

## Detecting Patterns in Energy Use and Greenhouse Gas Emissions of Cities Using Machine Learning

Kathleen B. Aviso<sup>a</sup>, Marc Joseph Capili<sup>a</sup>, Hon Huin Chin<sup>b</sup>, Yee Van Fan<sup>b</sup>, Jiří Jaromír Klemeš<sup>b</sup>, Raymond R. Tan<sup>a,\*</sup>

<sup>a</sup> Chemical Engineering Department, De La Salle University, 2401 Taft Avenue, 0922 Manila, Philippines

<sup>b</sup> Sustainable Process Integration Laboratory - SPIL, NETME Centre, Faculty of Mechanical Engineering, Brno University of Technology- VUT Brno, Technická 2896/2, 616 69 Brno, Czech Republic  
[raymond.tan@dlsu.edu.ph](mailto:raymond.tan@dlsu.edu.ph)

Cities are expected to play a major role in managing climate change in the coming decades. The actual environmental performance of urban centres is difficult to predict due to the complex interplay of technologies and infrastructure with social, economic, and political factors. Machine learning (ML) techniques can be used to detect patterns in high-level city data to determine factors that influence favourable climate performance. In this work, rough set-based ML (RSML) is used to identify such patterns in the Sustainable Cities Index (SCI), which ranks 100 of the world's major urban centres based on three broad criteria that cover social, environmental, and economic dimensions. These main criteria are further broken down into 18 detailed criteria that are used to calculate the aggregate SCI scores of the listed cities. Two of the environmental criteria measure energy intensity and greenhouse gas (GHG) emissions. RSML is used to generate interpretable rule-based (if/then) models that predict energy utilisation and GHG emissions performance of cities based on the other criteria in the database. Attribute reduction techniques are used to identify a set of 7 non-redundant criteria for energy use and 9 non-redundant criteria for GHG emissions; 6 criteria are common to these two sets. Then, RSML is used to generate rule-based models. A 10-rule model is determined for energy intensity, while an 11-rule model is found for GHG emissions. Both models were reduced further by eliminating rules with weak generalisation capability. A key insight from the rule-based models is that social, environmental, and economic attributes are associated with energy intensity and GHG emissions due to indirect effects.

### 1. Introduction

Cities play a major role in managing critical environmental issues such as energy resource consumption (Kennedy et al., 2015) and greenhouse gas (GHG) emissions (Ramaswami et al., 2021). The transition of modern urban centres into future sustainable smart cities is going to be critical to ensuring that these environmental issues are managed properly (Kamyab et al., 2020). Because the state of any given city is a complex function of many variables, data-driven insights are needed to formulate effective policies that drive the transition to improved sustainability (Gue et al., 2021).

Machine learning (ML) tools have become ubiquitous in the modern world (Jordan and Mitchell, 2015). They provide a powerful means of detecting patterns in data and then converting such patterns into models that can be used for analysis, classification, and prediction. Examples of recent applications include the analysis of alarm systems in process plants (Tamascelli et al., 2020) and the prediction of higher education outcomes (Aviso et al., 2020). ML tools have also been used to find patterns in the environmental data of cities and countries. For example, Mostafa (2010) used self-organising maps (SOM) to cluster countries based on ecological footprint; Ma et al. (2012) did a similar study using support vector machines (SVM). Sugiawan and Managi (2019) used ML to analyse patterns in inclusive wealth data of nations. Lopez et al. (2021) used rough set-based machine learning (RSML) to determine how country-level socioeconomic attributes predict emissions levels. Gue et al. (2021) used fuzzy cognitive maps (FCM) to determine Circular Economy (CE) drivers at the city level.

Conventional ML tools such as SOM and SVM have been criticised for lack of transparency and interpretability. As a result, there has been increased interest in explainable artificial intelligence (XAI) approaches that plausible

balance interpretability with sheer statistical performance (Calegari et al., 2020). These features are important in applications where clear insights need to be drawn from ML model outputs. One such approach is RSML (Mahajan et al., 2012), which generates models consisting of if/then rules for prediction or classification purposes. RSML is based on the theory of rough sets proposed by Pawlak (1982) and is suited for dealing with data with non-deterministic data with unclear patterns. A comprehensive tutorial on RSML can be found in Komorowski (2014). In addition to predicting country-level emissions (Lopez et al., 2021), recent sustainability-oriented applications include the prediction of reservoir integrity for CO<sub>2</sub> sequestration (Aviso et al., 2019) and the molecular design of environment-friendly specialty chemicals (Radhakrishnapany et al., 2020). The features of RSML make it well-suited to the problem of classifying sustainable cities, but no work has reported any such application to date. The purpose of this paper is to address this research gap.

In this work, RSML is used to generate ternary rule-based models to classify cities into good, moderate, and poor performers with respect to energy intensity and GHG aspects based on different social, environmental, and economic criteria from the Sustainable Cities Index (SCI) published by Arcadis (2020). The rule-based classifiers capture patterns in the data of the cities listed in SCI, using ternary logic due to its favourable statistical and decision-support properties (Yao, 2011). The rest of this paper is organised as follows. Section 2 gives the formal problem statement. Section 3 discusses an overview of RSML. Section 4 describes the key results of the application of RSML to the SCI data, while Section 5 discusses the practical implications of the results. Section 6 gives conclusions and prospects for future work.

## 2. Problem Statement

The formal problem statement is as follows:

- Given a set of SCI attributes which can be categorized either as condition, C, or decision, D, attributes of cities;
- Given a set of sample cities, S, where sample j has known performance C<sub>ij</sub> in the condition attribute i and known classification D<sub>kj</sub> in the decision attribute k;
- The problem is to determine a set of predictive rules which can adequately classify cities into decision class k in terms of energy intensity and GHG emissions.

## 3. Rough Set Based Machine Learning

It is normally assumed that objects can be classified based on information about them. However, this assumption is unable to deal with vague concepts at the boundaries of so-called crisp (well-defined) sets. Pawlak (1982) introduced the rough set theory (RST), which defines vague concepts to be confined between a lower, B<sub>\*</sub>(X), and upper approximation, B<sup>\*</sup>(X) of known crisp concepts. The lower approximation contains all objects that clearly belong to the set, while the upper approximation contains elements that can possibly belong to a set. The difference between the upper and lower approximation defines the boundary of the rough (vague) concept. To facilitate classification using RSML, the information about an object needs to be organised in an information table where the rows consist of examples, the columns consist of attributes (e.g., condition and decision attributes), and the elements describe the examples in the relevant attribute (Pawlak, 1984). Given a set of attributes A, objects may be clustered together because they are indiscernible from each other based on their performance in attribute set B where B ⊂ A. The objective then is to determine subset B, which will adequately cluster objects within the same decision class. The accuracy of the approximation, α<sub>B</sub>(X) of any subset, B is defined in Eq(1). Elements can also be defined according to their rough membership, μ<sub>X</sub><sup>B</sup>(x), as defined in Eq(2). This metric provides information with regards to the degree that example x belongs in class X based on the information in the attributes of subset B. Redundant information is also eliminated in RST by selecting subset B to contain only indispensable attributes; this is referred to as the reduct B, Red(B). The intersection of all reducts is known as the core, Core(B).

$$\alpha_B(X) = \frac{|B_*(X)|}{|B^*(X)|} \quad (1)$$

$$\mu_X^B(x) = \frac{|B(x) \cap X|}{|X|} \quad (2)$$

Decision rules can then be formed based on identified patterns in data using Eq(1) and using the definition of reducts. The support of a rule supp<sub>S</sub>(Φ, Ψ) refers to the number of samples which exhibit the characteristics of the condition, Φ, which results in a decision, Ψ (card(|Φ ∧ Ψ|)). The certainty factor, cer<sub>S</sub>(Φ, Ψ) of a rule also referred to as confidence coefficient is defined by Eq(3). It is the probability that objects exhibiting condition Φ results in decision Ψ. The coverage factor, cov<sub>S</sub>(Φ, Ψ) is defined by Eq(4) and refers to the degree to which samples belonging to a decision Ψ are explained by condition Φ. The strength of a rule, σ<sub>S</sub>(Φ, Ψ), is defined by

Eq(5) and just indicates the proportion of objects from the sample which is classified under the rule. Further details on the basic concepts of RST can be found in Pawlak (1997) and Pawlak (2002).

$$\text{cer}_S(\Phi, \Psi) = \frac{\text{card}(\|\Phi \wedge \Psi\|)}{\text{card}(\Phi)} \quad (3)$$

$$\text{cov}_S(\Phi, \Psi) = \frac{\text{card}(\|\Phi \wedge \Psi\|)}{\text{card}(\Psi)} \quad (4)$$

$$\sigma_S(\Phi, \Psi) = \frac{\text{supp}_S(\Phi, \Psi)}{\text{card}(U)} \quad (5)$$

#### 4. Case Study on Sustainable Cities Index

The case study considers the 2016 Sustainability Cities Index (SCI), which evaluates 100 major urban centres globally using social (“People”), environmental (“Planet”), and economic (“Profit”) indicators (Arcadis, 2020). The data was split into 18 condition attributes and 2 decision attributes; the latter are energy intensity per unit gross domestic product (GDP) and GHG emissions per capita. RSML is used to find patterns of association between the conditions and each decision attribute. The condition attributes are summarised in Table 1.

Table 1: Condition attributes of case study

Criteria	Sub-criteria	Reduct for energy intensity (D1)	Reduct for GHG emissions (D2)
Social	Demographics		
	Education		✓
	Income inequality		
	Work-life balance		
	Crime		
	Health		
	Affordability	✓	✓
Environmental	Environmental risks (EnvtlRisks)	✓	✓
	Green space (GreenSpace)	✓	✓
	Air pollution (AirPoll)	✓	✓
	Waste management (WasteMgmt)		✓
	Drinking water and sanitation		
Economic	Transport infrastructure (Transport)	✓	✓
	Economic development		
	Ease of doing business (Business)	✓	✓
	Tourism		
	Connectivity (Connect)	✓	
	Employment		✓

Of the 100 cities listed in SCI, 60 were selected for use as training data, while the remaining 40 was set aside as a validation data set. The original dimensionless SCI scores, which range from 0 to 1, were discretised into three categories, where the first category labelled as “Good” (coded as “1”) have scores in the upper quarter ( $C_{ij}$  or  $D_{kj} \geq 0.75$ ), the second category labelled as “Moderate” (coded as “2”) have scores in the two middle quarters ( $0.25 \leq C_{ij}$  or  $D_{kj} < 0.75$ ), and the third category labelled as “Poor” (code as “3”) have scores in the bottom quarter ( $C_{ij}$  or  $D_{kj} < 0.25$ ). RSML was implemented using the software tool ROSE 2.0 (Predki et al., 1998), which can be downloaded for free from the developer’s research group website (IDSS, 2020). More than 100 different reducts were generated for each decision attribute. Note that the presence of an attribute in a reduct indicates the existence of a recurring pattern in the data but does not necessarily indicate direct causality. Reduct selection was conducted by inspecting and screening the generated reducts for both plausibility and overlaps (i.e., factors influencing energy intensity are also likely to affect GHG emissions). In general, reduct selection should account for both data aspects and the user’s domain knowledge (Jia et al., 2016). The reducts selected for the two decision attributes are also shown in Table 1, showing that some sub-criteria appear in both reduct sets.

Using the selected reducts, training was conducted to generate the rules using the “satisfactory description” algorithm of ROSE 2.0 (i.e., maximum length = 3, minimum strength = 30 %, minimum discrimination = 90 %). The rules generated for energy intensity are summarised in Table 2. It can be seen that there are no rules that predict cities with good performance, which indicates that there are no detectable patterns among such cities in the training data. The rules for GHG emissions are shown in Table 3. The corresponding statistical performance

metrics are also shown in the tables. The rules can be interpreted linguistically as an if/then statement; for example, Rule 1 in Table 2 can be stated as: “If environmental risks score is poor then energy intensity score is moderate.” Note that the rules reflect patterns of association but do not necessarily imply causality. For each decision variable, the rules combine disjunctively (i.e., they are linked together via logical “or”) to give a rule-based model that summarises patterns in the training data. Within each rule-based model in Table 2 and 3, some rules or patterns are more consistent (i.e., with higher certainty) and more prevalent (i.e., with greater support, coverage, and strength) than others. The implications of such differences are discussed below.

*Table 2: Rules for classifying cities based on energy intensity (“Energy”) and associated statistics*

Rules	supp <sub>S</sub>	cer <sub>S</sub>	cov <sub>S</sub>	σ <sub>S</sub>
1 (EnvtlRisks = 3) → (Energy = 2)	16	0.94	0.33	0.27
2 (Business = 1) → (Energy = 2)	24	0.96	0.49	0.40
3 (GreenSpace = 2) & (Transport = 2) → (Energy = 2)	16	0.94	0.33	0.27
4 (GreenSpace = 2) & (Connect = 2) → (Energy = 2)	18	0.95	0.37	0.30
5 (EnvtlRisks = 1) & (Transport = 3) → (Energy = 3)	2	1.00	0.40	0.03
6 (AirPoll = 3) & (Connect = 1) → (Energy = 3)	3	1.00	0.60	0.05
7 (Affordability = 2) & (EnvtlRisks = 1) & (AirPoll = 3) → (Energy = 3)	3	1.00	0.60	0.05
8 (Affordability = 2) & (AirPoll = 3) & (Business = 2) → (Energy = 3)	3	1.00	0.60	0.05
9 (EnvtlRisks = 1) & (GreenSpace = 3) & (AirPoll = 3) → (Energy = 3)	3	1.00	0.60	0.05
10 (GreenSpace = 3) & (AirPoll = 3) & (Business = 2) → (Energy = 3)	3	1.00	0.60	0.05

*Table 3: Rules for classifying cities based on GHG emissions and associated statistics*

Rule	supp <sub>S</sub>	cer <sub>S</sub>	cov <sub>S</sub>	σ <sub>S</sub>
1 (Affordability = 1) & (Employment = 3) → (GHG = 1)	7	1.00	0.30	0.12
2 (EnvtlRisks = 2) & (AirPoll = 2) → (GHG = 1)	8	1.00	0.35	0.13
3 (AirPoll = 2) & (WasteMgmt = 3) → (GHG = 1)	9	1.00	0.39	0.15
4 (WasteMgmt = 3) & (Business = 3) → (GHG = 1)	8	1.00	0.35	0.13
5 (Transport = 3) & (Business = 3) → (GHG = 1)	9	0.90	0.39	0.15
6 (Affordability = 2) & (GreenSpace = 2) & (AirPoll = 1) & (WasteMgmt = 2) → (GHG = 2)	9	0.90	0.31	0.15
7 (Affordability = 2) & (EnvtlRisks = 1) & (AirPoll = 3) → (GHG = 3)	3	1.00	0.38	0.05
8 (Affordability = 2) & (AirPoll = 3) & (Business = 2) → (GHG = 3)	3	1.00	0.38	0.05
9 (EnvtlRisks = 1) & (GreenSpace = 3) & (AirPoll = 3) → (GHG = 3)	3	1.00	0.38	0.05
10 (GreenSpace = 3) & (AirPoll = 3) & (Business = 2) → (GHG = 3)	3	1.00	0.38	0.05
11 (Education = 2) & (EnvtlRisks = 1) & (WasteMgmt = 2) & (Employment = 1) → (GHG = 3)	3	1.00	0.38	0.05

*Table 4: Confusion matrix for energy intensity model*

Actual category	Predicted category			
	1	2	3	Unclassified
1	0	5	1	3
2	0	23	0	5
3	0	1.5	1.5	0

Both rule-based models were then tested on the 40 cities in the validation data set. The confusion matrices are shown in Tables 4 and 5. They show the classification performance and errors that occur when the models encounter a new set of data that was not used during training. It can be seen that many of the cities in the validation data are unclassified (i.e., their scores do not activate any of the rules in the models). For the energy intensity model, data of three cities activated multiple conflicting rules, which resulted in non-deterministic classification as either moderate or poor performers. The capability to handle such conflicts is a central feature of rough set theory and RSML; these cities were counted fractionally (i.e., 0.5 for moderate and 0.5 for poor) due to their simultaneous classification in these two categories. In the energy intensity rule-based model, Rules 6–10 did not match any of the cities in the validation data. In the GHG emissions model, Rules 7–10 were also not activated. Note that these rules are characterised in Tables 2 and 3 by small support sets and low strength.

Their low generalisation power can be attributed to spurious patterns in the training data; the final rule-based models can be reduced in size and complexity by eliminating these rules.

Table 5: Confusion matrix for GHG emissions model

Actual category	Predicted category			
	1	2	3	Unclassified
1	12	1	0	5
2	3	2	0	15
3	1	0	0	1

## 5. Practical Implications

In this section, some practical insights drawn from the two reduced rule-based models are discussed. Air pollution appears in the reducts for both energy intensity and GHG emissions. High (“poor”) energy intensity suggests the energy inefficiency of an industrialized economy, which is associated with high (“poor”) levels of air pollution. This link is seen in Rules 6–10 in Table 2. However, the results in Tables 3 suggest a stronger link to GHG emissions as indicated by Rules 2,3, 6–10. This suggests the validation set consist of cities with different combinations of GDP level and rules on the emission limit (application of filtration), offsetting the impact of air pollutants on energy intensity.

Rule 6 in Table 3 reflects a circumstance that “poor” air pollution performance is not associated with “poor” GHG emissions levels. This rule shows that a sound solution based on GHG emissions is insufficient in ensuring that the environmental sustainability, as the air pollution aspects could still be “poor”. It is inconclusive to assume a GHG-optimised solution is automatically environmentally sustainable. Simultaneous optimisation studies, considering the synergistic effects, should be encouraged.

Waste management (WasteMgmt) which accounts for the portion of landfilled solid waste and share of wastewater treated, shows is negatively associated with GHG emissions. Rules 3 and 4 (Table 3) show “good” waste management leads to high (“poor”) GHG emissions, for which two explanations can be drawn: (i) high recycling rates do not lead to low GHG emission; or (ii) “good” waste management could unintentionally lead to high waste generation, suggesting potential rebound effect. The burdening impact of recycling has been discussed in some of the studies (Bernstad Saraiva et al., 2018). This impact is apparent in less volatile waste (e.g., plastic), where the GHG emission of the waste ending up in landfills is less significant than the energy (with GHG emissions from energy sources) invested for recycling. However, it only reflects the global warming potential, while recycling is still preferable in terms of other impact categories such as eutrophication and land use. This rule emphasises the complex trade-offs in waste management and the dominant roles of waste generation in GHG emission, where the recovery has a less pronounced role in minimising the GHG emission. Other attributes with no clear causality appear in weak rules which may result from spurious patterns in the SCI data (e.g., education in Rule 11 of Table 3).

## 6. Conclusions

In this work, two rule-based ternary classification models for cities were generated from the SCI data using RSML. The models categorise cities into good, moderate, and poor performers with respect to energy intensity and GHG aspects, based on different social, environmental, and economic attributes. Both models exhibit good predictive performance for covered samples in both training and validation data; however, many cities fail to exhibit detectable patterns and are not covered by the models’ rules. These initial results show that there are no universal rules which could be applied globally. The rule-based models can be used as aids for urban planning to induce a shift to more sustainable development trajectories.

This work can be extended further using RSML or other ML tools. A more apparent set of trends might be observable by first clustering the cities based on decisive characteristics prior to using RSML, especially for energy intensity as it is relatively a localised issue compared to GHG emission. Future work also should focus on exploration of alternative reducts and rule sets coupled with k-fold validation.

## Acknowledgements

The financial support from the EU supported project Sustainable Process Integration Laboratory – SPIL funded as project No. CZ.02.1.01/0.0/0.0/15\_003/0000456, by Czech Republic Operational Programme Research and Development, Education, Priority 1: Strengthening capacity for quality research under the collaboration agreement with De La Salle University is gratefully acknowledged.

## References

- Arcadis, 2020, Citizen centric cities <[www.arcadis.com/campaigns/citizencentriccities/index.html](http://www.arcadis.com/campaigns/citizencentriccities/index.html)> accessed 10.10.2020.
- Aviso K.B., Janairo J.I.B., Promentilla M.A.B., Tan R.R., 2019, Prediction of CO<sub>2</sub> storage site integrity with rough set-based machine learning, *Clean Technologies and Environmental Policy*, 21, 1655–1664.
- Aviso K.B., Janairo J.I.B., Lucas R.I.G., Promentilla M.A.B., Yu D.E.C., Tan R.R., 2020, Predicting higher education outcomes with hyperbox machine learning: What factors influence graduate employability? *Chemical Engineering Transactions*, 81, 679–684.
- Bernstad Saraiva A., Souza R.G., Mahler C.F., Valle R.A.B., 2018, Consequential lifecycle modelling of solid waste management systems – Reviewing choices and exploring their consequences, *Journal of Cleaner Production*, 202, 488–496.
- Calegari R., Ciatto G., Omicini A., 2020, On the integration of symbolic and sub-symbolic techniques for XAI: A survey, *Intelligenza Artificiale*, 14, 7–32.
- Gue I.H.V., Tan R.R., Ubando A.T., 2021, Causal network maps of urban circular economies, *Clean Technologies and Environmental Policy*, in press, DOI: 10.1007/s10098-021-02117-9.
- IDSS, 2020, ROSE2, Laboratory of Intelligent Decision Support Systems, Poznań University of Technology <[hidss.cs.put.poznan.pl/site/rose.html](http://hidss.cs.put.poznan.pl/site/rose.html)> accessed 12.12.2020.
- Jia X., Shang L., Zhou B., Yao Y., 2016, Generalized attribute reduct in rough set theory, *Knowledge-Based Systems*, 91, 204–218.
- Jordan M.I., Mitchell T.M., 2015, Machine learning: Trends, perspectives, and prospects, *Science*, 349, 255–260.
- Kamyab H., Klemeš J.J., Fan Y.V., Lee C.T., 2020, Transition to sustainable energy system for smart cities and industries, *Energy*, 207, Article 118104.
- Kennedy C.A., Stewart I., Facchini A., Cersosimo I., Mele R., Chen B., Uda M., Kansal A., Chiu A., Kim K.-G., Dubeux C., La Rovere E.L., 2015, Energy and material flows of megacities, *Proceedings of the National Academy of Sciences of the United States of America*, 112, 5985–5990.
- Komorowski J., 2014, Learning rule-based models: The rough set approach, In: Brahme A (Ed.), *Comprehensive biomedical physics*, Vol. 6, Elsevier, Amsterdam, The Netherlands, 19–39.
- Lopez N.S., Mouy M., Africa A.D., 2021, Uncovering the significant socioeconomic attributes of low- and high-emission countries using rough set, *Clean Technologies and Environmental Policy*, in press, DOI: 10.1007/s10098-021-02067-2.
- Ma H., Chang W., Cui G., 2012, Ecological footprint model using the support vector machine technique, *PLoS ONE*, 7, Article e30396.
- Mahajan P., Kandwal R., Vijay R., 2012, Rough set approach in machine learning: a review, *International Journal of Computer Applications*, 56, 1–13.
- Mostafa M.M., 2010, Clustering the ecological footprint of nations using Kohonen's self-organising maps, *Expert Systems with Applications*, 37, 2747–2755.
- Pawlak Z., 1982, Rough sets, *International Journal of Computer and Information Sciences*, 11, 341–356.
- Pawlak Z., 1984, Rough classification, *International Journal of Man-Machine Studies*, 20, 469–483.
- Pawlak Z., 1997, Rough set approach to knowledge-based decision support, *European Journal of Operational Research*, 99, 48–57.
- Pawlak Z., 2002, Rough sets, decision algorithms and Bayes' theorem, *European Journal of Operational Research*, 136, 181–189.
- Predki B., Slowinski R., Stefanowski J., Susmaga R., Wilk S., 1998, ROSE - Software implementation of the rough set theory. In: L Polkowski, A Skowron (Eds.), *Rough sets and current trends in computing, Lecture notes in artificial intelligence*, Vol. 1424, Springer-Verlag, Berlin, Germany, 605–608.
- Radhakrishnapany K.T., Wong C.Y., Tan F.K., Chong J.W., Tan R.R., Aviso K.B., Janairo J.I.B., Chemmangattuvalappil N.G., 2020, Design of fragrant molecules through the incorporation of rough sets into computer-aided molecular design, *Molecular Systems Design and Engineering*, 5, 1391–1416.
- Ramaswami A., Tong K., Canadell J.G., Jackson R.B., Stokes E., Dhakal S., Finch M., Jittrapirom P., Singh N., Yamagata Y., Yewdall E., Yona L., Seto K.C., 2021, Carbon analytics for net-zero emissions sustainable cities, *Nature Sustainability*, 4, 460–463.
- Sugiawan Y., Managi S., 2019, New evidence of energy-growth nexus from inclusive wealth, *Renewable and Sustainable Energy Reviews*, 103, 40–48.
- Tamascelli N., Arslan T., Shah S.L., Paltrinieri N., Cozzani V., 2020, A machine learning approach to predict chattering alarms, *Chemical Engineering Transactions*, 82, 187–192.
- Yao Y., 2011, The superiority of three-way decisions in probabilistic rough set models, *Information Sciences*, 181, 1080–1096.