0,1,2\dots \tag{2}$$

One of these thresholds, denoted by h_{ref}, is set as reference. Its selection depends on the extension of beam dataset (e.g. $h_{ref} = h_7 = 0.2097$).

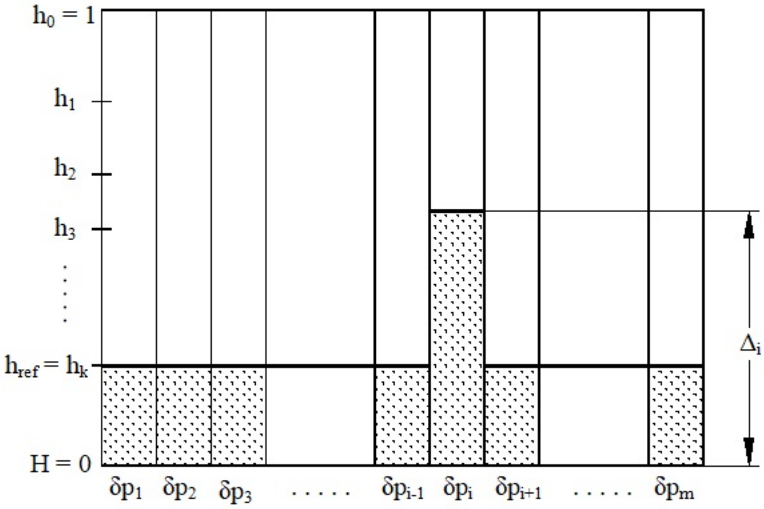

Fig. 1 Windowing sequence applied to C-component δp_i

The algorithm aims to discard the C-variables having high degree of dependence on other C-variables. It consists of the following actions, [23]:

- Windows having H = 0 and $h = h_{ref}$ are imposed to (m–1) of the C-components, excepting the i^{th} one, for which the image dimension Δ_i is measured (Fig. 1). Obviously, there are m possibilities to do this (i = 1, 2, ... m).
- The windowing sequence from above is run for each of the m C-components, hence m values of Δ_i will result.

- After that, the value of $\Delta_{\min} = \min\left(\Delta_i\right)_{i=1,2,\ldots m}$ is determined. If $\Delta_{\min} < h_{k-1}$ ($h_{ref} = h_k$), then the C-variable to which $\Delta_{\min}$ corresponds can be considered as highly dependent on other C-variables, and it may be discarded.
- The set of remaining C-variables is repetitively submitted to the previous three actions until the current value of $\Delta_{\min}$ becomes higher than h_{k-1}. The last set of C-variables forms *the maximal cluster of variables.*

Algorithm for assessing the modelling potential of variables

The modelling potential refers to C-variables and to their potential to model a certain E-variable. Three criteria, as presented below, can assess the modelling potential of a C-variable belonging to a given cluster:

- *The modelling power MP* representing the 'sensitivity' of E-variable relative to the considered C-variable;
- *The modelling capacity MC* referring to the measure in which C-variable is able to describe E-variable by itself only;
- *The modelling unevenness MU* reflecting the variability of the relation between C- and E-variables.

The algorithm, designed in order to assess the criteria from above in the case of C-variable p_i, belonging to the cluster (p_1, p_2 ... p_n), when aiming to model E-variable q, consists of the following steps, [23].

The windows having H = 0 and h = h_{ref} are imposed to (n–1) C-components corresponding to variables from the considered cluster, excepting the ith one, for which we impose a window whose dimension is successively modified from $h_0 = 1$, h_1, h_2 ... h_j, ... to $h_k = h_{ref}$. An image of E-component δq having the dimension Δ_{ij}, j = 1, 2, ... k, corresponds to h_j dimension of the window imposed to δp_i (see Fig. 2, where j = 3).

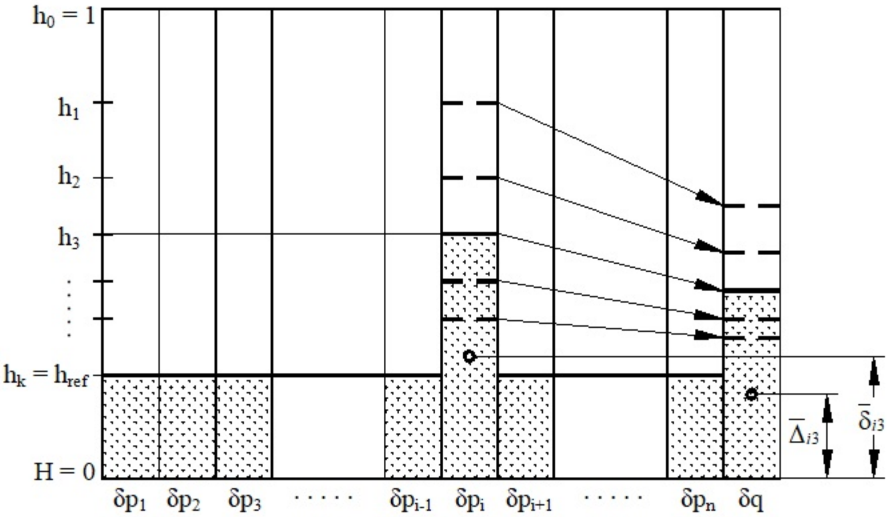

Fig. 2 Assessment of p_i modelling potential

- In the case of the window having dimension h_j imposed to δp_i component, the strand of beams passing through it and other (n-1) windows (for which $h = h_{ref}$) has the cardinal N_{ij}. Then, the average values $\bar{\delta}_{ij}$ and $\bar{\Delta}_{ij}$ are calculated as:

$$\bar{\delta}_{ij} = \frac{1}{N_{ij}} \sum_{l=1}^{N_{ij}} \delta p_{il} \,, \tag{3}$$

and $\overline{\Delta}_{ij} = \dfrac{1}{N_{ij}} \sum_{l=1}^{N_{ij}} \delta q_l$, respectively, (4)

where δp_{il} and δq_l refer to the values of δp_i and δq corresponding to the lth beam passing through the considered set of windows.

- A linear regression, having the form of:

$$y = a \cdot x + b \tag{5}$$

is fitted to the set of points $C_j\left(\overline{\delta}_{ij}, \overline{\Delta}_{ij}\right)_{j=1,2...k}$, the root mean square error (RMSE) being also calculated.

The local 'sensitivity' of a function to its argument, in a given point, is shown by the value of the function derivative, meaning the slope of the tangent to the function graphic drawn in that point. By analogy, we assume that the global 'sensitivity' of the q variable to the p_i variable may be expressed by slope a of the fitted straight line. Thus, slope a can be a metric for assessing *the modelling power MP*. At the same time, because $h_j \to 0$ (together with $\overline{\delta}_{ij}$) when $j \to \infty$, we assume that the causal relation 'strength' is reflected by the degree in which Δ_{ij} (hence $\overline{\Delta}_{ij}$ too) also tends to 0 when $j \to \infty$, this being in connection with b value. The closer to 0 it is, the 'stronger' the causal relation is. Thus, 1- b can be a metric for assessing *the modelling capacity MC*. Finally, the dispersion of C_j points relative to the fitted straight line can give relevant information on the variability of the relation between p_i and q. Thus, the RMSE of C_j points relative to the fitted straight line $y = a \cdot x + b$ can be a metric for assessing *the modelling unevenness MU.*

It should be noted that the values of these three criteria are relative because they depend on the composition of the cluster inside which they are considered. In other words, the same C-variable may show different modelling potentials if assessed in different clusters.

2.5 Identification of casual models

Let us consider the case of a causal model whose maximal cluster has n_{mc} C-variables, $p_1, p_2, ... p_{nmc}$. This should have, at least in principle, the highest potential for modelling the E-variable q. However, there might be situations when the values for one or more of cluster variables are not available, or a complicated model, involving all variables from the maximal cluster, might be useless. In both cases, the solution is to use a causal model defined by fewer C-variables. The selection of such most suitable model assumes that multiple causal models concerning q are identified. This can be realized by successively and repetitively applying a couple of algorithms [23], namely i) the algorithm for generating smaller clusters, and ii) the algorithm for assessing the modelling potential of a cluster.

Algorithm for generating smaller clusters

Let us suppose we have to deal with a cluster with n_c C-variables (which may be in particular the maximal cluster when $n_c = n_{mc}$). Any of the variables might be discarded in order to obtain a cluster with (n_c-1) variables, hence n_c clusters may result. If now we discard another C-variable from each smaller cluster, the total number of distinct clusters with (n_c-2) C-variables that could be obtained is $n_c\,(n_c-1)$. 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 the problem of assessing the potential for all of them very much. A reasonable solution is to consider only a part of possible eliminations, more specifically, to discard only the C-variables with lower modelling potential at each level. The algorithm applied for this purpose has three steps:

- Each of the n_c C-variables is analysed after a criterion for assessing the modelling potential (MP, MC or MU) has been selected.
- The number n_d of C-variables to be discarded is established in accordance to exigencies of the addressed modelling problem.
- After n_d C-variables with lowest modelling potential have been found, n_d clusters with (n_c-1) variables are generated by discarding them separately, one by one.

Algorithm for assessing the modelling potential of a cluster

The purpose of the algorithm is to assess the potential of a given cluster of C-variables for modelling the considered E-variable. The application of criteria defined in the previous subsection (MP, MC and MU) can be extended from assessing C-variables to assessing clusters of C-variables with respect to their modelling potential, after the needed adaptations have been made. In the case of a cluster, the values of the criteria (denoted by MP_c, MC_c and MU_c) are obtained by applying the following algorithm:

- The windows having H = 0 and the dimension successively modified from h_0 = 1, h_1, h_2 ... h_j ... to h_k = h_{ref} are concomitantly imposed to all n_c beam components corresponding to the C-variables from the cluster. An image of the beam component corresponding to E-component δq, having the dimension Δ_j, j = 1, 2, ... k is obtained for each dimension h_j of the window (see Fig. 3, where j = 3).

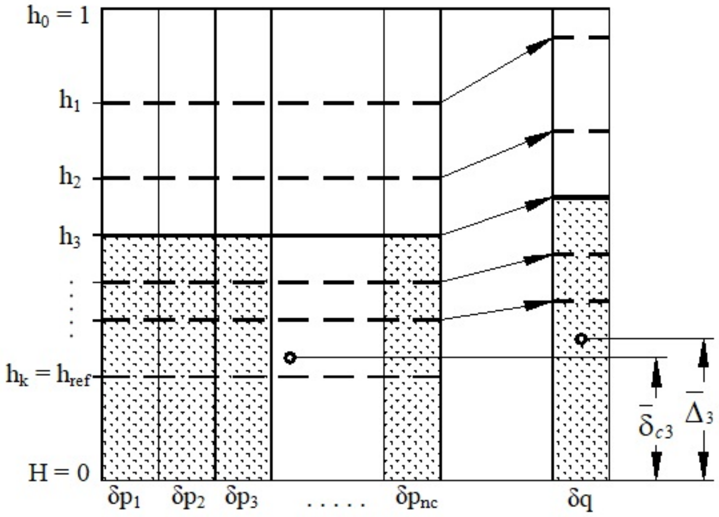

Fig. 3 Assessment of cluster modelling potential

- When the windows imposed to C-components have the dimension h_j, the strand of beams passing through them has the cardinal N_j. The average values $\bar{\delta}_{cj}$ and $\bar{\Delta}_j$ are calculated as:

$$\bar{\delta}_{cj} = \frac{1}{n_c \cdot N_j}\left(\sum_{l=1}^{N_j}\delta p_{1l} + \sum_{l=1}^{N_j}\delta p_{2l} + \ldots + \sum_{l=1}^{N_j}\delta p_{n_c l}\right), \qquad (6)$$

and $\bar{\Delta}_j = \frac{1}{N_j}\sum_{l=1}^{N_j}\delta q_l$, respectively, (7)

where δp_{il} and δq_l refer to the values of δp_i and δq corresponding to the lth beam (from the N_j) passing through the considered set of windows, i = 1, 2, ... n_c.

- A linear regression, having the form of:

$$y = A \cdot x + B \qquad (8)$$

is fitted to the set of points $\left(\overline{\delta}_{cj}, \overline{\Delta}_{j}\right)_{j=1,2\ldots k}$, the root mean square error (RMSE) being also calculated.

The values of the criteria evaluating the cluster potential are:

$$MP_c = A,\ MC_c = 1\text{-}\ B \text{ and } MU_c = \text{RMSE}. \tag{9}$$

As regards the procedure of identifying casual models, the maximal cluster represents the starting point, hence it defines the first causal model. Every time when a number of clusters are generated by reducing the dimension of the ones defining the already identified causal models by 1, the new clusters having values of a selected criterion higher than a pre-established threshold define new causal models of the q E-variable. The clusters of variables of the new models will be further submitted to the algorithms from above. The procedure of identifying new causal models is stopped when any of the smaller clusters generated does not show enough modelling potential.

As already mentioned, the tree of causal models is a representation of causal models concerning the same E-variable. The tree is constructed after a whole set of causal models concerning the addressed E-variable has been identified. The representation shows the value of the criterion assessing the modelling potential for each model and suggests the way in which this was obtained in the procedure of identifying models. It is a graph-type representation (Fig. 4), drawn according to the following rules:

- The cluster of each causal model is represented as a rectangle, inside which its C-variables are given.
- The arrow drawn between two rectangles shows that the second cluster results from the first cluster by discarding the variable whose symbol is mentioned near the arrow.
- The level (height) of the representation of a certain cluster shows values of the selected criterion (MP$_c$, MC$_c$, MU$_c$, or a weighted combination of them).

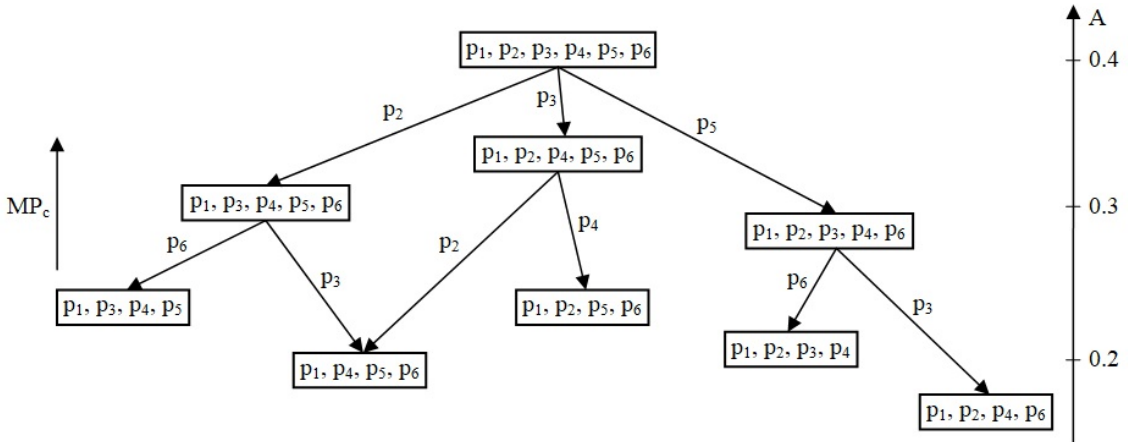

Fig. 4 Generic causal model tree associated with E-variable q, drawn by considering MP$_c$ criterion

3. Case studies

The two case studies presented below address MS jobs of turning. It is obvious that the larger the number of instances from the dataset, the more relevant the results of the case study. This is the reason why the first case study deals with a real MS job but with a dataset of artificial instances. The objectives of the case study are: (1) to demonstrate the causal modelling (CM) applicability and (2) to validate the CM results. The second case study uses a dataset of instances extracted from the industrial environment and concerns the job of turning roller bearing rings. Its objectives are: (1) to check the CM applicability in industrial conditions and (2) to assess the CM performance in the addressed industrial case.

3.1 Case study performed on the dataset of artificial instances

(1) Demonstration of CM applicability

Let us consider the nominal job of turning cylindrical parts. For its CM, 14 variables were considered, namely: length L and diameter D of the turned part, required level of part accuracy A, machinability of part material M, rigidity R, cutting speed *v*, feed f, cutting depth t, main cutting force F, power absorbed by lathe P, removed chips volume V, machining cost C, machining time span TS, and consumed energy E.

Some remarks should be made regarding these variables:

- The first five (L, D, A, M and R) are purely exogenous (independent) as their values are imposed by the part designer. The next six variables (v, f, t, F, P and V) depend on the first five variables, while the last three (C, TS and E) depend on all previous eleven variables.
- Most of the variables have clear physical meaning, so they can be expressed directly by their values. The exceptions (A, M and R) shall be expressed by conventional dimensionless values, from 1 to 10, assigned after synthesizing some features.

The relations describing the real turning process express v, f, t, F, P and V depending on L, D, A, M and R [23], as follows:

$$t = \frac{5.1 \cdot R - 0.1 \cdot A}{10} \text{ (mm)}, \tag{10}$$

$$f = \frac{4.4 - 0.4 \cdot A}{10} \text{ (mm/rot)}, \tag{11}$$

$$v = \frac{C_v}{f^{0.3} \cdot t^{0.2} \cdot T^m} \left(\frac{10}{M} \cdot x_v + \frac{R}{10} \cdot y_v \right) \text{ (m/min)}, \tag{12}$$

$$F = C_F \cdot f^{0.8} \cdot t \left(x_F + \frac{M}{10} \cdot y_F \right) \text{ (daN)}, \tag{13}$$

$$P = \frac{F \cdot v}{6000} \cdot \frac{1}{\eta} \text{ (kW)}, \tag{14}$$

$$V = \frac{\pi \cdot D \cdot L \cdot t}{10^3} \text{ (cm}^3\text{)}. \tag{15}$$

In relations (12) and (13) C_v, x_v and y_v, and C_F, x_F and y_F, respectively, refer to the constants to which values are pre-set. In relation (15), part dimensions L and D are expressed in millimetres.

The relations for calculating C, TS and E are [23]:

$$C = \frac{V}{v \cdot f \cdot t} \left[\left(1 + k + \frac{\tau_{sr}}{T} \right) c_\tau + \frac{\tau_{sr} \cdot c_\tau + c_s}{T} + \frac{P \cdot c_e}{60} \right] \text{ (Euro)}, \tag{16}$$

$$TS = \frac{V}{v \cdot f \cdot t} \left(1 + k + \frac{\tau_{sr}}{T} \right) \text{ (min)}, \tag{17}$$

$$E = \frac{P \cdot V}{v \cdot f \cdot t} \cdot \frac{1}{60} \text{ (kWh)}. \tag{18}$$

In relations from above, k refers to the ratio between the auxiliary time and the machining time, τ_{sr} – time for worn tool changing (min), T – tool durability (min), c_{τ} - wage specific costs (Euro/min), c_s - tool expenditure between two consecutive tool changes (Euro), c_e - energy price (Euro/kWh).

Variables C, TS and E have been selected as the variables to be evaluated; hence, each of them may be considered to be an E-variable; the rest of variables are C-variables.

In order to test its applicability, the modelling algorithm steps were followed as described in the previous section.

Data concatenation - the dataset of instances should actually include data collected from the addressed manufacturing system. Since here we deal with an artificial instance dataset, the clustering, updating and homogenization actions are replaced by dataset artificial generation as follows:

- Variation intervals have been established for each of the five variables L, D, A, M and R. The intervals [30, 300] and [20, 200], in millimetres, have been adopted for L and D, respectively. The conventional values of A, M and R are comprised in [1, 10] interval. Relative uniform divisions of these five intervals, composed by n = 150 points each, were adopted.
- The order of points from each of the five divisions has been separately randomized, thus resulting in the first five columns of the instance dataset with 150 lines.
- The values of v, f, t, F, P and V were calculated using formulae (10) – (15) for each set of L, D, A, M and R values.
- The values of C, TS and E have been calculated using formulae (16) – (18) on the base of the values previously found for v, f, t, F, P and V.
- The values of each variable have been separately scaled to the interval [0, 1].

The resulting artificial instance dataset has been stored as a Microsoft Excel file.

Comparison of instances - the beam dataset has been obtained with the help of a MatLab application written for this purpose by making comparisons between 150 instances from the dataset, as explained in subsection 2.3. Hereby, $\mathrm{N} = \mathrm{C}_{150}^{2} = 11{,}175$ beams compose the beam dataset, which has also been stored as a Microsoft Excel file.

Assessment of variables

<u>*Dimensionality reduction*</u>

At first, the reference threshold has been set to $h_{ref} = h_7 = 0.2097$, hence $h_{k-1} = h_6 = 0.2621$. According to the algorithm presented in 2.4, the windows having H = 0 and $h = h_{ref}$ were considered for the beam components corresponding to ten of the eleven C-variables, while for the eleventh the image dimension Δ_i was measured (i = 1, 2, ... 11, successively). The values obtained for Δ_i, by using a dedicated MatLab application are shown in Table 1, [23]. As it can be noticed, $\Delta_{min} = 0.2036$ (marked in bold) corresponds to variables A and f, hence one of them (e.g. f) may be discarded. In Step 2, the action from the previous step is repeated for the remaining ten C-variables and another variable is discarded, namely t, and so on. After Step 5, $\Delta_{min} = 0.3600 > h_6$, so the seven C-variables remaining until this point can be considered relatively independent and the maximal cluster is [L, D, A, M, R, v, F].

Table 1 Image dimensions and Δ_{min} values (in bold), [23]

Condition variable	Successive steps in dimensionality reduction				
	Step 1	Step 2	Step 3	Step 4	Step 5
L	0.9333	0.9333	0.9333	0.9333	0.9333
D	0.7889	0.7889	0.7889	0.7889	0.9056
A	0.2036	0.8409	0.8409	0.8409	0.8409
M	0.7611	0.7611	0.7611	0.7611	0.7611
R	0.2078	0.2078	0.3278	0.3278	**0.3600**
v	0.4084	0.4084	0.4084	0.4084	0.4084
f	**0.2036**	-	-	-	-
t	0.2043	**0.2043**	-	-	-
F	0.3842	0.3842	0.3851	0.3851	0.3851
P	0.2050	0.2050	**0.2050**	-	-
V	0.2385	0.2385	0.2385	**0.2385**	-

One can notice that the actually independent C-variables (the first five from Table 1) retrieve themselves all in the maximal cluster, which confirms what we knew from the very beginning (when the artificial instance database had been built) and proves the reliability of the proposed algorithm. Another important remark is that only 7/11 C-variables remained for modelling E-variables, which results in a significant ease of the modelling problem.

Assessment of variable modelling potential

The C-variable modelling potential was assessed according to the EC criterion (hence, after the values of b). This has been determined for each C-variable of the maximal cluster, according to the algorithm presented in 2.4, after considering the costs C as E-variable, with the help of *Curve fitting tool* from MatLab. The resulting values are presented in the first line of Table 2.

Identification of causal models

Smaller cluster generation

According to the algorithm from 2.5, the C-variables with lowest EC (highest values of *b*) are suitable to be discarded. After generating each set of smaller clusters, their modelling potential has been assessed in order to identify new causal models.

Three clusters with six C-variables each have been generated from the maximal cluster at the first dimension reduction by discarding one of the n_d = 3 variables having the lowest potential. Then, two clusters with five C-variables (this time, n_d = 2) resulted from each of these three at the second dimension reduction. Finally, two clusters of four C-variables (when n_d = 2) were obtained from each cluster with five variables. The clusters with three C-variables did not show enough potential for identifying any causal models, so smaller cluster generation was stopped there. The C-variables selection and the resulting clusters are presented in Table 2.

In the table, the variables having the highest value of b (the lowest *MC*) have been marked in bold. It should be noticed that because of discarding the same variables in different successions, some clusters were obtained twice in the same structure. This is the reason why there are only four (instead of six) distinct clusters with five C-variables and six (instead of 12) clusters with four C-variables. Hereby, the causal model tree will include 14 clusters.

Table 2 Smaller cluster generation, [23]

Generation of 6-variable clusters							
Variables	L	D	**A**	M	**R**	**v**	F
b	0.0269	0.0358	**0.0475**	0.0378	**0.0486**	**0.0476**	0.0437
Resulting clusters	[L, D, M, R, v, F]		[L, D, A, M, v, F]		[L, D, A, M, R, F]		
Generation of 5-variable clusters							
Variables	L	D	M	**R**	**v**	F	
b	0.0289	0.0352	0.0415	**0.0557**	**0.0492**	0.0332	
Resulting clusters	[L, D, M, v, F]			[L, D, M, R, F]			
Variables	L	D	A	M	**v**	**F**	
b	0.0244	0.0313	0.0519	0.0474	**0.0532**	**0.0601**	
Resulting clusters	[L, D, A, M, F]			[L, D, A, M, v]			
Variables	L	D	**A**	M	**R**	F	
b	0.0278	0.0375	**0.0486**	0.0427	**0.0493**	0.0444	
Resulting clusters	*[L, D, M, R, F]*			*[L, D, A, M, F]*			
Generation of 4-variable clusters							
Variables	L	D	M	**v**	**F**		
b	0.0276	0.0303	0.0456	**0.0619**	**0.0487**		
Resulting clusters	[L, D, M, F]			[L, D, M, v]			
Variables	L	D	M	**R**	**F**		
b	0.0278	0.0375	0.0474	**0.0557**	**0.0601**		
Resulting clusters	*[L, D, M, F]*			[L, D, M, R]			
Variables	L	D	A	**M**	**F**		
b	0.0284	0.0291	0.0446	**0.0507**	**0.0604**		
Resulting clusters	[L, D, A, F]			[L, D, A, M]			
Variables	L	D	A	**M**	**v**		
b	0.0358	0.0411	0.0522	**0.0554**	**0.0649**		
Resulting clusters	[L, D, A, v]			*[L, D, A, M]*			

Assessment of cluster modelling potential

After having found the clusters of C-variables that will compose the causal model tree, the values of A, B and RMSE were found with the *Curve fitting* tool from MatLab, by applying the algorithm introduced in subsection 2.5. These values are presented in Table 3.

The clusters from Table 4 define the identified causal models. The causal model tree from Figure 5 was drawn according to the criterion MC_c (which is very similar to the one drawn according to the MC_c criterion). In our view, the third modelling potential criterion, namely MU_c, is preferable to be used for the discrimination between two clusters with close MP_c or MC_c.

It is very important to notice that the causal model tree *corresponds to what we knew from the very beginning*: the models including the five independent C-variables show better potential. *As it was also expected*, the models with more C-variables prove, in general, better potential than the ones with fewer C-variables. Finally, yet importantly, the model [L, D, A, M] still has a reasonable potential, hence a significant reduction in the model dimension (4 instead of 7) can be obtained by applying the presented algorithm. All remarks from above prove the method applicability.

Table 3 Values of A, B and RMSE, [23]

Cluster variables	A	B	RMSE
L, D, A, M, R, v, F	0.5051	0.0163	0.0040
L, D, M, R, v, F	0.5110	0.0217	0.0086
L, D, A, M, v, F	0.4800	0.0227	0.0032
L, D, A, M, R, F	0.4923	0.0116	0.0032
L, D, M, v, F	0.4510	0.0382	0.0046
L, D, M, R, F	0.4954	0.0166	0.0076
L, D, A, M, F	0.4657	0.0180	0.0031
L, D, A, M, v	0.4769	0.0235	0.0035
L, D, M, v	0.3654	0.0641	0.0045
L, D, M, F	0.3879	0.0472	0.0011
L, D, M, R	0.3114	0.0730	0.0054
L, D, A, F	0.3818	0.0453	0.0033
L, D, A, M	0.4522	0.0197	0.0031
L, D, A, v	0.4275	0.0398	0.0026

Table 4 R-coefficients values, [23]

Crt. no.	Cluster variables	R-coefficient			
		Training	Validation	Test	All
1	L, D, A, M, R, v, F	0.9990	0.9985	0.9972	0.9982
2	L, D, M, R, v, F	0.9973	0.9857	0.9874	0.9872
3	L, D, A, M, v, F	0.9990	0.9897	0.9910	0.9953
4	L, D, A, M, R, F	0.9999	0.9980	0.9885	0.9962
5	L, D, M, v, F	0.9938	0.9338	0.9477	0.9751
6	L, D, M, R, F	0.9992	0.9787	0.9808	0.9917
7	L, D, A, M, F	0.9992	0.9938	0.9878	0.9944
8	L, D, A, M, v	0.9829	0.9339	0.9181	0.9697
9	L, D, M, v	0.9393	0.7560	0.9202	0.9124
10	L, D, M, F	0.9311	0.8369	0.9277	0.9164
11	L, D, M, R	0.8879	0.8058	0.7182	0.8389
12	L, D, A, F	0.9443	0.8997	0.8315	0.9176
13	L, D, A, M	0.9500	0.9531	0.8456	0.9374
14	L, D, A, v	0.9997	0.9932	0.9926	0.9977

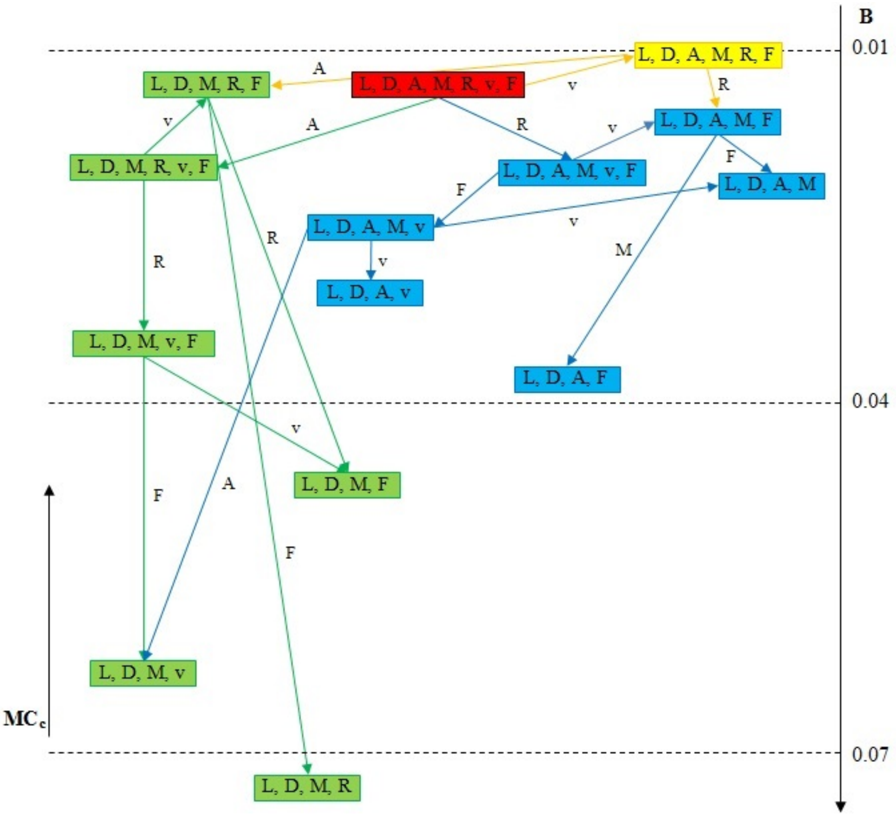

Fig. 5 Causal model tree drawn according to criterion MCc, [23]

(2) Validation of CM results

By starting from the observation that the modelling potential of a given cluster is directly related to the performance in modelling the E-variable on its base, the CM results were validated by neural networks (NN) modelling. The performance of NN models having the structure according to the found causal models has been evaluated and compared to the potential of clusters corresponding to these causal models. NN models were built on the base of the 14 causal models composing the causal tree, and the *Neural Networks* tool from MatLab was used.

The *NN*-models with two layers (the hidden layer with 10 neurons) were built on the base of the dataset of artificial instances with 150 lines (104 lines were used for training the network, 23 for validation and 23 for the testing of the model). The NN models performance was evaluated by using the values of the correlation coefficient *R* between output and target values.

In order to enable a comparison between the modelling potential of casual models and the corresponding performance of NN models, first, the values of A, 1-B (Table 3) and R (*All*, Table 4) were scaled to interval [0, 1]. Then, the three sets of values were represented versus the current number of the model (Table 4) by joining each set of resulting points into a polyline (Fig. 6).

After examining the diagrams from Fig. 6, we can draw the following conclusions:

- The profile of the three polylines is similar in terms of the general appearance and most of the trends between successive points.
- The three polylines have common points (8th and 11th point) or very close points (1st, 6th, 10th, and 12th point).
- There is an obvious grouping of the points in the domains (e.g. between 0.8 and 1, or below 0.6) showing either both good or both low modelling capacity and model performance.

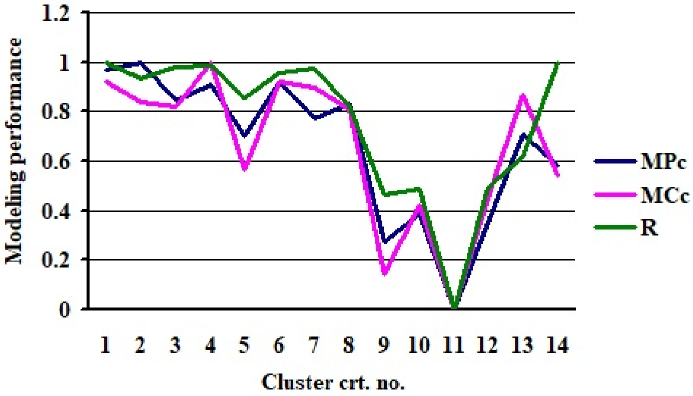

Fig. 6 Comparison between causal model potential and NN model performance

This entitles us to consider the result of the validation of the CM results to be positive. The differences appearing in the cases of some clusters might be explained by the difference between the dimensions of the application domains: in the case of the proposed CM the dataset has 11,175 beams, while in the case of finding the NN models, there were only 150 instances.

Note: The procedure for dimensionality reduction (see 2.4) can also be applied to E-variables C, TS and E. If the windows having H = 0 and $h = h_{ref} = 0.2097$ are successively imposed to two of E-variables, then the image of the third E-variable results in: $\Delta_{12} = 0.2094$, $\Delta_{13} = 0.2099$ and $\Delta_{12} = 0.7800$. This entitles us to conclude that the variables C and TS are highly dependent and they should have similar causal model trees, while E is relatively independent with respect to the other two and a distinct causal model tree needs to be drawn in this case.

3.2 Case study performed on the dataset of industrial instances

(1) Checking CM applicability in industrial conditions

The CM has also been performed in the case of a dataset extracted from the industrial environment and it concerns the job of turning roller bearing rings.

Data concatenation

After the clustering action, the values of ten C-variables were available, namely, exterior and interior diameters D_e and D_i of the ring, ring width L, cutting speed v, feed f, and depth t, volume of removed material V, maximum cutting force F and power P, and complexity index I_c (in connection with the ring profile). The time span TS was selected as E-variable.

The dataset has 155 instances; some of them are presented in Table 5.

Table 5 Real instance dataset (actual values, excerpt), [23]

Instance crt. no.	D_e (mm)	D_i (mm)	L (mm)	v (m/min)	f (mm/rot)	t (mm)	V (mm^3)	F (N)	P (kW)	I_c (-)	TS (min)
1	89.18	69.46	31.2	259.69	0.35	0.82	48.399	970.876	5.502	3.75	3.599
2	125.12	108.64	31.22	261.81	0.17	0.76	40.333	523.539	3.371	2.5	3.791
3	170.02	139.96	39.24	277.05	0.304	0.19	144.046	205.594	1.888	7.7	2.52
4	215.14	181.86	47.24	300.08	0.32	0.17	217.414	188.196	1.879	2.21	3.841
5	190	157.45	36.2	286.53	0.34	0.11	167.733	127.438	1.509	4.2	2.792
. . .	. . . .										

Comparison of instances

This time, the beam dataset generated with the same MatLab application has $N = C_{155}^{2} = 11{,}935$ beams.

Assessment of variables

Dimensionality reduction

The procedure has been applied in the same conditions as in the case of the dataset of artificial instances: $h_{ref} = h_7 = 0.2097$, $h_{k-1} = h_6 = 0.2621$. This time only four steps were necessary because after the fourth step, $\Delta_{min} = 0.5183 >> h_{k-1}$. Three C-variables ($D_e$, F and P) were found as dependent on other variables and discarded. The remaining seven C-variables lead to the maximal cluster: [D_i, L, v, f, t, V, I_c].

Assessment of variable modelling potential

The three criteria for assessing the modelling potential have been evaluated for each C-variable of the maximal cluster by calculating the values of a, b, and RMSE relative to TS as E-variable. The results are presented in Table 6.

Table 6 Values of a, b and RMSE, [23]

	D_i	L	v	f	t	V	I_c
a	0.1272	0.2241	0.05765	0.0041	0.02191	0.3234	0.01558
b	0.04458	0.03765	0.04864	0.05269	0.04961	0.03479	0.05466
RMSE	0.00116	0.00058	0.00176	0.00112	0.00185	0.00171	0.00269

Identification of casual models

The *MC* (assessed by means of the values of b) was adopted as a criterion for selecting the C-variables to be discarded when generating smaller clusters. Three clusters with six C-variables each have been generated from the maximal cluster in the first stage. Then, in the second stage, two clusters with five C-variables resulted from each of these three. Finally, two clusters of four C-variables were obtained from each cluster with five variables. After assessing the clusters potential in order to identify causal models, the process of generating smaller clusters had to be stopped at the level of 4-variable clusters. It should be noted that only three (instead of six) distinct clusters with five variables and four (instead of six) clusters with four variables were obtained. Hereby, the causal model tree will be formed from the 11 clusters presented in Table 7.

Table 7 Values of A, B and RMSE, [23]

Condition variable from cluster	A	B	RMSE
D_i, L, v, f, t, V, I_c	0.4058	0.0261	0.0034
D_i, L, v, t, V, I_c	0.4013	0.0266	0.0030
D_i, L, v, f, V, I_c	0.4263	0.0240	0.0030
D_i, L, v, f, t, V	0.4178	0.0287	0.0041
D_i, L, v, V, I_c	0.4167	0.0263	0.0028
D_i, L, v, t, V	0.4064	0.0314	0.0048
D_i, L, v, f, V	0.4253	0.0310	0.0044
D_i, L, V, I_c	0.3722	0.0344	0.0023
D_i, L, v, V	0.4339	0.0314	0.0047
D_i, L, t, V	0.3774	0.0365	0.0365
D_i, L, f, V	0.3919	0.0372	0.0039

(2) Assessment of CM performance in the addressed industrial case

The criteria for assessing the modelling potential of the remained clusters were evaluated by using the values of A, B and RMSE, presented in Table 7. The causal model tree drawn according to the MC_c criterion is depicted in Fig. 7.

The results of the case study performed on the dataset of industrial instances are very similar and prove the validity of the proposed CM once again. They reveal the existence of a model with six C-variables (D_i, L, v, f, V, I_c) that shows a very good potential for modelling the turning time span (B = 0.024), as well as a significantly simpler solution for doing the same thing with reasonably good results, namely the use of only four C-variables - D_i, L, v and V (when B = 0.0314).

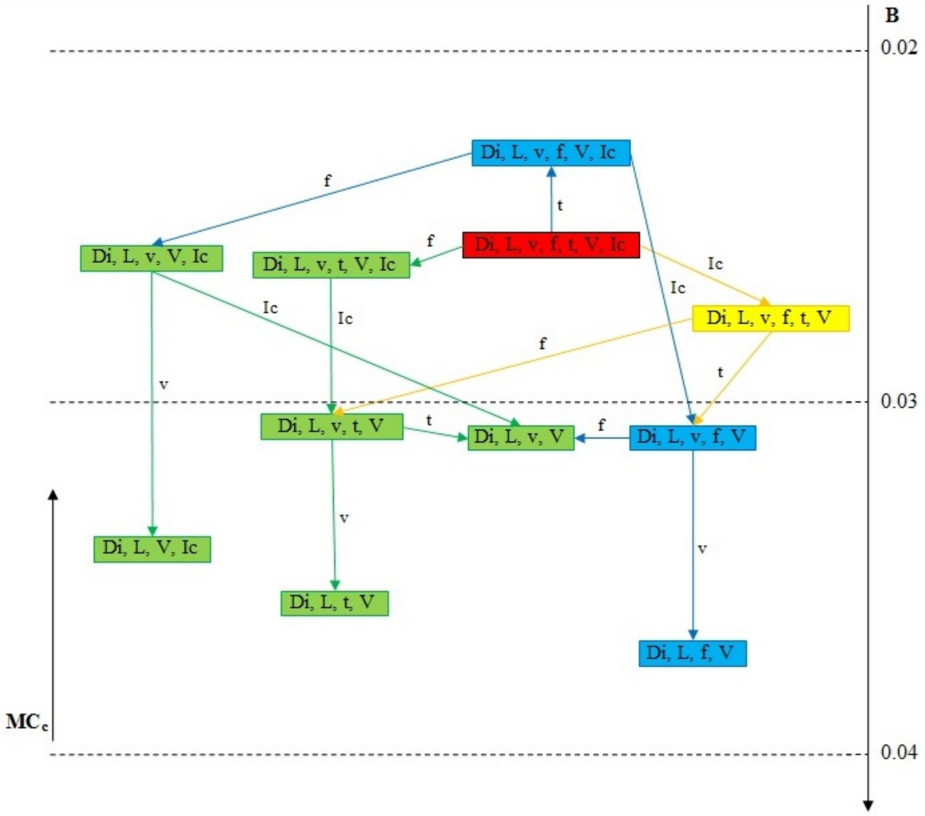

Fig. 7 Causal model tree in the case of industrial instance dataset, [23]

4. Conclusion

The MS management involves a ‘what if’ analysis, where the capability of evaluating the predicted values of a large number of variables is crucial. Appropriate models are necessary for evaluation. Finding the most suitable structure based on which a future model will be built is critical and it yet represents a notable challenge. The ultimate goal of the paper is to answer to this challenge. The answer is the proposed MS causal modelling.

The objective of CM is to enable a correlation of the model structure to the multiple conditionality imposed during the ‘what if’ analysis, and it concerns: i) the criteria based on which the analysis actually needs to be performed, ii) the synthetic/analytic character of the analysis, iii) the evaluation accuracy, iv) the unavailability of some independent variable values, v) the stringency in decision making process, and vi) the applicability when MS structure modifies in time quickly.

For reaching the objective, the MS is described by nominal jobs that it could involve. The here proposed MS causal modelling is realized by building causal models of MS nominal jobs. The CM supposes four steps: data concatenation, comparison of instances, analysis of variables, and identification of causal models. The delivered result is a causal model trees for all E-variables of the MS nominal jobs.

The degree of fulfilling the CM objective was assessed by the results obtained from the two case studies performed. The first case study was performed on a dataset of artificial instances and the other on a dataset of industrial instances. After the results have been analysed, the following conclusions may be drawn:

- The proposed MS causal modelling is applicable and works well. Unlike the existing approaches (see [8], [14], [19]), it delivers a significant number of model structure forms and facilitates the selection of the most suitable model.
- The proposed method for CM allows correlating the model accuracy to the actually required level by choosing an appropriate feature cluster from the delivered clusters.
- As opposed to current approaches (e.g. [12]) where the difficulties in evaluating MS features or possible unavailability of some features are ignored, the proposed CM enables us to choose C-variables from the available MS features.
- The CM can be applied with good results even when a small past casuistry is available (unlike in [7], [11]) for example, after reducing the number of instances from the database to one third, the CM leads to the same result.
- The comparison of the CM results to the results obtained by the NN modelling validates the CM feasibility.
- The MS causal modelling proves to have good potential for implementation in the industrial environment with a very wide range of possible applications.
- The performance obtained when applying the CM in industrial conditions is good.

Acknowledgement

This work was supported by the Romanian Ministry of Research and Innovation, CCCDI –UEFISCDI, project number PD-III-P1-1.2-PCCDI-2017-0446 / *Intelligent manufacturing technologies for advanced production of parts from automobiles and aeronautics industries* (TFI PMAIAA) – 82 PCCDI/2018 within PNCDI III.

REFERENCES

[1] C.-X. J. Feng, Z.-G. S. Yu, U. Kingi, and M. Pervaiz Baig, ‘Threefold vs. fivefold cross validation in one-hidden-layer and two-hidden-layer predictive neural network modeling of machining surface roughness data’, *J. Manuf. Syst.* **2005**, vol. 24, no. 2, pp. 93–107. https://doi.org/10.1016/S0278-6125(05)80010-X.

[2] N. Rehman, ‘Data Mining Techniques Methods Algorithms and Tools’ **2017**, vol. 6, no. 7, pp. 227–231.

[3] Su W., Bo M. (2006) Ant Colony Optimization for Manufacturing Resource Scheduling Problem. In: Wang K., Kovacs G.L., Wozny M., Fang M. (eds) Knowledge Enterprise: Intelligent Strategies in Product Design, Manufacturing, and Management. PROLAMAT 2006. IFIP International Federation for Information Processing, vol 207. Springer, Boston, MA. https://doi.org/10.1007/0-387-34403-9_120

[4] T. R. Paul, A. Saha, H. Majumder, V. Dey, and P. Dutta, ‘Multi-objective optimization of some correlated process parameters in EDM of Inconel 800 using a hybrid approach’, *J. Brazilian Soc. Mech. Sci. Eng.* **2019**, vol. 41, no. 7, pp. 1–11. https://doi.org/10.1007/s40430-019-1805-9.

[5] P. Denno, C. Dickerson and J. A. Harding, ‘Dynamic production system identification for smart manufacturing systems’, *J. Manuf. Syst.* **2018**, vol. 48, no. November 2017, pp. 192–203. https://doi.org/10.1016/j.jmsy.2018.04.006.

[6] S. Chun and T.-H. Kim, ‘Simultaneous identification of model structure and the associated parameters for linear systems based on particle swarm optimization’, *Complexity* **2018**, vol. 2018, pp. 1–17. https://doi.org/10.1155/2018/2713684.

[7] D. Kozjek, R. Vrabič, D. Kralj, and P. Butala, ‘A Data-Driven Holistic Approach to Fault Prognostics in a Cyclic Manufacturing Process’, *Procedia CIRP* **2017**, vol. 63, pp. 664–669. https://doi.org/10.1016/j.procir.2017.03.109.

[8] J. Zou, Q. Chang, J. Arinez, and G. Xiao, ‘Data-driven modeling and real-time distributed control for energy efficient manufacturing systems’, *Energy,***2017**, vol. 127, pp. 247–257. https://doi.org/10.1016/j.energy.2017.03.123.

[9] P. Oborski, ‘Integrated monitoring system of production processes’, *Manag. Prod. Eng. Rev.* **2016**, vol. 7, no. 4, pp. 86–96. https://doi.org/10.1515/mper-2016-0039.

[10] L. Chen, H. Qiu, L. Gao, C. Jiang, and Z. Yang, ‘A screening-based gradient-enhanced Kriging modeling method for high-dimensional problems’, *Appl. Math. Model.* **2019**, vol. 69, pp. 15–31. https://doi.org/10.1016/j.apm.2018.11.048.

[11] N. Lu, Y. Li, C. Liu, and W. Mou, ‘Cutting Tool Condition Recognition in NC Machining Process of Structural Parts Based on Machining Features’, *Procedia CIRP* **2016**, vol. 56, pp. 321–325. https://doi.org/10.1016/j.procir.2016.10.028.

[12] D. Kong, Y. Chen and N. Li ‘Force-based tool wear estimation for milling process using Gaussian mixture hidden Markov models’, *Int. J. Adv. Manuf. Technol.* **2017**, vol. 92, no. 5–8, pp. 2853–2865. https://doi.org/10.1007/s00170-017-0367-1.

[13] Y. Liu, X.H. Luan, and H. Liu, ‘Feature selection and sampling uncertainty analysis for variation sources identification in the assembly process online sensing’, *Int J. Adv. Manuf. Technol.* **2017**, vol. 92, pp. 2777-2785. https://doi.org/10.1007/s00170-017-0361-7

[14] J. Zhou, X. Li, O. P. Gan, S. Han, and W. K. Ng, ‘Genetic Algorithms for Feature Subset Selection in Equipment Fault Diagnosis’, *First World Congr. Eng. Asset Manag.* **2006**, pp. 1104–1113. https://doi.org/10.1007/978-1-84628-814-2_121

[15] C. C. Wang, Y.-C. Lin, Y.-C. Lin, S. R. Jhang, and C. W. Tung, ‘Identification of informative features for predicting proinflammatory potentials of engine exhausts’, *Biomed. Eng. Online* **2017**, vol. 16, no. s1, pp. 1–10. https://doi.org/10.1186/s12938-017-0355-6.

[16] R. Merrell and D. Diaz, ‘Comparison of Data Mining Methods on Different Applications: Clustering and Classification Methods’, *Inf. Sci. Lett.* **2015**, vol. 4, no. 2, pp. 61–66.

[17] M. Ester, H. P. Kriegel, J. Sander, and X. Xu, ‘A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise’, *Second Int. Conf. Knowl. Discov. Data Min.* **1996**, vol. 2, pp. 226–231.

[18] F. Klawonn, ‘Exploring Data Sets for Clusters and Validating Single Clusters’, *Procedia Comput. Sci.* **2016**, vol. 96, pp. 1381–1390. https://doi.org/10.1016/j.procs.2016.08.183.

[19] W. Guo and A. G. Banerjee, ‘Identification of key features using topological data analysis for accurate prediction of manufacturing system outputs’, *J. Manuf. Syst.* **2017**, vol. 43, pp. 225–234. https://doi.org/10.1016/j.jmsy.2017.02.015.

[20] G. Singh, F. Mémoli and G. Carlsson, ‘Topological Methods for the Analysis of High Dimensional Data Sets and 3D Object Recognition’, *Ned. Tijdschr. Geneeskd.* **2007**, vol. 151, no. 46, pp. 2551–2.

[21] C. Hitchcock, ‘Causal Models’, in *The Stanford Encyclopedia of Philosophy* (Fall 2018 ed.), Stanford University, **2018**.

[22] J. Pearl, ‘An introduction to causal inference’, *Int. J. Biostat.* **2010**, vol. 6, no. 2. https://doi.org/10.2202/1557-4679.1203.

[23] C. Afteni, ‘Holistic optimization of manufacturing process’, PhD Thesis, *‘Dunarea de Jos’ University of Galati*, **2020**, Series I 4: Industrial Engineering No. 70.

Submitted: 28.7.2020

Accepted: 11.01.2021

Prof. Gabriel-Radu Frumusanu*
Eng. Cezarina Afteni
Prof. Alexandru Epureanu
Department of Manufacturing Engineering
‘Dunarea de Jos’ University Faculty of
Engineering, Galati, Romania
*Corresponding author:
gabriel.frumusanu@ugal.ro