A R C H I V E S O F M E T A L L U R G Y A N D M A T E R I A L S
Volume 55 2010 Issue 3

M. PERZYK* A. SOROCZYNSKI*, J. KOZLOWSKI*

APPLICATIONS OF ROUGH SETS THEORY IN CONTROL OF FOUNDRY PROCESSES

ZASTOSOWANIA TEORII ZBIORÓW PRZYBLIŻONYCH W STEROWANIU PROCESAMI ODLEWNICZYMI

In contemporary manufacturing industry many varied methods are used for controlling process parameters, ranging from paper SPC charts to automated closed loop systems. In recent years a remarkably increasing interest in application of supporting systems based on computational intelligence methods, utilizing company recorded data, is observed. Most frequently, methods of automated knowledge extraction from data in the form of logic rules are utilized, based on classification learning systems. A growing interest of industrial applications of rough sets theory (RST) is observed in that area, which can provide with various types of important information concerning complex manufacturing processes. In the present paper the main elements and characteristics of RST are expounded and compared to another learning system – widely used decision trees (DTs). Current applications of the both methods in manufacturing industry are briefly reviewed. Results of the paper authors' research concerning two important issues of RST and DTs applications in foundry production are presented. They include assessment of correctness of relative significances of process parameters of arbitrary nature (e.g. physical, human, organizational etc.) and evaluation of reliability of engineering knowledge in the form of logic rules. The numerical experiments, carried out on simulated and real data, related to foundry technology, have shown that RST can be a valuable tool for control of complex manufacturing processes and it performs remarkably better then DTs.

Keywords: foundry production, process control, computational intelligence, rough sets theory, decision trees

We współczesnym przemyśle wytwórczym stosowanych jest wiele różnorodnych metod do sterowania i kontroli parametrów procesów, począwszy od papierowych kart kontrolnych SSP, a skończywszy na systemów automatycznych, pracujących w zamkniętych pętlach. W ostatnich latach obserwuje się znacznie zwiększone zainteresowanie stosowaniem systemów wspomagających opartych na metodach inteligencji obliczeniowej, wykorzystujących dane zarejestrowane w przedsiębiorstwie. Najczęściej wykorzystywane są metody zautomatyzowanego pozyskiwania wiedzy z danych w postaci reguł logicznych, oparte na klasyfikacyjnych systemach uczących się. W tym zakresie obserwuje się wzrastające zainteresowanie przemysłowymi zastosowaniami teorii zbiorów przybliżonych (ang. skrót RST), które są w stanie dostarczyć różnego typu informacji o złożonych procesach wytwórczych. W mniejszym artykule przedstawiono główne elementy RST i porównano z innymi systemami uczącymi się – szeroko stosowanymi drzewami decyzyjnymi (ang. skrót DTs). Dokonano krótkiego przeglądu aktualnych zastosowań obu systemów w przemyśle wytwórczym. Przedstawiono wyniki badań autorów artykułu dotyczące dwóch istotnych rodzajów zastosowań RST i DTs w produkcji odlewniczej. Obejmują one ocenę poprawności wyznaczania względnych istotności parametrów procesu o dowolnej naturze (np. fizycznej, ludzkiej, organizacyjnej) oraz ocenę wiarygodności wiedzy inżynierskiej w postaci reguł logicznych. Eksperymenty numeryczne, przeprowadzone na danych symulowanych i rzeczywistych, związanych z technologią odlewniczą pokazały, że RST może być wartościowym narzędziem sterowania złożonymi procesami wytwarzania, spełniając swoje zadania istotnie lepiej niż DTs.

1. Introduction

In contemporary manufacturing industry many varied methods are used for controlling process parameters, ranging from paper SPC charts to automated closed loop systems. In spite of the degree of automation of the control system, it is essential to formulate the relationships between process parameters as inputs and process results as outputs. In recent years a remarkably increasing interest in application of supporting systems based on artificial intelligence (AI) or computational intelligence (CI) methods is observed. They utilize data recorded in the company and can help engineers and operators to understand and diagnose manufacturing process problems, to provide procedures which identify the input parameters that can be effectively used to control the process and, finally, to develop the appropriate relationships. The CI tools are the key methods of data mining, a multidisciplinary field, which

* INSTITUTE OF MANUFACTURING TECHNOLOGIES, WARSAW UNIVERSITY OF TECHNOLOGY 02-524 WARSZAWA, NARBUTTA 85, POLAND

includes methodologies and tools from several disciplines, including also database systems, visualization and statistics. The CI algorithms can provide various types of information, much work has been done to develop methods of automated knowledge extraction in the form of logic rules of the type "if … then …".

The input and output process variables can be of several different types: numerical continuous – represented by real numbers (e.g. temperature in degrees Celsius) or discrete ones, which include ordinal type (expressed verbally or by integral numbers, e.g. temperature can be "low", "medium" or "high") as well as nominal type (values are represented by names, e.g. a grade of an alloy, a number of working group etc.). Variables of the ordinal type are of particular interest in control of various industrial processes as they can be used for expressing some uncertainties and approximations of the quantities involved in them. The approach based on fuzzy logic, using linguistic variables, is widely used, including foundry industry (e.g. [1]). Whereas application of the fuzzy calculus requires that the input – output relationships are assumed, based on human's knowledge or intuition, the CI methods are capable of semi-automated finding such dependencies, using data collected in the normal production [2].

The ordinal or nominal types of variables can be handled by several kinds of the learning systems, such as Bayesian classifiers, decision trees (DTs), some types of artificial neural networks, support vector machines and the systems based on rough sets theory (RST). For manufacturing problems, DTs are probably the most frequently used tools for knowledge extraction from data (e.g. [3-8]), whereas the RST-based methods seem to be their newer alternative (e.g. [9-12]). Both algorithms are relative simple, especially compared to neural or fuzzy-neural systems, and easy to interpret by the users. Both of them treat the data in a natural way however, they are based on completely different principles and algorithms. They are widely discussed in the world literature, and will be only briefly characterized below.

DTs are non-parametric classification models, constructed from data by successive splits of the data records (learning examples), starting from the whole set. The splits are made in such a way that in the resulting subsets the classes of the decision variable are possibly homogeneous (preferably identical). The best splitting point is based on one input variable, called a splitting variable. This procedure is repeated for successive subsets and it leads to a model structure represented by an oriented graph, resembling a tree. The splitting points are nodes of the graph, the first node is called a core and the lines connecting the nodes are called branches or edges. The subsets which are not further divided are called leaves, and they provide results of classification (the dominant class in a leave is decisive). A tree model usually requires restrictions on its size. It is done either in the course of constructing the tree (e.g. by stopping further splits when the assumed minimum number of examples in a node is achieved) or in a special simplification procedure of an already induced, too complex tree, called pruning. There are several algorithms for tree induction, which differ in the criterion of the class homogeneity in splits and in the criterion of the tree complexity. It is worth noticing, that each route leading from a core to a leaf can be expressed by a logic rule of the previously mentioned structure. DTs also allow for evaluation of relative significances of the attributes, based on the so called purity of the splits. The large increments of the class homogeneity resulting from a split based on a given variable indicate its large significance.

In RST approach each discernible learning example (data record) can basically make a rule. The set of rules thus obtained can be usually reduced and the rules can be simplified (i.e. their conditional part can be shortened), by deleting attributes which do not contribute to classification. The rules can be evaluated, first of all from the standpoint of uniqueness of the classification. This is expressed by the confidence, defined as a ratio of the number of examples in which occurs this same combination of attributes' values and this same class variable as in the rule, to the number of examples in which occurs that combination of attributes values only (i.e. regardless the output class). Another parameter used for rules evaluation is number (or fraction) of examples compatible with a rule, called rule's support. If it is not possible to obtain rules of 100% confidence, then some not fully unique rules are utilized (rough classification). RST also makes possible an easy evaluation of relative significances of the attributes, based on reduction of uniqueness of classification resulting from deleting a given attribute in all rules.

The practical aspects of application of the above CI tools are also different. The computation times necessary for DTs are generally short and the interpretation of rules obtained from DT can be facilitated by the graphical representation of the trees. The RST theory may require long computational times and may lead to much larger number of rules, compared to DTs, if one seeks a detailed information from the knowledge system. It should be noticed, that whereas DTs are widely spread both in handbooks and in commercially available software, the RST can be rather seldom found, except for scientific literature.

Making a right choice of the knowledge extraction algorithm is important, particularly in construction of control systems. However, there are very little comparative studies available, which could show the advantages and weakness of individual tools [9, 10]. The purpose of the present paper is to show important advantages of RST-based algorithms, compared to those utilizing DTs and some statistical methods, from the standpoint of control of complex manufacturing processes, such as foundry and metallurgical ones. Two important features are considered and evaluated: correctness of relative significances of process parameters and reliability of engineering knowledge in the form of logic rules. Determination of the most significant process parameters can help not only to detect root causes of deteriorating product quality but also to indicate optimal or critical parameters that can be used for controlling the process. The logic rule set can be used directly in a control system for the discrete-type output variables.

2. Methodology

The methods discussed in the previous chapter were assessed with a use of simulated and industrial data sets. The synthetic sets were obtained by assuming analytical formulas of the type Y=f(X1, X2, ...), from which, for random values of continuous-type input variables X1, X2, ..., the continuous-type dependent variable Y was calculated. Then a Gaussian-type noise with maximum deviation ±20% was imposed on the input variables, and finally all the continuous values were converted to categorical ones, assuming the equal intervals method. Two numbers of the intervals for discretization (i.e. numbers of categories) were assumed: 5 and 10. In most cases, sets comprising 1000 records were generated in this way. Three basic formulas were used, giving simulated data sets of the characteristics described below.

Sim 1, obtained from the basic formula:
$Y=X1+2\cdot X2+3\cdot X3+4\cdot X4+5\cdot X5$; linearly increasing significances of variables, in additive manner, without interactions.

Sim 2, obtained from the basic formula:
$Y=X1\cdot X2+X3+X4+X5$; strong interactions between two variables of equal significances, the remaining variables have significances equal to the joint significance of the first two, without interactions among them.

Sim 3, obtained from the basic formula:
$Y=\tanh(0{,}1\cdot X1+0{,}2\cdot X2+0{,}4\cdot X3+0{,}8\cdot X4+1{,}6\cdot X5)$; increasing significances of variables, additive model with asymptotic output limit (saturation value) resulting in a specific form of interaction between all input variables.

Similar situations as represented in the above relationships often appear in practice. For example, Sim 3 may reflect simultaneous action of several chemical elements, which change the alloy microstructure and properties in the same direction. These cannot exceed certain physical limits and the actual effect of each variable depends on the structure and properties produced by the other elements.

All the industrial data sets were related to metal casting processes. Ind 1, Ind 2 and Ind 3 data sets were collected in a regular production of ductile cast iron in a cooperating foundry; the numbers of records were 861.

Ind 1 correlates chemical composition of ductile cast iron, defined by 5 elements, often considered as most important for its microstructure and mechanical properties (Mn, Si, Cr, Ni and Cu), with the material tensile strength;

Ind 2 correlates chemical composition of ductile cast iron, defined by all 9 elements controlled in the foundry (C, Mn, Si, P, S, Cr, Ni, Cu and Mg), with its tensile strength;

Ind 3 correlates chemical composition of ductile cast iron, defined by 5 elements as in Ind 1 with its four grades, assumed as the output class variable: 400/18, 500/07, special 500/07 with increased hardness and 'not classified'.

The rest two data sets: Ind 4 and Ind 5 were obtained as readouts from a semi-empirical nomograph [13] which permits to calculate solidification shrinkage of grey cast iron as a function of four variables: carbon contents (5 different values – categories), sum of silicon and phosphorus content (4 values), casting modulus (4 values) and pouring temperature (4 values). The outputs were the decisions concerning necessity and size of application of feeders to avoid the shrinkage defects. In Ind 4 data set the output, named 'Feeder', had 2 classes ('No' – when the volume change between poring and the end of solidification was zero or positive and 'Yes' – when the overall volume change was negative). In Ind 5 data set the output had 3 classes ('No', 'Small' and 'Large', dependent on the magnitude of the shrinkage). Numbers of records in these data sets were 190. It is worth noticing that, unlike the previously described simulated and industrial data sets, Ind 4 and Ind 5 have an important feature: a very low level of noise, which could be only a result of inaccuracies in the readouts. The discretization of the continuous – by nature – input variables was not required, as the readouts of the nomograph were made for selected, fixed values of these variables. The noise existing in typical production data as well as the simulated data generated as described above, may result in their inconsistency, defined as an occurrence of different output variable values (decision classes) for an identical combination of input values. In Ind 4 and Ind 5 data sets such inconsistencies were absent.

The RST-based procedure, oriented at generation of full set of logic rules, was written by the present authors with a somewhat similar approach as used in the Explore algorithm [14]. First, all the combinations of single input variables appearing in the data are placed in the rules (i.e. rules including only one condition are generated) and their confidences are calculated. Then the further conditions are added, providing the confidence of a rule thus obtained is increased, compared to the rule with shorter conditional part. The relative significances of input variables were calculated on the basis of the reduction of so called positive region of data (i.e. giving rules of 100% confidence) resulting from ignoring a given variable. More details concerning RST computational details can be found in [15].

The DTs were created and evaluated with a use of the well known C&RT algorithm, included in the commercial software Statistica (details of the computational procedures can be found in the software manual [16]. In order to obtain possibly the largest choice of logic rules from the data, various splitting conditions, stopping criteria and pruning parameters were tried out. The smallest trees which ensured the smallest fraction of false predictions for training sets were chosen. For the purpose of computing the relative significances of input variables, two different stopping criteria were used in most cases: the Statistica default, giving relatively 'small' trees, and the user criterion of minimum records in leaves equal to 5, leading to relatively 'large' trees.

The relative significances of input variables were also calculated using the statistical method appropriate for discrete-type variables, based on contingency tables. The Cramer's V statistics was used as a measure of the significance.

3. Assessment of the relative significances of process input variables

For the simulated data sets the expected values of relative significances obtained by various methods are known as equal to the coefficients appearing in the basic formulas used for generation of the data. Hence, the evaluation of the precision of the methods can be easily made. In Fig. 1 selected comparisons of the calculated relative significances with the expected ones are shown and in Fig. 2 the average errors, defined as absolute differences between calculated and expected values, are presented for several cases.

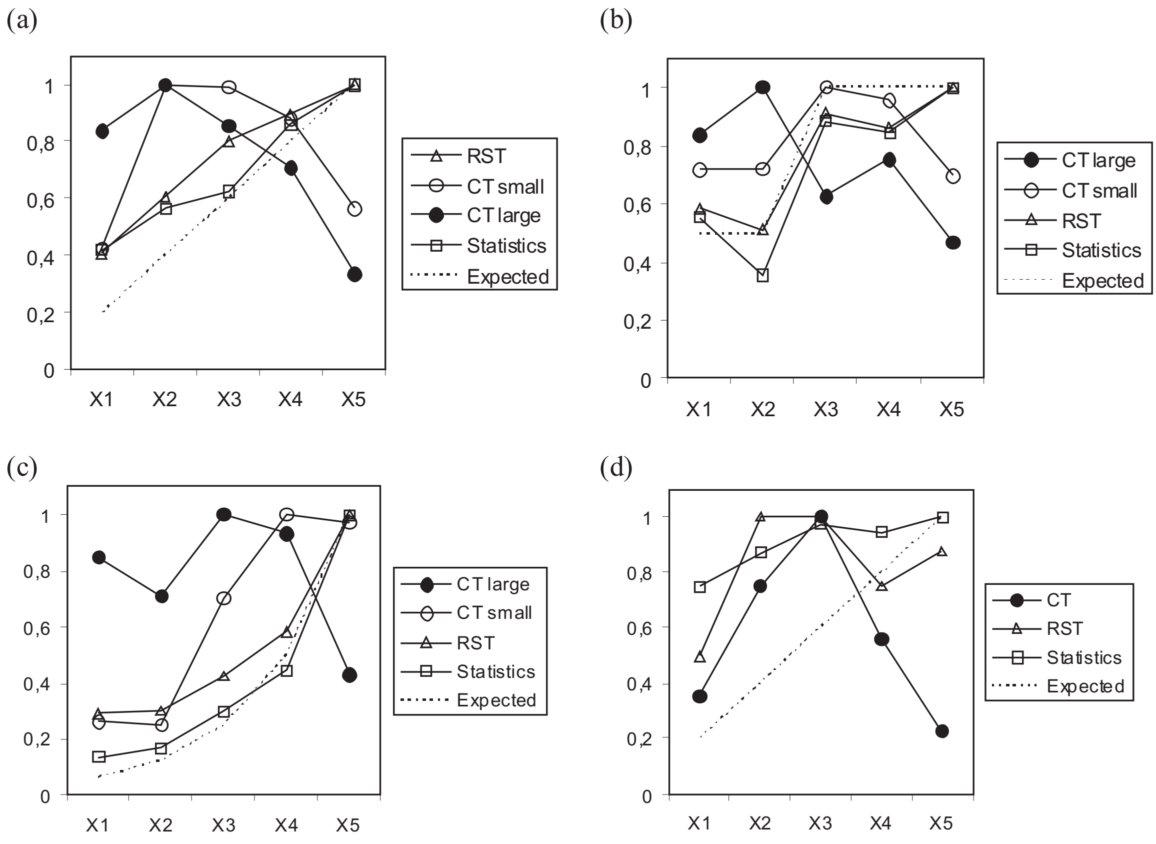


Fig. 1. Relative significances of input variables, obtained by various methods and expected, for simulated data sets with the assumed number of categories equal 5: (a) Sim 1, 1000 records; (b) Sim 2, 1000 records; (c) Sim 3, 1000 records; (d) Sim 1, 100 records

It can be seen that for all the simulated data sets with 1000 records the DTs predictions are very poor, compared to RST and the statistic method: not only the errors are much higher but it is also important that DTs often do not reflect the expected tendencies of the variables' significances. However, the good performance of RST and statistical method is not confirmed for small data sets (Figs. 1d and 2c).

For the industrial data the expected values of relative significances are generally not known. For the data sets related to ductile cast iron production (Ind 1, Ind 2 and Ind 3) the main problem is that they were collected in particular foundry, where some of the chemical elements could be kept at the levels which do not allow them to exhibit their full effect on the mechanical properties of the alloy. The only reliable information obtained from that foundry was that copper was the main element used for control of the matrix of SG iron (increasing the pearlite content and, consequently, the tensile strength) and it can be expected to have the largest significance. This was confirmed by another type of the significance analysis, based on a more precise regression modeling for that data which allowed to avoid conversion of the real numbers to categories [17]. In Fig. 3 the present work results assuming 5 input variables are presented and in Fig. 4 the results for 9 elements are shown, together with the above mentioned results obtained from regression modeling.

It can be seen that for the case of 5 elements assumed as inputs and tensile strength as output (Fig. 3a) the three methods brought generally divergent results and only the statistical approach pointed at copper contents as the most significant variable. When the SG iron grade was assumed as the output (Fig. 3b) the results obtained by the three methods are fairly similar and indicate copper as a significant element. For the case of 9 elements assumed as inputs and tensile strength as output (Fig. 4) the three methods studied in the present work give differentiated predictions for most input variables (except sulfur as the least significant element and manganese as a very important one). None of the present methods pointed at copper as the most significant element, as indicated by the regression analysis. It is important to notice, that the latter have also shown divergent results for some variables (e.g. Mn and Si).

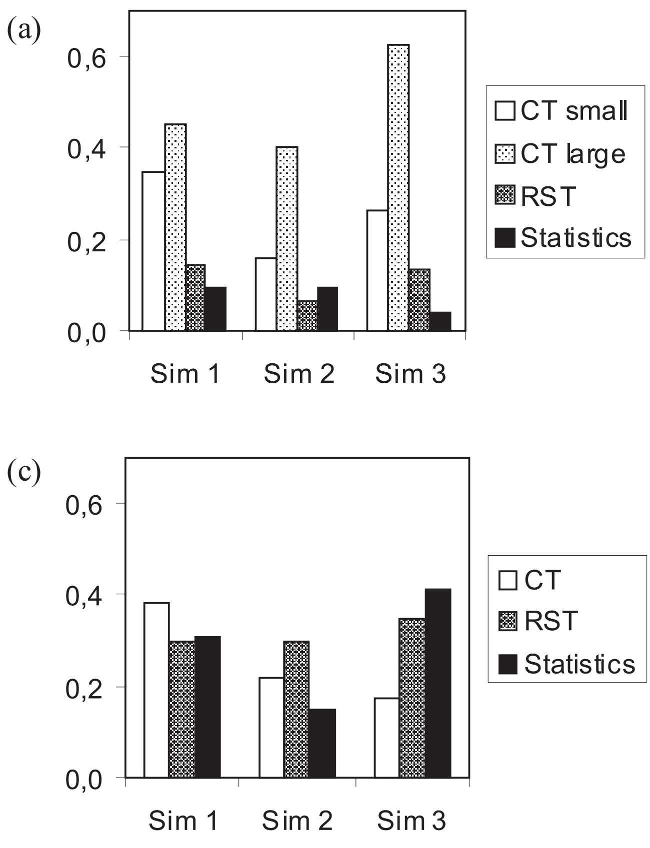
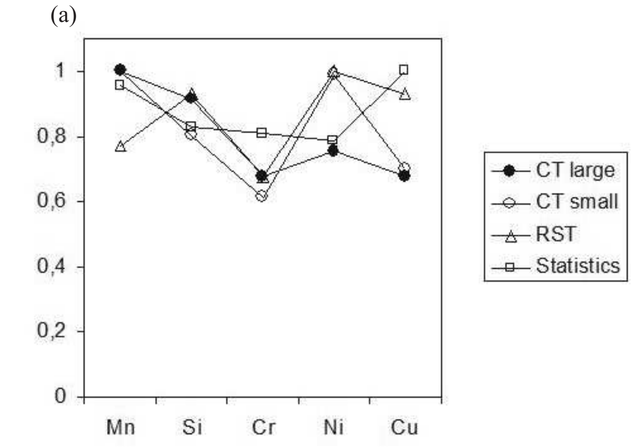


Fig. 2. Average errors of relative significances obtained by various methods and different numbers of records and categories: (a) 1000 records and 5 categories, (b) 1000 records and 10 categories, (c) 100 records and 5 categories



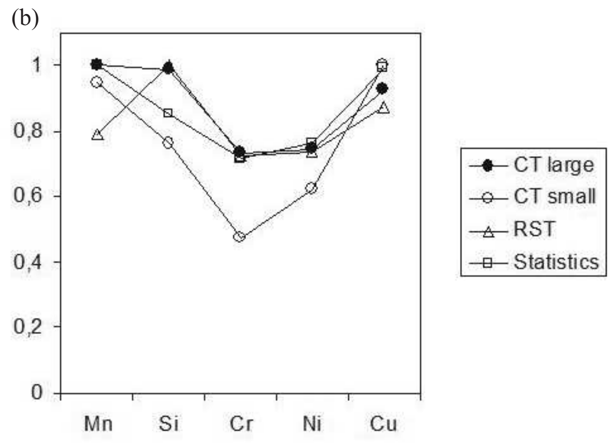


Fig. 3. Relative significances of input variables, obtained by various methods for two industrial data sets related to SG iron production: (a) Ind 1 – tensile strength assumed as the output (5 categories), (b) Ind 2 – alloy grade assumed as the output (4 classes)

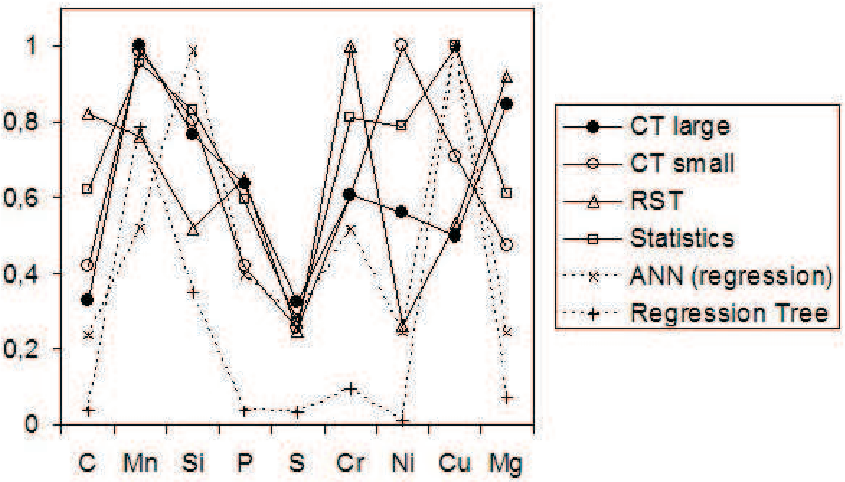
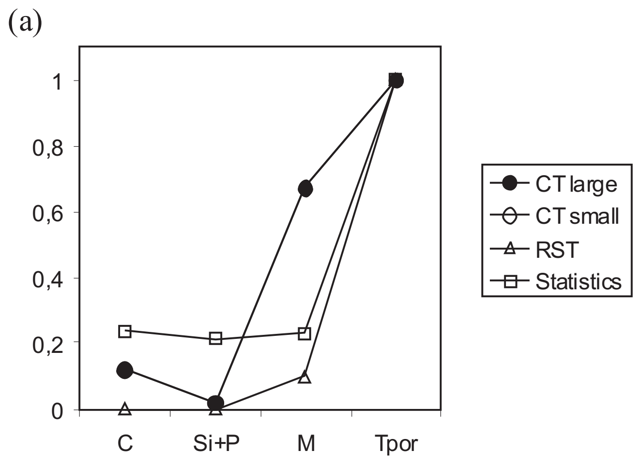


Fig. 4. Relative significances of input variables obtained in the present work by various methods for Ind 3 data set – SG iron tensile strength assumed as the output (5 categories, solid lines) and obtained in [16] by regression modeling for continuous variables (dotted lines)



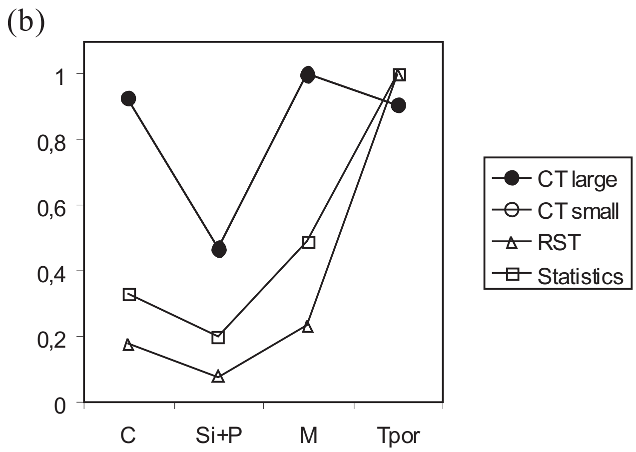


Fig. 5. Relative significances of input variables, obtained by various methods for two industrial data sets related to feeding of grey cast iron castings: (a) Ind 4 – requirement of feeder application assumed as the output (2 categories), (b) Ind 5 – requirement and size of feeder assumed as the output (3 classes)

In Fig. 5 the results for two data sets related to feeding of grey cast iron castings (Ind 4 and Ind 5) are shown. According to industrial experience, it can be expected that the alloy chemical composition, commonly expressed by its carbon content and the sum of silicon and phosphorus contents, has minor effect on shrinkage and, consequently, feeding requirements. The main influencing factors should be the poring temperature which directly decides on the magnitude of volume change from poring to the beginning of solidification and the casting modulus, expressing the casting cooling rate influencing kinetics of the volume changes during solidification. The results obtained by all the three methods fully confirmed these expectations for the case of two output classes, discerning sign of the volumetric changes from poring to the end of solidification (Fig. 5a). However, for the more complex output (Fig. 5b), the DTs predictions generally failed.

4. Assessment of engineering knowledge systems obtained from RST and DTs

4.1. Requirements for knowledge rules applicable for control of manufacturing processes

General requirements for knowledge rules which could be useful in manufacturing industry are rather obvious and similar to those for other areas of applications. First, the rules should be reliable, which means that there is a real chance that an application of the rule will bring the predicted result. This can be expressed by the previously explained rule quality parameters: confidence and support. Second, the rules should not be unnecessarily demanding, i.e. they should not comprise conditions which are not important, particularly redundant. The algorithms used for knowledge extraction are first of all oriented at generation of a set of rules which best characterize the problem, i.e. most reliable ones. However, in many industrial processes, particularly in manufacturing, some more specific requirements are relevant, related to design and development of new processes or control of currently running ones.

Hence, typical questions to be answered by using the rules can be formulated as follows:

- What are the most effective and reliable ways (i.e. process parameters - input values) to achieve an assumed result (class variable)?
- What would happen if we are not able to apply certain input values, i.e. what would we get if we use different ones? Do we still have a chance (and how big) to get the required result?
- What will be the predictions (and how reliable) in the case some input variables cannot be specified, e.g. they may be out of control?
- What are all alternative ways to achieve our goal and how reliable are they?

The requirements for rules system and the knowledge extraction tools, suitable for manufacturing industry applications, are not only a consequence of the issues described above, but also the specificity of available data. Typically, the number of independent variables (i.e. problem dimensionality) is not large, it seldom exceeds 10. Number of available records can vary within broad ranges, from only a few to many thousands, especially when the automatic data acquisition system is utilized. Typical industrial data are noisy, which results in their inconsistency, i.e. an occurrence of different output variable values for an identical combination of input values (conditions in a rule).

From the characteristics of industrial processes problems presented above the following requirements for rule systems seem to be essential or at least important:

- The rules should make use of all information in data. This means, for example, that all output values (classes) must be represented. Even single cases can be valuable and therefore they should be reflected in the rule system.
- The rules should not contain redundant conditions as they can be misleading for the user.
- It should be possible to find a rule 'tailored' to the user specifications, including combinations of input variables values which are not represented in the data.
- Reliability of all rules should be evaluated, using the confidence and support as the primary parameters.

4.2. Characteristic behavior of DTs and RST in rules extraction

A structure of a DT model is uniquely defined by a set of the logic expressions, corresponding to the knowledge rules. The nature of DT models, based on recursive partitioning of the data records, results in a set of conditions, which may be different from the combinations of input variables in the training data records. Some of the combinations appearing in the data set may be absent in the tree and vice versa, also some sequences of conditions occurring in the data may be abbreviated in the tree. The lack of some combinations of input values in DTs which are present in training data, may result in the rule system in which some important rules are missing.

Another consequence is that DTs can give wrong predictions for training data. In case of consistent data, this may be a result of improper tree structure, i.e. in which the given combination of input values (attributes) is connected with a class of the output variable which is different from that which appears in the data. Partly incorrect predictions may be a consequence of the fact, that DTs are able to give only one prediction for a given combination of input variables values. For noisy, inconsistent data it must always lead to a fraction of false predictions. Considering a DT as a knowledge rules system it means that for that type of data DTs must omit some rules, potentially also important for a user. In particular, those omitted rules can be the only ones which give a certain output.

Rules obtained from DTs may include redundant conditions as the splitting variable used in the core must appear in all rules (generally, the splitting variable in a node must appear in all rules resulting from subsequent splits). In contrast, RST provides 'fitted' rules, i.e. without unnecessary conditions. That type of behavior of the both algorithms was commented in detail in [9].

It is essential that all of the above discussed drawbacks of the rule systems obtained from DTs are absent in the RST-based systems. Below, some results of numerical tests are presented, which demonstrate how much this fundamental difference can be significant.

In Fig. 6 the fractions of wrong predictions obtained from DTs for all consistent data subsets (i.e. all the discernible input values combinations pointing at one output value only) are shown, for selected training data sets.

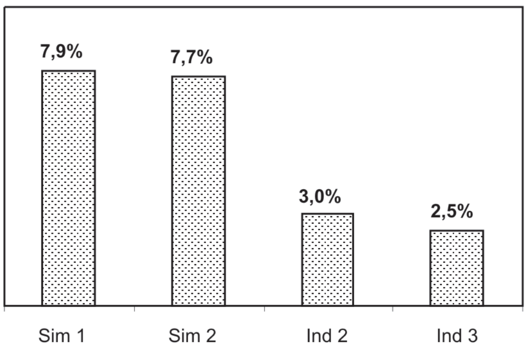


Fig. 6. Average fractions of false predictions obtained from DTs for consistent data subsets (including single records)

The general level of false predictions for real data is much lower, compared to simulated data. An interpretation of this observation would require a deeper analysis of the data sets structures, e.g. representativeness of the classes of input and output variables.

In Fig. 7 some statistical information obtained for inconsistent data subsets is shown. It is interesting to note that in several cases DTs have pointed at the decision classes which are not predominant for the given combination of input values.

In Fig. 8 the fractions of rules included in DTs which are not supported by data are shown, exhibiting quite large values in several cases. In principle, this can be a positive feature of DTs as such rules may be desired by a user. However, the usefulness of such rules may be questionable.

First, because they do not necessarily meet the user's specific needs and second, because their reliability, defined by confidence and support, is not determined.

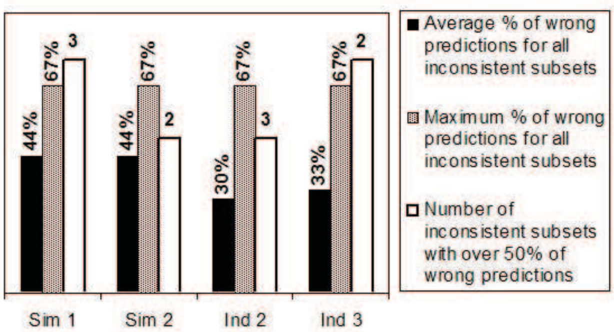


Fig. 7. Statistics of false predictions obtained from DTs for inconsistent data subsets

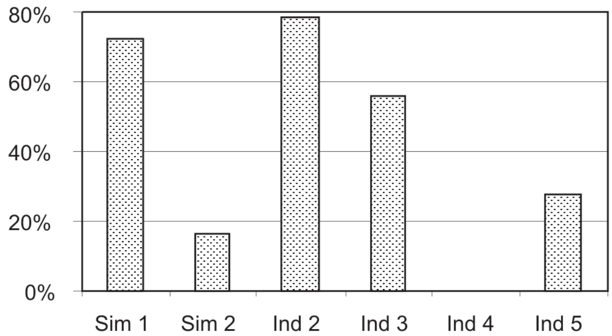


Fig. 8. Fractions of rules in DTs not supported by data

In Fig. 9 the numbers of rules absent in DTs, but extracted by RST, are presented, together with total numbers of rules in DTs and from RST. The missing rules may be valuable for a user, as it was found that their confidences are relatively high and comparable with those obtained for the rules which are included in DTs. It is worth noticing that for some of the simulated data sets, some of the missing rules had 100% confidence.

In Fig. 10 fractions of DT rules with redundant conditions are shown. Obviously, the RST rules, taken as reference, had this same confidence values. It was also found that the average number of redundant conditions was similar to the number of important conditions. The conclusion is that the presence of redundant conditions in rules obtained from DTs, being a result of the nature of that type models, may be their significant disadvantage.

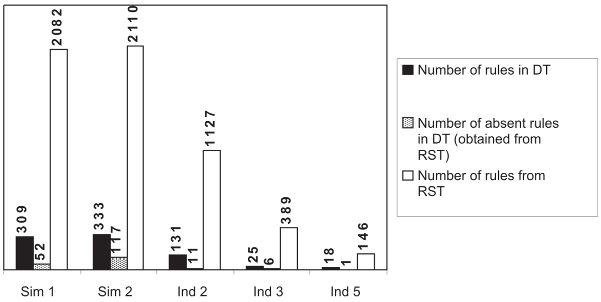


Fig. 9. Quantities of rules in DTs and obtained from RST – total and missing in DTs

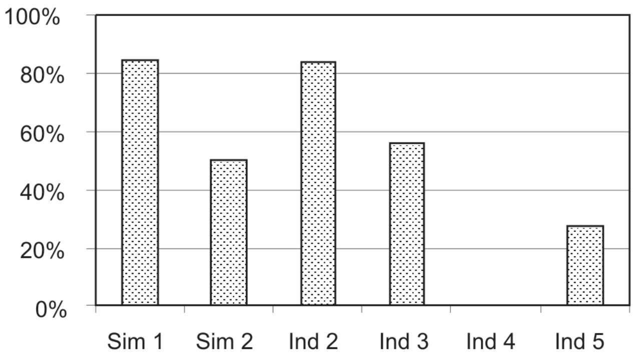


Fig. 10. Fractions of rules in DTs with redundant conditions

An important feature of a rule system is its predictive capability for new data, i.e. combinations of the input variables' values which have not appeared in the past. Some preliminary simple tests confirmed that for some cases DTs are unable to give predictions for the desired new input values combinations, as discussed earlier. Also, relative large fractions of false predictions by RST-based rule systems were found; this requires to treat this problem in a more detail in a separate study.

5. Summary and conclusions

The comparative analysis of two types of classification learning systems, have revealed substantial advantages of RST over DTs in two important aspects of their applications in control of complex manufacturing processes, characteristic for foundry and metallurgical production.

For the simulated data, identification of significances of process parameters by the RST-based systems generally appeared to be much more precise and reliable, compared to DTs. The widely used statistical method, based on contingency tables, also demonstrated a good performance and turned out to be the best in most cases. This substantial advantage of RST-based and statistical methods was partly confirmed by the real data, related to foundry production. However, this general observation does not concern small data sets, for which the errors of those two methods errors increased 2 to 3 times, compared to the corresponding large sets. These errors were comparable to those obtained from DTs and may be regarded as non-acceptable for many applications. The methods of finding the relative significances of discrete type variables for small data sets require further analyses.

RST-based rule systems seem to be fundamentally better in every respect, compared to those obtained from DTs, including completeness, reliability and lack of redundant conditions of the rules. A further research, involving combinations of RST and fuzzy sets approaches in control of production processes, would be desirable.

Acknowledgements

This work was supported by grant N R07 0015 04 from Ministry of Science and Higher Education, Poland.

REFERENCES

[1] E. Ziółkowski, R. Wrona, Using Fuzzy Optimisation Method in Calculation of Charge Burden to Correct the Chemical Composition of Metal Melt, Archives of Foundry Engineering **7**, 183-186 (2007).

[2] E. Czogała, A. Mrózek, Z. Pawlak, The idea of a rough fuzzy controller and its application to the stabilization of a pendulum-car system, Fuzzy Sets and Systems **72**, 61-73 (1995).

[3] K. Wang, Applying data mining to manufacturing: The nature and implications, J Intell Manuf **18**, 487-495 (2007).

[4] R. S. Chen, R. C. Wu, C. C. Chang, Using data mining technology to design an intelligent CIM system for IC manufacturing, Proc. Sixth Int. Conf. Softw. Eng. Atif. Intell. Netw. Parallel/Distr. Comput. Self-Assemb. Wireless Netw., SNPD/SAWN Towson, MD, USA, 70-75 (2005).

[5] J. Hur, H. Lee, F. G. Baek, An intelligent manufacturing process diagnosis system using hybrid data mining, Lect. Notes Comput. Sci. **4065 LNAI**, 561-575 (2006).

[6] H. Huang, D. Wu, Product quality improvement analysis using data mining: A case study in ultra-precision manufacturing industry, Lect. Notes Comput. Sci. **3614 LNAI**, 566-580 (2006).

[7] L. Rokach, O. Maimon, Data mining for improving the quality of manufacturing: a feature set decomposition approach, J Intell Manuf **17**, 285-299 (2006).

[8] D. Koonce, C. H. Fang, S. C. Tsai, Data mining tool for learning from manufacturing systems, Comput Ind Eng **33**, 27-30 (1997).

[9] A. Kusiak, C. Kurasek, Data mining of printed-circuit board defects, IEEE Trans Rob Autom **17**, 191-196 (2001).
[10] T. L. Tseng, M. C. Jothishankar, T. Wu, G. Xing, F. Jiang, Applying data mining approaches for defect diagnosis in manufacturing industry, IIE Annual Conference and Exhibition, Houston, 1441-1447 (2004).
[11] L. Shen, F. E. H. Tay, L. Qu, Y. Shen, Fault diagnosis using Rough Sets Theory", Computers in Industry, **43**, 61-72 (2000).
[12] H. Sadoyan, A. Zakarian, P. Mohanty, Data mining algorithm for manufacturing process control, Int J Adv Manuf Technol **28**, 342-350 (2006).
[13] M. Perzyk, A. Soroczynski, Comparison of selected tools for generation of knowledge for foundry production, Archives of Foundry Engineering **8**, 263-266 (2008).
[14] J. Stefanowski and D. Vanderpooten, Induction of decision rules in classification and discovery-oriented perspectives, Int J Intell Syst **16**, 13-27 (2001).
[15] L. Polkowski, Rough Sets: mathematical Foundations. Physica-Verlag Heidelberg New York (2002).
[16] StatSoft, Inc. (2008). STATISTICA (data analysis software system), version 8.0. www.statsoft.com.
[17] M. Perzyk, R. Biernacki, J. Kozlowski, Data mining in manufacturing: Significance analysis of process parameters, Proc. IMechE Part B: J. Engineering Manufacture **222**, 1503-1516 (2008).

Received: 10 February 2010.