



Article citation info:

Rusek J, Firek K, Słowik L. Extracting structure of Bayesian network from data in predicting the damage of prefabricated reinforced concrete buildings in mining areas. *Eksploracja i Niezawodność – Maintenance and Reliability* 2020; 22 (4): 658–666, <http://dx.doi.org/10.17531/ein.2020.4.9>.

Extracting structure of Bayesian network from data in predicting the damage of prefabricated reinforced concrete buildings in mining areas

Indexed by:



Janusz Rusek^{a*}, Karol Firek^a, Leszek Słowik^b

^aAGH University of Science and Technology, al. Mickiewicza 30, 30-059 Krakow, Poland

^bBuilding Research Institute ITB, ul. Filtrowa 1, 00-611 Warsaw, Poland

Highlights

- Damage risk model for reinforced concrete (RC) buildings subject to mining impacts.
- The Bayesian network methodology was applied
- The network structure was selected using greedy-search learning algorithms.
- The results for the Hill-Climbing and Tabu learning methods were compared.
- The final network structure and the optimal learning criterion were determined.

Abstract

This article presents the results of the research on the construction of a model for assessing the risk of damage to building structures located in mining areas. The research was based on the database on the structure, technical condition and mining impacts regarding 129 prefabricated reinforced concrete buildings erected in the industrialised large-block system, located in the mining area of the Legnica-Glogow Copper District (LGCD). The methodology of the Bayesian Belief Network (BBN) was used for the analysis. Using the score-based Bayesian structure learning approach (Hill-Climbing and Tabu-Search) as well as the selected optimisation criteria, 16 Bayesian network structures were induced. All models were subjected to quantitative and qualitative evaluation by verifying their features in the context of accuracy of prediction, generalisation of acquired knowledge and cause-effect relationships. This allowed to select the best network structure together with the corresponding optimisation criterion. The analysis of the results demonstrated that the Tabu-Search method adopting the optimisation criterion in the form of Locally Averaged Bayesian Dirichlet score (BDla) led to obtaining a model with the best features among all the selected models. The results justified the adoption of the BBN methodology as effective in the context of assessing the extent of damage to building structures in mining areas.

Keywords

Bayesian network, damage, risk, prefabricated reinforced concrete structures.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

1. Introduction

Damage to existing building structures in mining areas can be caused by a number of construction and environmental factors, including mining impacts. They can be seen on the surface in the form of continuous or non-continuous ground deformations [6, 38] and mining tremors [23, 41]. The damage may concern both structural and secondary members [8]. This is also the case with prefabricated reinforced concrete buildings erected in industrialised construction systems [15], currently accommodating circa 1/3 of Poland's inhabitants (approx. 12 million people). In addition, these are usually large-sized multi-family or public buildings, where disturbances in the comfort of use resulting from their damage are of great importance.

Most frequently, the process of damage is initiated not by a single factor but by many factors at the same time. This makes the problem regarding the evaluation of the causes of damage complex and difficult to describe from an analytical point of view. In the context of mining impacts, the situation is further complicated by uncertainty regarding predictions of deformations that may occur on the ground surface as well as mining tremors [19].

In practice, it is often necessary to carry out such an evaluation procedure for a large number of building structures erected on a given mining area. This fact disqualifies the FEM numerical approach, which is ineffective in this case. Statistical models seem to be the only way to deal with the problem.

A group of popular and very effective methods for detecting damage in concrete structures also includes the so-called non-destructive methods, such as the *Acoustic Emission* (AE) [22], the *Digital Image Correlation* (DIC) [10] or the ultrasonic [31] methods. In addition to crack detection itself, non-destructive methods allow a more detailed analysis of reinforced concrete member degradation, which also includes determination of reinforcing bar corrosion [28]. This is particularly important in assessing damage in prefabricated large-block buildings, where the basic issue is to determine the technical condition of joints and their reinforcement. Moreover, in recent years, alternative approaches based on the use of *Machine Learning* (ML) methods have become increasingly popular, e.g. [28, 29].

Given the specificity of the problem being analysed, especially in the context of predicting damage to a large number of building struc-

(*) Corresponding author.

E-mail addresses: J. Rusek - rusek@agh.edu.pl, K. Firek - kfirek@agh.edu.pl, L. Słowik - l.slowik@itb.pl

tures, ML methods have been selected for further research, as they allow the presentation of the phenomenon in probability notation.

The effectiveness of ML tools has been demonstrated in the previously performed research studies [25, 27]. However, all of the tested methods assumed mutual independence of potential causes that might affect the damage process. For this reason, it was decided to use the so-called *Bayesian Network* (BN) or *Bayesian Belief Network* (BBN) methodology which, according to [24], allows the possibility of taking conditional independence into account. This allows to build a model in which the correlations between individual factors are closer to the physical reality associated with the damage emergence process. As a result, the established BN structure can be identified with the damage risk model.

An additional feature of *Bayesian networks*, which is a very important advantage as far as the issue of assessing damage to building structures subject to e.g. mining impacts is concerned, is the possibility to use the created model not only for prediction, but also for the diagnosis of damage. This allows for a much broader implementation of such a system in practice.

Unfortunately, in the case of a large number of variables describing the modelled process, the lack of knowledge of the *Bayesian network* structure is frequently a significant problem. There are methods, however, that allow to induce the network structure from data. According to [18], they are divided into 3 groups: *constraint-based* structure learning, *score-based* structure learning and *hybrid* approach. The *score-based* approach has been used in this research paper, and the reasons for this selection are explained in Chapter 4.

In order to identify the optimal structure of the *Bayesian network*, two algorithms were tested: *Hill-Climbing* (HC) and *Tabu-Search* (TS), which belong to the score-based structure learning methods [18].

A database containing information on damage to a group of 129 prefabricated reinforced concrete structures erected in the WBL (large-block) industrialised system located in the LGCD mining area formed the basis for the analysis (Fig. 1). The information on the technical condition, including damage, was collected during the “*in-situ*” surveys and based on the reports. These data were used to induce the structures of *Bayesian networks* and for their subsequent testing. The quality of the created models was verified in the context of correct classification and generalisation features. Additionally, having an explicit representation of correlations between network nodes, verification was carried out in terms of model compliance with the observed, real nature of the modelled process. It was a subjective evaluation based on expert knowledge.

Based on the results obtained, the structure of the *Bayesian network* was induced, which could represent a complex damage risk model.



The basis for the evaluation here was primarily the compliance of model classification with learning and testing data.

2. Literature overview

The methodology of *Bayesian Networks* (BN) or *Bayesian Belief Networks* (BBN) is currently widely used in the analysis of risk [3], security [12], reliability [43] and predicting the extent of damage [16]. In recent years, implementation of this type of methodology in the field of civil engineering has become increasingly popular. However, the expert approach dominates here, where the *Bayesian network* structure is imposed, and only model parameters are subject to learning. When specifying areas convergent to the subject matter discussed in this research paper, the issues related to the analysis of the risk of damage in building structures could be mentioned. The subject of these analyses are mainly linear structures, such as pipelines, bridge structures [1] or tunnels [40], but there are also examples regarding the evaluation of the risk of damage in buildings [30]. *Bayesian networks* are also used in a broader sense regarding reliability analyses [39], including safety assessment [5]. This methodology is used to construct diagnostic systems for building structure maintenance management [20], as well as for the assessment of their technical condition [21]. In the narrower range, they also appear as systems that allow the evaluation of strength parameters of existing structures, both static [4] and fatigue [44] ones.

As far as the diversity of building construction types is concerned, the *Bayesian network* methodology is also used for masonry [17] and steel structures [7]. In a broader context, it is also frequently used as a tool to predict random environmental impacts such as floods, earthquakes, tsunamis [13], climate change [37] or land settlement [42]. The information obtained from such analyses is, in turn, used at both the design stage and the risk assessment of existing building structures.

It should be emphasised that the *Bayesian network* methodology has been recently applied in construction engineering also in Poland, e.g. to assess the influence of traffic vibrations on surface development [36] and to analyse the risk in construction investment of a tunnel under the *Martwa Wisła* river in *Gdańsk* [14].

However, despite numerous applications of *Bayesian network* methodology in civil engineering, the use and development of methods for network structure discovery from data is still the domain of the sciences such as medicine, biology, genetics, e.g. [9]. For this reason, the research presented in this paper is an extension of the methodology used so far in civil engineering, which may allow for the analysis of more complex engineering problems regarding damage risk analysis or structural reliability. The issue being discussed, on *Bayesian network* structure discovery from the data for constructing a damage risk model



Fig. 1. Examples of buildings constructed using WBL (large-block) industrialised technology (own source)

for building structures subject to mining impacts is, according to the authors, an interdisciplinary and innovative approach to the subject.

It should be noted that the *Bayesian network* methodology, in addition to the issues related to the reliability and safety of building structures, is also an indirect tool used in the risk analysis.

3. Characteristics of the database

The database which formed the basis of the research contained detailed information on a group of 129 prefabricated reinforced concrete large-sized multi-family and public buildings erected in the industrialised large-block WBL system, e.g. [26]. In each case, at the location of the building, indices describing the intensity of mining impacts for the duration of the survey were determined (**MC**, **AH** and **ASG** - see

Table 1). The **ASG** variable was determined based on the original index of mining tremors a_{sg} [41] and it expresses the effect of multiple impacts of mining tremors on the technical wear of buildings. The database was supplemented with information on construction and material properties, history and quality of maintenance, repairs and recorded damage. For damage, the damage index w_u developed by one of the authors of this research paper was used, referred to both structural members and finishing elements [8].

During the initial database analysis, all variables were categorised for further use during learning *Bayesian networks*. It involved assigning labels to individual categories. A list of all variables with the assigned range of labels is presented in Table 1. At the later stage, each variable will reflect individual nodes in the *Bayesian network*.

Table 1. List of variables in the database with the assigned label range

LIST OF VARIABLES		
DENOTATION	DESCRIPTION	TYPE OF VARIABLE/ VARIABLE LABEL
DATA ON MINING IMPACTS AT THE LOCATION OF A GIVEN BUILDING STRUCTURE		
MC	mining area category	category / 3 categories
AH	maximum horizontal component of acceleration of ground vibrations	category / 4 categories
ASG	mining tremors intensity index [41]	category / 4 categories
INDICES OF DAMAGE TO STRUCTURAL MEMBERS AND FINISHING ELEMENTS		
w_{u2}	index of damage to basement or foundation walls	category / 4 categories
w_{u3}	index of damage to overground internal and external load-bearing walls (including lintels and spandrels)	category / 4 categories
w_{u7}	index of damage to higher ceilings, flat roof (covering)	category / 4 categories
w_{u11}	index of damage to partition walls	category / 2 categories
w_{u12}	index of damage to internal plasters and wall coverings	category / 3 categories
w_{u13}	index of damage to floors (floor layers)	category / 3 categories
w_{u17}	index of damage to façade (façade layers)	category / 4 categories
w_{u19}	index of damage to roofing	category / 4 categories
DATA ON MAINTENANCE AND REPAIRS		
CR	current repairs	category / 4 categories
FR	façade repairs	category / 5 categories
RR	roof repairs	category / 6 categories
IR	interior renovations	category / 7 categories
CONSTRUCTION AND BUILDING GEOMETRY DATA		
LEN	building length (longer of the dimensions)	category / 6 categories
NoST	number of storeys	category / 6 categories
NoSE	number of segments	category / 8 categories
LoSE	length of segment	category / 6 categories
DIL	dilatation (width)	category / 3 categories
SHA	building shape	category / 4 categories
ToF	type of foundation	category / 5 categories
SW	curtain walls	category / 4 categories
BC	basement ceiling	category / 4 categories
DP	design protection	category / 3 categories
EP	existing protection	category / 5 categories
DATA ON DURABILITY		
AGE	building age	category / 4 categories
DUR	durability	category / 5 categories
TS	technical state	category / 4 categories

4. Research methodology

Bayesian Networks (BN), also known as Bayesian Belief Networks (BBN) can be represented in the form of a Directed Acyclic Graph (DAG) [18]. The structure of the graph (G) encodes information about correlations between individual variables $X = \{X_1, \dots, X_N\}$, which is represented by the edges of the graph (E) and nodes (V). In general, BBN represents the total probability distribution over a set of random variables X , which can be presented as [34]:

$$P(X | G, \Theta) = \prod_{i=1}^N P(X_i | \Pi_{X_i}, \Theta_{X_i}) \quad (1)$$

gdzie:

$G = G(X, E, V)$ – a structure of an acyclic directed graph

$X = \{X_1, \dots, X_N\}$ – a set of all variables present in the graph nodes

$X_i = \{x_i^{(1)}, \dots, x_i^{(k_i)}\}$ – states of j -th variable

E – a set of all edges

V – a set of all nodes

$\Pi_{X_i} = \{x_i^{(q_1)}, \dots, x_i^{(q_i)}\}$ – a set of parents, i.e. all nodes of the graph determining the state of the node X_i

$\theta = \{\theta_{X_1}, \dots, \theta_{X_N}\}$ – a set of all parameters of conditional relations between individual nodes X_i , and a set of their parents Π_{X_i}

In the case of discrete variables, the parameters of the model $\theta_{X_j} = \{\theta_{ijk}\}$ are represented in the form of a multivariate Conditional Probability Table (CPT), whose elements are expressed as [11]:

$$\theta_{ijk} = P(X_j = x_j^{(i)} | \Pi_{X_j} = \pi_j^{(k)}) \quad (2)$$

According to the relationship (1), the total probability distribution $P(X | G, \Theta)$ is subject to decomposition based on conditional local

distributions $P(X_i | \Pi_{X_i}, \Theta_{X_i})$, described over each random variable X_i relative to the set of conditioning variables corresponding to it, the so-called parents Π_{X_i} . This formulation is possible thanks to the concept of conditional independence introduced by Pearl in [24]. It also allows for an effective analysis of changes in the value of the adopted criterion during the search for the optimal network structure.

In the described problem concerning the construction of the damage risk model, both BBN structure and its parameters are unknown. In the previous research, significant, but subtle influence of individual factors on the damage process has been confirmed. Therefore, at the stage of selection of the BBN structure learning method from data, it was important that the final model had the largest possible number of variables potentially affecting the process of damage initiation. Based on the literature [18], constraint-based structure learning methods and a hybrid approach were rejected. It was found that in these learning methods, more weight is attributed to individual correlations between network nodes than to the global response of the model. Thus, the focus was placed on score-based structure learning methods. Eventually, two methods were used: Hill-Climbing (HC) and Tabu-Search (TS).

The HC optimisation approach is one of the greedy-search methods [35]. It consists in searching the closest environment around the current point in space to which a given value of the adopted optimisation criterion corresponds. Then, a point in the search space whose criterion value is higher than in the previous step is selected from the environment. The search space is the space of the structures of Directed Acyclic Graphs (DAGs), and the algorithm advances by adding, subtracting or swapping correlations between nodes.

The TS approach is a modification of the HC algorithm. This modification consists in storing a certain number of forbidden paths in the search space that have already been traversed in previous iterations, and thus limiting blind searching and getting stuck in the local optimum [33].

In this research paper, using the *bnlearn* package, a comparative analysis of both approaches was undertaken. In addition, their performance was tested when various optimisation criteria were adopted. For this purpose, criteria belonging to two groups were selected: Information-Theoretic scores (IT) [2] and Bayesian Dirichlet scores (BD) [32]. The first group included: Log-Likelihood (LL), Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). From the Bayesian Dirichlet scores (BD) group, according to [33], the following were used: Bayesian Dirichlet equivalent uniform (BDeu), modified Bayesian Dirichlet equivalent score (mBDe), Bayesian Dirichlet sparse score (BDs), locally averaged Bayesian Dirichlet score (BDla) and the K2 score. The chronological diagram of the performed research is presented in Figure 2.

5. Results of the conducted analyses

The research used the methodology described in Chapter 3, which assumed two approaches to the construction of the Bayesian belief network (BBN) structure from data. Both approaches belong to the group of score-based structure learning methods [18]. The Hill-Climbing (HC) algorithm was used as the first one. In the second one, its modification, the Tabu-Search (TS) algorithm was used.

The data set on building structures described in Chapter 2 was divided into training and test sets. The key here was to separate the sets in such a way that they maintained an even distribution of the values of the categorised variables listed in Table 1, i.e.: w_{u2} , w_{u3} , w_{u7} , w_{u11} , w_{u12} , w_{u13} , w_{u17} and w_{u19} . This was finally done using the Stratified Sampling (SS) method. As a result, a training set with 105 patterns (81.36% of the total number of patterns in the database) and a test set whose number of patterns was 24 (18.6%) were obtained.

As far as the *bnlearn* package used in the analyses is concerned, it is possible to take the expert knowledge into account. This is carried out by entering a list representing forbidden (Blacklist) and forced

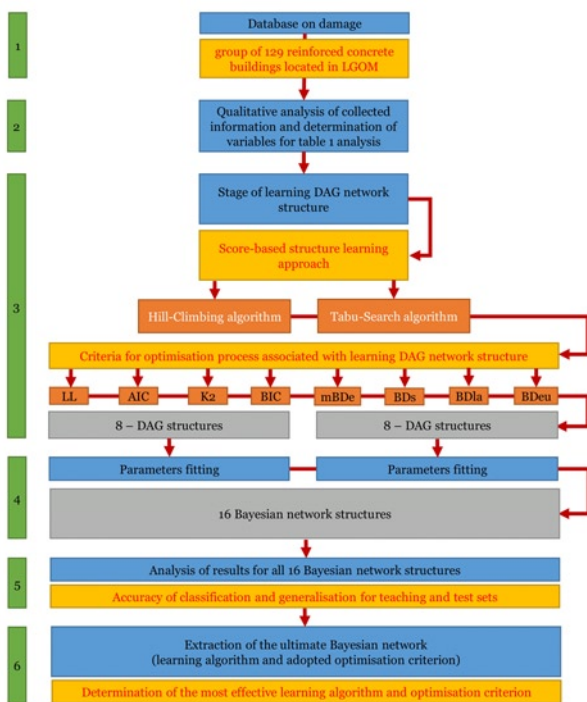


Fig. 2. Chronological diagram of the carried out analyses

Table 2. Accuracy of classification for individual nodes corresponding to damage indices - training set - algorithm teaching HC

criterion	Accuracy level ac^{HC-TR} [%]							
	wu ₂	wu ₃	wu ₇	wu ₁₁	wu ₁₂	wu ₁₃	wu ₁₇	wu ₁₉
LL	100,00	99,05	99,05	100,00	100,00	100,00	99,05	99,05
AIC	92,38	90,48	85,71	92,38	94,29	95,24	90,48	90,48
BIC	93,33	90,48	89,52	92,38	94,29	90,48	90,48	90,48
BDeu	89,52	91,43	86,67	92,38	94,29	94,29	91,43	93,33
BDs	92,38	93,33	89,52	97,14	97,14	99,05	94,29	92,38
mBDe	91,43	91,43	85,71	92,38	94,29	94,29	91,43	93,33
BDla	90,48	93,33	85,71	97,14	97,14	95,24	95,24	94,29
K2	93,33	90,48	90,48	93,33	91,43	94,29	95,24	90,48

Table 3. Accuracy of classification for individual nodes corresponding to damage indices - test set - algorithm teaching HC

criterion	Accuracy level ac^{HC-TS} [%]							
	wu ₂	wu ₃	wu ₇	wu ₁₁	wu ₁₂	wu ₁₃	wu ₁₇	wu ₁₉
LL	95,83	95,83	79,17	100,00	100,00	100,00	79,17	95,83
AIC	83,33	87,50	75,00	95,83	95,83	100,00	37,50	95,83
BIC	87,50	87,50	79,17	91,67	91,67	91,67	41,67	95,83
BDeu	83,33	87,50	79,17	95,83	95,83	100,00	62,50	100,00
BDs	91,67	91,67	79,17	100,00	100,00	100,00	75,00	100,00
mBDe	83,33	87,50	83,33	95,83	95,83	100,00	66,67	100,00
BDla	83,33	91,67	83,33	100,00	100,00	100,00	87,50	100,00
K2	95,83	83,33	75,00	95,83	100,00	91,67	87,50	95,83

Table 4. Accuracy of classification for individual nodes corresponding to damage indices - training set - algorithm teaching TS

criterion	Accuracy level $ac^{TABU-TR}$ [%]							
	wu ₂	wu ₃	wu ₇	wu ₁₁	wu ₁₂	wu ₁₃	wu ₁₇	wu ₁₉
LL	100,00	99,05	100,00	100,00	100,00	100,00	100,00	100,00
AIC	92,38	90,48	85,71	92,38	94,29	96,19	90,48	90,48
BIC	93,33	90,48	89,52	92,38	94,29	90,48	90,48	90,48
BDeu	91,43	90,48	88,57	92,38	94,29	96,19	91,43	93,33
BDs	95,24	96,19	89,52	97,14	97,14	98,10	94,29	93,33
mBDe	91,43	89,52	86,67	93,33	94,29	94,29	90,48	93,33
BDla	91,43	93,33	86,67	97,14	97,14	95,24	95,24	93,33
K2	93,33	90,48	86,67	89,52	94,29	92,38	95,24	90,48

(Whitelist) correlations. These are strong constraints that are not subject to modification during the learning process.

This research paper summarises the results of the performed analyses, which aimed to compare the two approaches used. Eventually, 16 BBN structures were induced and submitted to evaluation. This

Table 5. Accuracy of classification for individual nodes corresponding to damage indices - test set - algorithm teaching TS

criterion	Accuracy level $ac^{TABU-TS}$ [%]							
	wu ₂	wu ₃	wu ₇	wu ₁₁	wu ₁₂	wu ₁₃	wu ₁₇	wu ₁₉
LL	95,83	95,83	79,17	100,00	100,00	100,00	79,17	95,83
AIC	83,33	87,50	70,83	95,83	95,83	100,00	50,00	95,83
BIC	87,50	87,50	79,17	91,67	91,67	91,67	54,17	95,83
BDeu	87,50	87,50	87,50	95,83	95,83	100,00	66,67	100,00
BDs	91,67	91,67	83,33	100,00	100,00	100,00	66,67	100,00
mBDe	83,33	87,50	83,33	95,83	95,83	100,00	66,67	100,00
BDla	83,33	91,67	87,50	95,83	100,00	100,00	79,17	100,00
K2	95,83	91,67	79,17	91,67	91,67	91,67	87,50	95,83

resulted from the adoption of two learning algorithms and eight optimisation criteria (*Information-Theoretic* scores (IT) and *Bayesian Dirichlet* scores (BD) - according to Chapter 3).

Before starting the calculations, strong constraints were introduced. They excluded from the network those correlations that were contrary to logic and were not observed in reality. In total, a set of 427 pairs between nodes was created, in which correlations were forbidden. It should be noted that no list of inductions was created, leaving freedom to individual learning algorithms.

In the **first stage**, individual structures were evaluated in terms of quantity. The basis here was the analysis of the accuracy of classification and generalisation features of the selected BBN structures. A relative measure was used here in the form of a percentage share of correctly classified patterns relative to the size of the entire data set (training and test set, respectively). These results for individual combinations resulting from the adopted learning method and optimisation criterion are contained in Tables 2 ÷ 5. It allowed to assess the accuracy of classification and generalisation features of the created models. The basis here was simulating network responses in nodes corresponding to individual damage indices: w_{u2} , w_{u3} , w_{u7} , w_{u11} , w_{u12} , w_{u13} , w_{u17} and w_{u19} .

In the second stage, the obtained BBN structures were verified in terms of quality. This was dictated by practical considerations resulting from the possibility of later use of the created model to assess risk in construction. Here, the focus was placed on the detected cause-effect relationships. The evaluation was made in an expert manner, based on the authors' experience in the field of damage to building structures in mining areas.

Based on the results of the quantitative analysis, which is represented by the level of correctly classified patterns for training and test sets described by: ac^{HC-TR} , ac^{HC-TS} , $ac^{TABU-TR}$, $ac^{TABU-TS}$ (Tables 2 ÷ 5), it was found that:

- LL optimisation criterion leads to the best fitting of training patterns (Tables 2 and 4), but generates the largest classification errors for test sets (Tables 3 and 5). It results therefrom that for both the HC and TS methods, the induced BBN structures do not have good generalisation features and lead to overfitting a model,
- the worst results were obtained for the prediction of damage indices w_{u7} and w_{u17} - see Tables 3 and 5. The corresponding results in the training set reached the value of the correct classification at the accuracy level $ac^{HC-TR} \approx ac^{HC-TS} \approx 90$. The difference between the results for the training and test sets indicates overfitting the model in the prediction of w_{u7} and w_{u17} indices.

- the best results in relation to the prediction of the values of damage indices w_{u7} and w_{u17} , both in the training and test sets, were obtained by adopting the BDla and K2 measures as the optimisation criteria - see Tables 3 and 5,
- in the case of prediction of other damage indices: w_{u2} , w_{u3} , w_{u11} , w_{u12} , w_{u13} and w_{u19} , apart from the LL criterion, both learning methods generate models achieving a very high level of accuracy in simulations, both in the training and test sets. The choice of individual optimisation criteria has no significant influence here.

As part of the **first stage**, i.e. the quantitative evaluation, a total of 4 structures were induced for which the best prediction results were obtained for all damage indices (including the w_{u7} and w_{u17} indices) and their generalisation features were confirmed. These were structures created by the HC and TS methods using the optimisation criteria BDla and K2. The results in graphical form are presented in Figures 3 to 6.

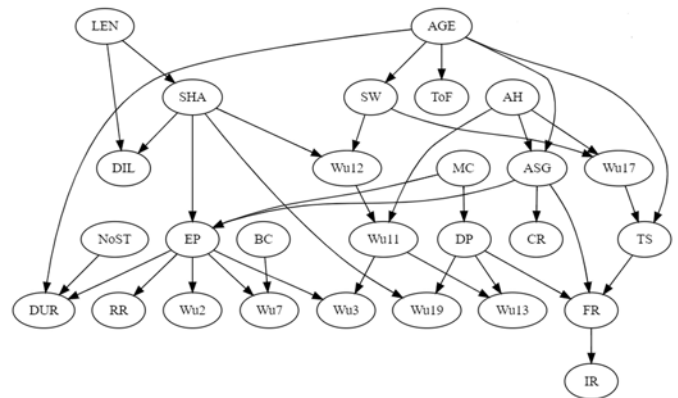


Fig. 3. The BBN structure extracted by the HC method using the BDla optimisation criterion

In the case of the BBN structure obtained by the HC method and BDla criterion (Fig. 3), it can be noticed that two nodes representing variables are omitted: **NoSE** and **LoSE**. The same applies to the structure selected by the TS method and the BDla criterion. Only one node indicating the **LoSE** variable is omitted here. The **NoSE** variable appears in the overall structure of the BBN network - Fig. 4.

On the other hand, in the case of the K2 criterion, regardless of the learning algorithm used, the **LoSE**, **NoSE**, **LEN**, **DIL**, and **SHA** nodes remain outside the selected structures (Figures 5 and 6).

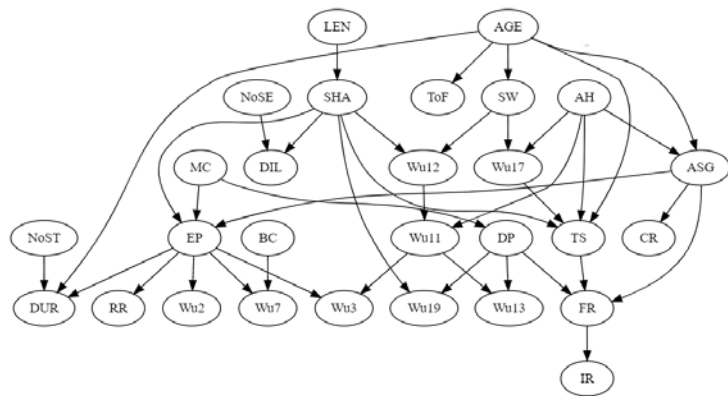


Fig. 4. The BBN structure extracted by the TS method using the BDla optimisation criterion

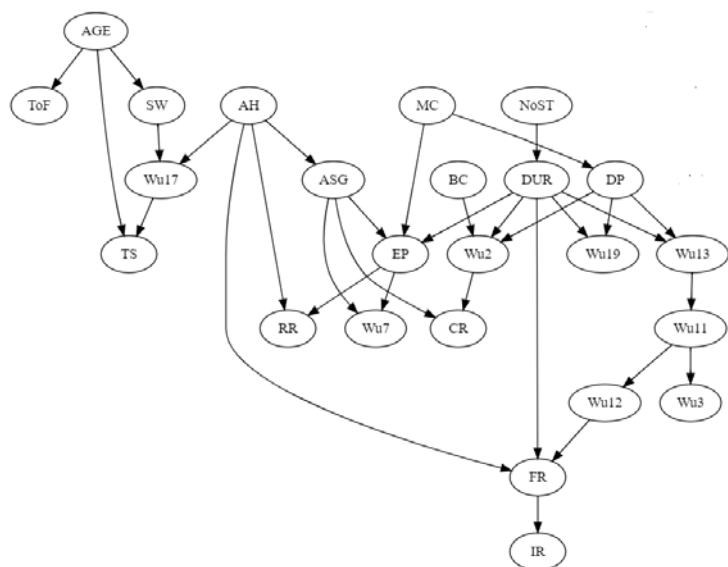


Fig. 5. The BBN structure extracted by the HC method using the K2 optimisation criterion

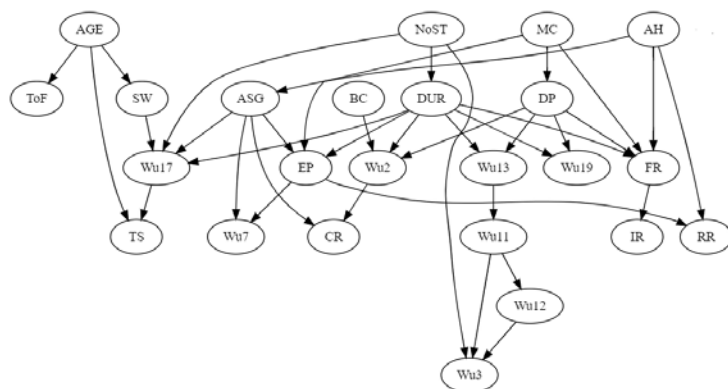


Fig. 6. The BBN structure extracted by the TS method using the K2 optimisation criterion

Geometrical features of a given object (see Table 1) influence its durability [27], and thus also on the risk of damage. Therefore, based on the authors' many years of experience, it seems that models represented by BBN network structures selected using the TS algorithm and the BDla criterion are closer to reality - see Figure 4. In this case,

the most complete description of the analysed problem was obtained, taking into account the set of active variables included in the BBN network represented by individual nodes.

6. Summary

The paper presents the results of analyses regarding the construction of the *Bayesian network* (BN) for predicting the extent and evaluation of damage to building structures subject to mining impacts. Due to the fact that the established *Bayesian network* represents the resulting likelihood of the occurrence of damage and all factors determining their emergence, it can become a model for assessing the risk of damage. Ultimately, this model was developed based on the information on 129 large-sized multi-family and public buildings with a prefabricated reinforced concrete structure erected in the WBL (large-block) industrialised construction system. The database collected information about the construction and material properties, quality of maintenance, as well as data on the evaluation of damage to members and mining impacts. Then, assuming two algorithms for learning *Bayesian network* structures from data (*Hill-Climbing* and *Tabu-Search*) and 8 different optimisation criteria, 16 *Bayesian network* structures were induced for the purpose of predicting the extent of damage and probability of its occurrence.

The *Bayesian networks* obtained were subjected to quantitative and qualitative evaluation. The quantitative evaluation consisted of verifying the compliance of prediction of individual networks with the data from training and test sets. Thus, the degree of generalisation of the acquired knowledge was also checked. As a result of these analyses, four structures were induced which were subjected to further qualitative analysis. During the qualitative analysis, the completeness of the created structures was assessed in terms of actively connected nodes representing individual variables potentially influencing the occurrence of the damage process. The result was a narrowing of pre-selected networks from the quantitative analysis stage to one model. This model was created using the *Tabu-Search* algorithm using the BDla optimisation criterion.

The risk model created in the form of a *Bayesian network*, resulting from learning the DAG structure based on data, has several very significant advantages:

- It offers the possibility to both predict the probability of damage occurrence as well as diagnose the causes of its occurrence. Therefore, it can be used as a tool to estimate the risk of damage for a large number of building structures located in a mining area.
- It enables the interpretation of cause-and-effect relationships, which broaden the knowledge about the modelled phenomenon, especially in the case of the effect of variables regarding the quality of object maintenance.
- It allows to make inferences about any variable contained in the DAG structure representing the *Bayesian network*.
- It can be used in the absence of accurate information on the state of variables in individual nodes. Therefore, it enables the model to be used in the uncertainty area.
- It can be easily updated when new data is recorded resulting from observation of the actual course of a given process.

In addition, this type of methodology has a much broader implementation than just for the assessment of mining impacts. Examples include: *Structure Health Monitoring* or *Maintenance Management*.

The obtained results give rise to further research in this area, both in terms of analysing other types of building structures as well as testing other methods that allow learning *Bayesian network* structures from data.

References

1. Abdelkader EM, Zayed T, Marzouk M. A computerized hybrid Bayesian-based approach for modelling the deterioration of concrete bridge decks. *Structure and Infrastructure Engineering* 2019; 15: 1178-99, <https://doi.org/10.1080/15732479.2019.1619782>.
2. Carvalho AM. Scoring functions for learning Bayesian networks. *Inesc-id Tec. Rep* 2009; 12.
3. Chung H, Lee I-M, Jung J-H, Park J. Bayesian networks-based shield TBM risk management system: Methodology development and application. *KSCE Journal of Civil Engineering* 2019; 23: 452-65, <https://doi.org/10.1007/s12205-018-0912-y>.
4. Dahire S, Tahir F, Jiao Y, Liu Y. Bayesian Network inference for probabilistic strength estimation of aging pipeline systems. *International Journal of Pressure Vessels and Piping* 2018; 162: 30-9, <https://doi.org/10.1016/j.ijpvp.2018.01.004>.
5. Fang S-E, Tan J, Zhang X-H. Safety evaluation of truss structures using nested discrete Bayesian networks. *Structural Health Monitoring-an International Journal* 2020, <https://doi.org/10.1177/1475921720907888>.
6. Fedorowicz J, Słowik L. Interpretation of the behavior of a system building object-difficult subsoil in modern numerical modeling. *Proceedings of the 11th International Conference on New Trends in Statics and Dynamics of Buildings* 2013; 3.
7. Fereshtehnejad E, Banazadeh M, Shafieezadeh A. System reliability-based seismic collapse assessment of steel moment frames using incremental dynamic analysis and Bayesian probability network. *Engineering Structures* 2016; 118: 274-86, <https://doi.org/10.1016/j.engstruct.2016.03.057>.
8. Firek K. Proposal for classification of prefabricated panel building damage intensity rate in mining areas. *Archives of Mining Sciences* 2009.
9. Gogoshin G, Boerwinkle E, Rodin AS. New algorithm and software (BNOmics) for inferring and visualizing bayesian networks from heterogeneous big biological and genetic data. *Journal of Computational Biology* 2017; 24: 340-56, <https://doi.org/10.1089/cmb.2016.0100>.
10. Golewski GL. Measurement of fracture mechanics parameters of concrete containing fly ash thanks to use of Digital Image Correlation (DIC) method. *Measurement* 2019; 135: 96-105, <https://doi.org/10.1016/j.measurement.2018.11.032>.
11. Heckerman D, Geiger D, Chickering DM. Learning Bayesian networks: The combination of knowledge and statistical data. *Machine learning* 1995; 20: 197-243, <https://doi.org/10.1007/BF00994016>.
12. Jin Y, Zhang J, Sun L. Safety Risk Assessment of Prefabricated Building Construction Based on Bayesian Network. *IOP Conference Series: Earth and Environmental Science* 2019; 371: 32052, <https://doi.org/10.1088/1755-1315/371/3/032052>.
13. Ravikumar K, Rajiv Kannan A. Survey of Spatial Datamining Methods for Natural Disaster Management. *Middle-East Journal of Scientific Research* 2017; 25: 217-27.
14. Kembłowski MW, Grzyl B, Kristowski A, Siemaszko A. Risk Modelling with Bayesian Networks - Case Study: Construction of Tunnel under the Dead Vistula River in Gdansk. *Procedia Engineering* 2017; 196: 585-91, <https://doi.org/10.1016/j.proeng.2017.08.046>.
15. Knyziak P, Kanoniczak M. Difficulties in Operation of Elevations in Large-Panel Buildings. *IOP Conference Series: Materials Science and Engineering* 2019; 661: 12059, <https://doi.org/10.1088/1757-899X/661/1/012059>.
16. Lee S, Choi M, Lee H-S, Park M. Bayesian Network-based Seismic Damage Estimation for Power and Potable Water Supply Systems. *Reliability Engineering & System Safety* 2020: 106796, <https://doi.org/10.1016/j.res.2020.106796>.
17. Marsili F, Croce P, Klawonn F, Vignoli A, Boschi S, Landi F. A Bayesian network for the definition of probability models for masonry mechanical parameters. *14th International Probabilistic Workshop*. 2017: 253-68, https://doi.org/10.1007/978-3-319-47886-9_18.
18. Nagarajan R, Scutari M, Lèbre S. Bayesian networks in r. *Springer* 2013; 122: 125-7, <https://doi.org/10.1007/978-1-4614-6446-4>.
19. Niedojadło Z, Gruszczyński W. The impact of the estimation of the parameters values on the accuracy of predicting the impacts of mining exploitation. *Archives of Mining Sciences* 2015, <https://doi.org/10.1515/amsc-2015-0012>.
20. Nielsen JS, Sørensen JD. Computational framework for risk-based planning of inspections, maintenance, and condition monitoring using discrete Bayesian networks. *Structure and Infrastructure Engineering* 2018; 14: 1082-94, <https://doi.org/10.1080/15732479.2017.1387155>.
21. Nielsen JS, Tcherniak D, Ulriksen MD. A case study on risk-based maintenance of wind turbine blades with structural health monitoring. *Structure and Infrastructure Engineering* 2020: 1-17, <https://doi.org/10.1080/15732479.2020.1743326>.
22. Niewiadomski P, Hoła J. Failure process of compressed self-compacting concrete modified with nanoparticles assessed by acoustic emission method. *Automation in Construction* 2020; 112: 103111, <https://doi.org/10.1016/j.autcon.2020.103111>.
23. Pachla F, Tataru T. Dynamic Resistance of Residential Masonry Building with Structural Irregularities. *Seismic Behaviour and Design of Irregular and Complex Civil Structures III* 2020: 335-47, https://doi.org/10.1007/978-3-030-33532-8_26.
24. Pearl J. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Elsevier, 2014.
25. Rusek J. Application of Support Vector Machine in the analysis of the technical state of development in the LGOM mining area. *Eksploracja i Niezawodność – Maintenance and Reliability* 2017; 19 (1): 54–61, <http://dx.doi.org/10.17531/ein.2017.1.8>.
26. Rusek J, Firek K. Assessment of technical condition of prefabricated large-block building structures located in mining area using the Naive Bayes classifier. *International Multidisciplinary Scientific GeoConference Surveying Geology and Mining Ecology Management SGEM* 2016; 2.
27. Rusek J. The Point Nuisance Method as a Decision-Support System Based on Bayesian Inference Approach. *Archives of Mining Sciences* 2020; 65: 117-27.
28. Sadowski L. Non-destructive investigation of corrosion current density in steel reinforced concrete by artificial neural networks. *Archives of Civil and Mechanical Engineering* 2013; 13: 104-11, <https://doi.org/10.1016/j.acme.2012.10.007>.
29. Sadowski Ł, Hoła J, Czarniecki S. Non-destructive neural identification of the bond between concrete layers in existing elements. *Construction and Building Materials* 2016; 127: 49-58, <https://doi.org/10.1016/j.conbuildmat.2016.09.146>.
30. Sari DP, Rosadi D, Effendie AR, others. Application of Bayesian network model in determining the risk of building damage caused by earthquakes. *2018 International Conference on Information and Communications Technology (ICOIACT)* 2018: 131-5, <https://doi.org/10.1109/ICOIACT.2018.8350776>.
31. Schabowicz K, Gorzelańczyk T. A nondestructive methodology for the testing of fibre cement boards by means of a non-contact ultrasound scanner. *Construction and Building Materials* 2016; 102: 200-7, <https://doi.org/10.1016/j.conbuildmat.2015.10.170>.
32. Scutari M. Dirichlet Bayesian network scores and the maximum relative entropy principle. *Behaviormetrika* 2018; 45: 337-62, <https://doi.org/10.1007/s41237-018-0048-x>.

33. Scutari M. Learning Bayesian networks with the bnlearn R package. <https://doi.org/10.18637/jss.v035.i03>.
34. Scutari M, Graafland CE, Gutiérrez JM. Who learns better Bayesian network structures: Accuracy and speed of structure learning algorithms. *International Journal of Approximate Reasoning* 2019; 115: 235-53, <https://doi.org/10.1016/j.ijar.2019.10.003>.
35. Scutari M, Vitolo C, Tucker A. Learning Bayesian networks from big data with greedy search: computational complexity and efficient implementation. *Statistics and Computing* 2019; 1-14, <https://doi.org/10.1007/s11222-019-09857-1>.
36. Siemaszko A, Jakubczyk-Gańczyńska A, Jankowski R. The Idea of Using Bayesian Networks in Forecasting Impact of Traffic-Induced Vibrations Transmitted through the Ground on Residential Buildings. *Geosciences* 2019; 9: 339, <https://doi.org/10.3390/geosciences9080339>.
37. Sperotto A, Molina J-L, Torresan S, Critto A, Marcomini A. Reviewing Bayesian Networks potentials for climate change impacts assessment and management: A multi-risk perspective. *Journal of Environmental Management* 2017; 202: 320-31, <https://doi.org/10.1016/j.jenvman.2017.07.044>.
38. Tajduś K, Misa R, Sroka A. Analysis of the surface horizontal displacement changes due to longwall panel advance. *International Journal of Rock Mechanics and Mining Sciences* 2018; <https://doi.org/10.1016/j.ijrmms.2018.02.005>.
39. Tran T-B, Bastidas-Arteaga E, Aoues Y et al. Reliability assessment and updating of notched timber components subjected to environmental and mechanical loading. *Engineering Structures* 2018; 166: 107-16, <https://doi.org/10.1016/j.engstruct.2018.03.053>.
40. Wang F, Li H, Dong C, Ding L. Knowledge representation using non-parametric Bayesian networks for tunneling risk analysis. *Reliability Engineering & System Safety* 2019; 191: 106529, <https://doi.org/10.1016/j.res.2019.106529>.
41. Wodyński A, Lasocki S. Assessment of mining tremor influence on the technical wear of building. *Acta Geodynamica et Geomaterialia*. Ser. A and Ser. B 2004; 50: 187-94.
42. Wu L, Wang J, Zhou J et al. Multi-scale geotechnical features of dredger fills and subsidence risk evaluation in reclaimed land using BN. *Marine Georesources & Geotechnology* 2019: 1-23, <https://doi.org/10.1080/1064119X.2019.1644406>.
43. Yazdani A, Shahidzadeh M-S, Takada T. Bayesian networks for disaggregation of structural reliability. *Structural Safety* 2020; 82: 101892, <https://doi.org/10.1016/j.strusafe.2019.101892>.
44. Yuan M, Liu Y, Yan D, Liu Y. Probabilistic fatigue life prediction for concrete bridges using Bayesian inference. *Advances in Structural Engineering* 2019; 22: 765-78, <https://doi.org/10.1177/1369433218799545>.

Janusz Rusek, dr hab. inż., profesor AGH

Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie
al. Mickiewicza 30, 30-059 Kraków, Polska
tel. +48 12 617 22 80, e-mail: rusek@agh.edu.pl

Karol Firek, dr hab. inż., profesor AGH

Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie
al. Mickiewicza 30, 30-059 Kraków, Polska
tel. +48 12 617 51 56, e-mail: kfirek@agh.edu.pl

Leszek Słowik, dr inż.

Instytut Techniki Budowlanej
ul. Filtrowa 1, 00-611 Warszawa, Polska
tel. +48 32 730 29 64, e-mail: l.slowik@itb.pl

Wyodrębnianie struktury sieci Bayesowskiej z danych w prognozowaniu uszkodzeń żelbetowych budynków prefabrykowanych na terenach górniczych

Słowa kluczowe: sieci Bayesowskie, uszkodzenia, ryzyko, prefabrykowane budynki żelbetowe

Streszczenie: W artykule zaprezentowano wyniki badań dotyczących budowy modelu do oceny ryzyka powstawania uszkodzeń budynków usytuowanych na terenach górniczych. Podstawą do badań była baza danych nt. konstrukcji, stanu technicznego oraz wpływów górniczych dla 129 żelbetowych prefabrykowanych budynków wznoszonych w uprzemysłowionym systemie wielkoblokowym zlokalizowanych na terenie górniczym Legnicko-Głódzkiego Okręgu Miedziowego (LGOM). Do analiz zastosowano metodykę sieci przekonania Bayesa (BBN – *Belief Bayesian Networks*). Stosując podejście *score-based Bayesian structure learning* (*Hill-Climbing* oraz *Tabu-Search*) oraz wyselekcjonowane kryteria optymalizacyjne, wyłoniono 16 struktur sieci *Bayesowskich*. Wszystkie modele poddano ocenie ilościowej i jakościowej, weryfikując ich własności w kontekście trafności predykcji, generalizacji nabytej wiedzy oraz zależności przyczynowo-skutkowych. Pozwoliło to na wyselekcjonowanie najlepszej struktury sieci wraz z odpowiadającym kryterium optymalizacyjnym. Analiza wyników wykazała, że metoda *Tabu-Search* przy przyjęciu kryterium optymalizacyjnego w postaci *Locally Averaged Bayesian Dirichlet score* (BDla), prowadzi do uzyskania modelu o najlepszych własnościach spośród wszystkich wyłonionych modeli. Uzyskane rezultaty uzasadniają przyjęcie metodyki BBN, jako efektywnej w kontekście oceny zakresu uszkodzeń budynków na terenach górniczych.

1. Wprowadzenie

Uszkodzenia istniejących budynków na terenach górniczych mogą być spowodowane przez szereg czynników budowlanych i środowiskowych, do których zalicza się również wpływy górnicze. Objawiają się one na powierzchni terenu w postaci ciągłych lub nieciągłych deformacji terenu [6,38] oraz wstrząsów górniczych [23,41]. Uszkodzenia te mogą dotyczyć zarówno elementów konstrukcyjnych, jak i drugorzędnych [8]. Taka sytuacja ma miejsce również w przypadku budynków o prefabrykowanej konstrukcji żelbetowej, wzniesionych w systemach budownictwa uprzemysłowionego [15], w których mieszka około 1/3 mieszkańców Polski (ok. 12 mln ludzi). Dodatkowo są to z reguły wielkogabarytowe budynki wielorodzinne lub użyteczności publicznej, gdzie zakłócenia komfortu użytkownika z tytułu uszkodzeń są bardzo istotne.

Najczęściej zainicjowanie procesu powstawania uszkodzeń zależy nie od jednego lecz od wielu czynników jednocześnie. Powoduje to, że problem dotyczący oceny przyczyn zaistniałego stanu uszkodzeń jest złożony i trudny do opisu z analitycznego punktu widzenia. W kontekście wpływów górniczych, sytuację komplikuje dodatkowo niepewność dotycząca prognoz deformacji na powierzchni terenu oraz wstrząsów [19].

W praktyce, niejednokrotnie konieczne jest przeprowadzenie takiej oceny dla dużej liczby budynków stanowiących zabudowę danego terenu górniczego. Fakt ten dyskwalifikuje w ocenie podejście numeryczne MES, które w takim przypadku jest nieefektywne. Wydaje się, że jedyną możliwością zmierzenia się z problemem dają modele statystyczne.

Do grona popularnych i bardzo efektywnych metod pozwalających na detekcję uszkodzeń w konstrukcjach betonowych można zaliczyć również grupę tzw. metod nieniszczących. Tutaj wymienić należy metody akustyczne (Acoustic Emission - AE) [22], obrazowe (Digital Image Correlation - DIC) [10] lub ultradźwiękowe [31]. Oprócz samej detekcji pęknięć, metody nieniszczące umożliwiają bardziej szczegółową analizę degradacji elementów żelbetowych, która obejmuje również wyznaczenie korozji prętów zbrojeniowych [28]. Ma to szczególne znaczenie w przypadku oceny uszkodzeń w prefabrykowanych budynkach wielkopłytowych, gdzie podstawową kwestią jest rozstrzygnięcie stanu technicznego połączeń oraz ich zbrojenia. Dodatkowo, w ostatnich latach coraz częściej wykorzystywane są również alternatywne podejścia bazujące na wykorzystaniu metod z obszaru uczenia maszynowego np. [28,29].

Biorąc pod uwagę specyfikę analizowanego problemu, szczególnie w kontekście predykcji uszkodzeń dla dużej liczby budynków, wyodrębniono do dalszych badań metody uczenia maszynowego umożliwiające przedstawienie badanego zjawiska w notacji probabilistycznej.

W ramach dotychczasowych badań wykazano efektywność stosowania narzędzi z obszaru *Machine Learning* (ML) [25,27]. Jednak wszystkie z testowanych metod zakładały wzajemną niezależność potencjalnych przyczyn mogących mieć wpływ na proces powstawania uszkodzeń. Z tego powodu zdecydowano o wykorzystaniu metodyki tzw. *sieci Bayesowskich* (*Bayesian Network* - BN) lub *sieci przekonań Bayesa* (*Bayesian Belief Network* - BBN), która zgodnie z [24], dopuszcza możliwość uwzględnienia niezależności warunkowej. To z kolei pozwala na budowę modelu, w którym relacje między poszczególnymi czynnikami, są bliższe rzeczywistości fizycznej związanej z przebiegiem procesu powstawania uszkodzeń. W efekcie ustalona struktura BN może być utożsamiona z modelem ryzyka powstania uszkodzeń.

Dodatkową własnością sieci *Bayesowskich*, która stanowi bardzo istotną zaletę w odniesieniu do zagadnienia powstawania uszkodzeń w budynkach poddanych m.in. wpływom górniczym, jest możliwość wykorzystania utworzonego modelu nie tylko w kontekście predykcji, ale i diagnozy uszkodzeń. To z kolei pozwala na dużo szerszą implementację takiego systemu w praktyce.

Niestety istotnym problemem w przypadku dużej liczby zmiennych opisujących modelowany proces, jest często brak znajomości struktury sieci *Bayesowskiej*. Istnieją jednak metody, które pozwalają na wyodrębnienie struktury sieci z danych. Zgodnie z [18] podzielono je na 3 grupy: *constraint based structure learning*, *score-based structure learning* oraz *podejście hybrydowe*. W pracy zastosowano podejście *score-based*, dla którego uzasadnienie podano w rozdziale 4.

W celu wyłonienia optymalnej struktury sieci *Bayesowskiej* przetestowano dwa algorytmy: *Hill-Climbing* (HC) oraz *Tabu-Search* (TS), które należą do grupy metod *score-based structure learning* [18].

Podstawą do analiz była baza danych o uszkodzeniach grupy 129 żelbetowych, prefabrykowanych budynków wzniesionych w systemie uprzemysłowionym WBL zlokalizowanych na terenie górniczym LGOM (rys. 1). Informacje o stanie technicznym, w tym

o uszkodzeniach były zbierane w trakcie inwentaryzacji dokonywanej „in-situ” oraz na podstawie zgłoszeń. Dane te posłużyły do wyłonienia struktur sieci *Bayesowskich* oraz późniejszego ich testowania. Weryfikowano jakość utworzonych modeli w kontekście poprawnej klasyfikacji oraz własności generalizujących. Dodatkowo, dysponując jawną reprezentacją powiązań między węzłami sieci dokonywano weryfikacji pod kątem zgodności modelu z obserwowaną, rzeczywistą naturą modelowanego procesu. Była to ocena subiektywna, dokonywana na podstawie wiedzy eksperckiej.

Na podstawie uzyskanych wyników wyłoniono strukturę sieci *Bayesowskiej*, która może reprezentować złożony model ryzyka powstania uszkodzeń. Podstawą oceny była tutaj przede wszystkim zgodność klasyfikacji modelu z danymi uczącymi i testowymi.

2. Przegląd literaturowy

Metodyka sieci *Bayesowskich* (BN) lub inaczej *sieci przekonań Bayesa* (BBN) znajduje aktualnie szerokie zastosowanie w analizie ryzyka [3], bezpieczeństwa [12], niezawodności [43] oraz przewidywania zakresu uszkodzeń [16]. W ostatnich latach coraz częściej spotyka się implementację tego typu metodyki w obszarze inżynierii lądowej. Dominuje tutaj jednak podejście ekspertowe, gdzie struktura sieci *Bayesa* nadawana jest w sposób odgórny, a uczeniu podlegają jedynie parametry modelu. Wyszczególniając obszary zbieżne do podejmowanej w pracy tematyki można wymienić zagadnienia związane z analizą ryzyka powstania uszkodzeń w obiektach budowlanych. Przedmiotem tych analiz są przeważnie budowlane obiekty liniowe takie jak: rurociągi, obiekty mostowe [1] lub tunele [40]. Chociaż można spotkać również przykłady dotyczące oceny ryzyka powstania uszkodzeń w budynkach [30]. Zastosowanie sieci *Bayesowskich* występuje również w szerszym ujęciu dotyczącym analiz niezawodnościowych [39], w tym również oceny bezpieczeństwa obiektów budowlanych [5]. Metodyka ta wykorzystywana jest także do budowy systemów diagnostycznych w obszarze zarządzania utrzymaniem obiektów budowlanych [20], jak i do oceny ich stanu technicznego [21]. W węższym zakresie występują również, jako systemy pozwalające na ocenę parametrów wytrzymałościowych istniejących konstrukcji, zarówno statycznych [4], jak i zmęczeniowych [44].

W kwestii różnorodności typów konstrukcji budowlanych, metodyka sieci *Bayesowskich* stosowana jest również w odniesieniu do budynków murowanych [17], jak i stalowych [7]. Natomiast w szerszym kontekście, bardzo często wykorzystywana jest jako narzędzie do prognozowania wpływów środowiskowych o charakterze losowym takich jak: powódzie, trzęsienia ziemi, tsunami [13], zmiany klimatyczne [37] lub osiadania terenu [42]. Wyniesione z takich analiz informacje, są z kolei są wykorzystywane zarówno na etapie projektowania, jaki oceny zagrożenia istniejących konstrukcji budowlanych.

Należy podkreślić, że w ostatnim czasie, metodyka sieci *Bayesowskich* znajduje zastosowanie w obszarze budownictwa również na terenie Polski. Przykładem może być tutaj wykorzystane jej do oceny wpływu drgań komunikacyjnych na zabudowę powierzchni [36] oraz analiza ryzyka w ramach inwestycji budowlanej dotyczącej budowy tunel pod Martwa Wisłą w Gdańsku [14].

Jednak pomimo licznych zastosowań metodyki opartej na sieciach *Bayesowskich* w obszarze inżynierii lądowej, stosowanie i rozwijanie metod wykrywania struktury z danych jest nadal domeną nauk tj. medycyna, biologia, genetyka np.[9]. Z tej przyczyny, prezentowane w pracy badania, stanowią poszerzenie metodyki stosowanej dotychczas w obszarze inżynierii lądowej, które może pozwolić na analizę bardziej złożonych problemów inżynierskich z obszaru analizy ryzyka uszkodzeń lub niezawodności konstrukcji. Podjęte zagadnienie, dotyczące wykrywania struktury sieci *Bayesowskiej* z danych do budowy modelu ryzyka

powstania uszkodzeń w budynkach poddanych wpływow eksploatacji górniczej stanowi zdaniem autorów, interdyscyplinarne i nowatorskie ujęcie tematu.



Rys. 1. Przykłady budynków zrealizowanych w technologii uprzemysłowionej WBL
(own source)

Fig. 1. Examples of buildings constructed using WBL industrialised technology

3. Charakterystyka bazy danych

W bazie danych, która była podstawą badań, zebrano szczegółowe informacje dotyczące grupy 129 żelbetowych prefabrykowanych wielkogabarytowych budynków wielorodzinnych i użyteczności publicznej, wzniesionych uprzemysłowionym systemem wielkoblokowym WBL np. [26]. W każdym przypadku, w miejscu lokalizacji budynku, określone zostały wskaźniki opisujące intensywność wpływów górniczych na czas dokonywanej inwentaryzacji (**MC**, **AH** i **ASG** – por. tab. 1). Zmienna **ASG** została ustalona na podstawie oryginalnego wskaźnika wstrząsów górniczych a_{sg} [41] i wyraża wpływ wielokrotnego oddziaływania wstrząsów górniczych na zużycie techniczne budynków. Bazę uzupełniają informacje dotyczące cech konstrukcyjno-materiałowych, historii i jakości utrzymania, remontów oraz odnotowanych uszkodzeń. W przypadku uszkodzeń wykorzystano autorski wskaźnik uszkodzeń w_u odniesiony zarówno do elementów konstrukcyjnych, jak i wykończeniowych [8].

W trakcie wstępnej analizy bazy danych, dokonano kategoryzacji wszystkich zmiennych pod kątem dalszego ich wykorzystania w trakcie uczenia sieci *Bayesowskich*. Polegało to na przypisaniu etykiet dla poszczególnych kategorii. Wykaz wszystkich zmiennych wraz z przypisanym zakresem etykiet podano w tabeli 1. W późniejszym etapie, każda ze zmiennych będzie odzwierciedlać poszczególne węzły w sieci *Bayesowskiej*.

4. Metodyka badań

Sieci Bayesowskie (BN – *Bayesian Network*), zwane także *Sieciami przekonań Bayesa* (BBN – *Bayesian Belief Networks*) można przedstawić w postaci *acyklicznego grafu skierowanego* (DAG – *Directed Acyclic Graph*) [18]. W strukturze grafu (G) zakodowana jest informacja o wzajemnych zależnościach między poszczególnymi zmiennymi $\mathbf{X} = \{X_1, \dots, X_N\}$, którą reprezentują krawędzie grafu (E) oraz węzły (V). W ogólnym ujęciu, BBN reprezentuje łączny rozkład prawdopodobieństwa nad zbiorem zmiennych losowych \mathbf{X} , który można przedstawić jako [34]:

$$P(\mathbf{X}|G, \boldsymbol{\theta}) = \prod_{i=1}^N P(X_i | \Pi_{X_i}, \theta_{X_i}) \quad (1)$$

gdzie:

$G = G(\mathbf{X}, \mathbf{E}, \mathbf{V})$ – struktura acyklicznego grafu skierowanego

$\mathbf{X} = \{X_1, \dots, X_N\}$ – zbiór wszystkich zmiennych występujących w węzłach grafu

$X_i = \{x_i^{(1)}, \dots, x_i^{(k_i)}\}$ – stany j -tej zmiennej

\mathbf{E} – zbiór wszystkich krawędzi

\mathbf{V} – zbiór wszystkich węzłów

$\Pi_{X_i} = \{x_i^{(q_1)}, \dots, x_i^{(q_i)}\}$ – zbiór rodziców, czyli wszystkich węzłów grafu warunkujących stan węzła X_i

$\boldsymbol{\theta} = \{\theta_{X_1}, \dots, \theta_{X_N}\}$ – zbiór wszystkich parametrów relacji warunkowych między poszczególnymi węzłami X_i , a zbiorem ich rodziców Π_{X_i}

W przypadku zmiennych dyskretnych, parametry modelu $\theta_{X_j} = \{\theta_{ijk}\}$ są reprezentowane w postaci wielonomialnej tablicy prawdopodobieństwa warunkowego (CPT - *Conditional Probability Table*), której elementy wyraża się jako [11]:

$$\theta_{ijk} = P(X_j = x_j^{(i)} | \Pi_{X_j} = \pi_j^{(k)}) \quad (2)$$

Zgodnie z zależnością (1) łączny rozkład prawdopodobieństwa $P(\mathbf{X}|G, \boldsymbol{\theta})$ podlega dekompozycji w oparciu o warunkowe rozkłady lokalne $P(X_i | \Pi_{X_i}, \boldsymbol{\theta}_{X_i})$, opisane nad każdą zmienną losową X_i względem odpowiadającego jej zbioru zmiennych warunkujących, tzw. rodziców Π_{X_i} . Takie sformułowanie jest możliwe dzięki pojęciu warunkowej niezależności wprowadzonej przez *Perla* w pracy [24]. Pozwala to dodatkowo na efektywną analizę zmian wartości przyjętego kryterium w trakcie wyszukiwania optymalnej struktury sieci.

W opisywanym problemie, dotyczącym budowy modelu ryzyka powstawania uszkodzeń, niewiadomymi są zarówno struktura BBN, jak również jej parametry. W ramach dotychczasowych badań potwierdzono istotny, lecz subtelny wpływ poszczególnych czynników na proces powstawania uszkodzeń. Dlatego, na etapie selekcji metody uczenia struktury BBN z danych, kierowano się tym, żeby w ostatecznym modelu występowała jak największa liczba zmiennych potencjalnie wpływających na proces inicjowania uszkodzeń. Na podstawie literatury [18] odrzucono metody *constraint-based structure learning* oraz *podejście hybrydowe*. Stwierdzono bowiem, że w ramach tych metod uczenia większa waga przypisywana jest indywidualnym połączeniom między węzłami sieci niż globalnej odpowiedzi modelu. Tym samym ukierunkowano się na metody *score-based structure learning*. Finalnie zastosowano dwie metody: *Hill-Climbing* (HC) i *Tabu-Search* (TS).

Podejście optymalizacyjne HC zaliczane jest do grona metod *greedy-search* [35]. Polega ono na przeszukiwaniu najbliższego otoczenia wokół aktualnego punktu w przestrzeni, któremu odpowiada dana wartość przyjętego kryterium optymalizacyjnego. Następnie, z otoczenia wybierany jest punkt w przestrzeni poszukiwań, którego wartość kryterialna jest wyższa niż w kroku poprzednim. Obszar poszukiwań stanowi przestrzeń struktur acyklicznych grafów skierowanych (DAG), a postęp algorytmu następuje poprzez dodawanie, odejmowanie lub zamianę powiązań między węzłami.

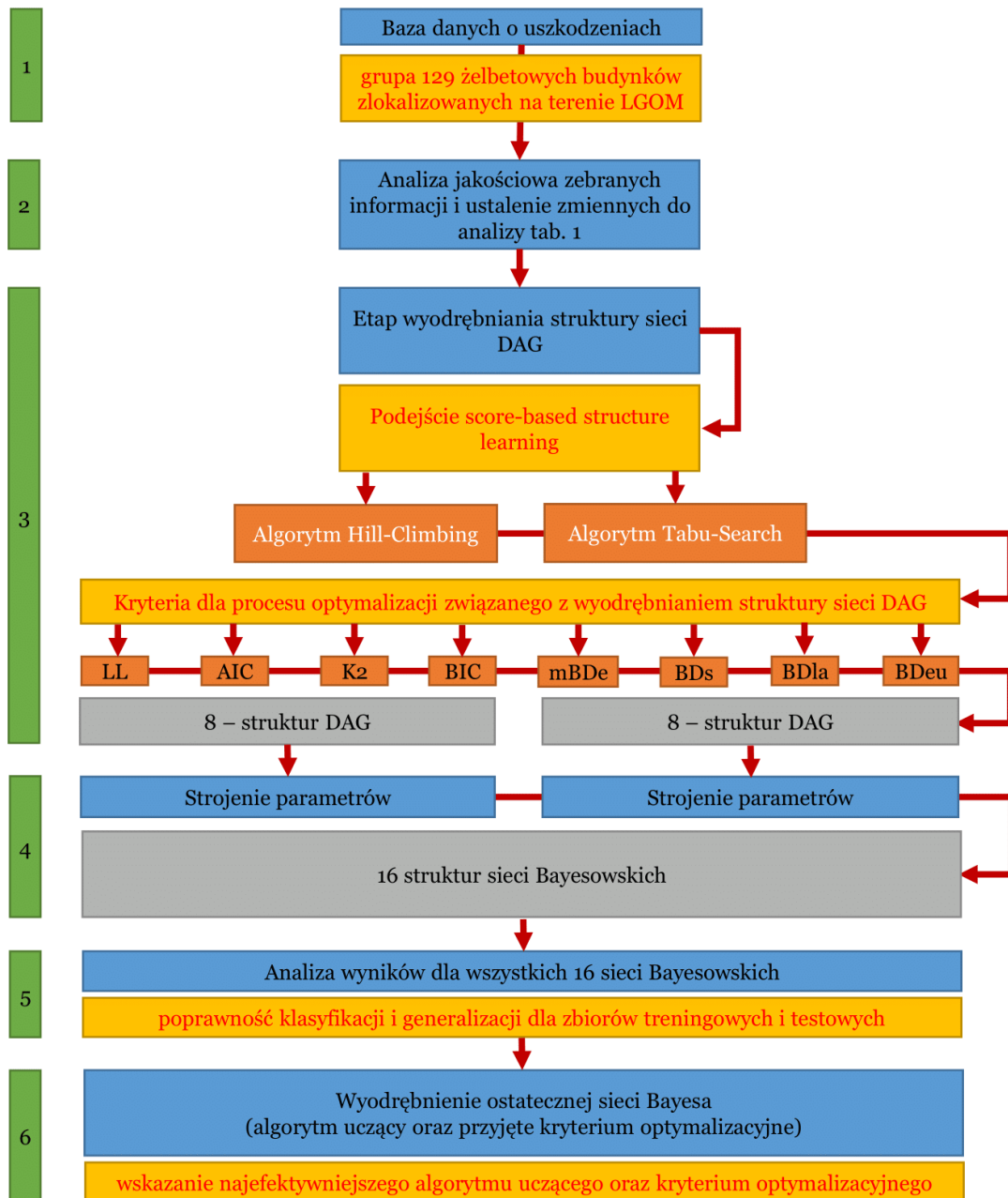
Z kolei podejście TS stanowi modyfikację algorytmu HC. Modyfikacja ta polega na przechowywaniu określonej liczby zakazanych ścieżek w przestrzeni poszukiwań, które zostały już przebyte w ramach poprzednich iteracji, a tym samym ogranicza się ślepe błądzenie oraz utknięcie w optimum lokalnym [33].

Tabela 1. Wykaz zmiennych w bazie danych wraz z przypisanym zakresem etykiet
 Table 1. List of variables in the database with the assigned label range

WYKAZ ZMIENNYCH		
OZN	OPIS	TYP ZMIENNEJ/ ETYKIETY ZMIENNYCH
DANE O ODDZIAŁYWANIACH GÓRNICZYCH W MIEJSCU LOKALIZACJI DANEGO BUDYNKU		
MC	kategoria terenu górniczego	kategor. / 3 kategorie
AH	maksymalna składowa pozioma przyspieszenia drgań gruntu	kategor. / 4 kategorie
ASG	wskaźnik oceny wpływu wstrząsów górniczych [41]	kategor. / 4 kategorie
WSKAŹNIKI USZKODZEŃ ELEMENTÓW KONSTRUKCYJNYCH I WYKOŃCZENIOWYCH		
W _{u2}	wskaźnik uszkodzeń ścian piwnicznych lub fundamentowych	kategor. / 4 kategorie
W _{u3}	wskaźnik uszkodzeń ścian nośnych nadziemna wewn. i zewn. (w tym nadproży i murów podokiennych)	kategor. / 4 kategorie
W _{u7}	wskaźnik uszkodzeń stropów wyżej, stropodachu (przekrycia)	kategor. / 4 kategorie
W _{u11}	wskaźnik uszkodzeń ścian działowych	kategor. / 2 kategorie
W _{u12}	wskaźnik uszkodzeń tynków wewnętrznych i okładzin ściennych	kategor. / 3 kategorie
W _{u13}	wskaźnik uszkodzeń podłogi (warstw podłogowych)	kategor. / 3 kategorie
W _{u17}	wskaźnik uszkodzeń elewacji (warstw elewacyjnych)	kategor. / 4 kategorie
W _{u19}	wskaźnik uszkodzeń pokrycia dachowego	kategor. / 4 kategorie
DANE O UTRZYMANIU I REMONTACH		
CR	remonty bieżące	kategor. / 4 kategorie
FR	remonty elewacji	kategor. / 5 kategorie
RR	remonty dachu	kategor. / 6 kategorie
IR	remonty wnętrza	kategor. / 7 kategorie
DANE O KONSTRUKCJI I GEOMETRII BUDYNKU		
LEN	długość budynku (dłuższy z wymiarów)	kategor. / 6 kategorie
NoST	liczba kondygnacji	kategor. / 6 kategorie
NoSE	liczba segmentów	kategor. / 8 kategorie
LoSE	długość segmentów	kategor. / 6 kategorie
DIL	dylatacja (szerokość)	kategor. / 3 kategorie
SHA	kształt budynku	kategor. / 4 kategorie
ToF	rodzaj fundamentu	kategor. / 5 kategorie
SW	ściany osłonowe	kategor. / 4 kategorie
BC	strop nad piwnicą	kategor. / 4 kategorie
DP	zabezpieczenia projektowe	kategor. / 3 kategorie
EP	zabezpieczenia istniejące	kategor. / 5 kategorie
DANE DOTYCZĄCE TRWAŁOŚCI		
AGE	wiek budynku	kategor. / 4 kategorie
DUR	trwałość	kategor. / 5 kategorie
TS	stan techniczny	kategor. / 4 kategorie

W niniejszej pracy, wykorzystując pakiet *bnlearn*, podjęto analizę porównawczą obydwu podejść. Dodatkowo przetestowano ich działanie w sytuacji przyjęcia różnych kryteriów optymalizacyjnych. W tym celu przyjęto kryteria należące do dwóch grup: *Information-Theoretic scores* (IT) [2] and *Bayesian Dirichlet scores* (BD) [32]. Do pierwszej z nich należą: *Log-Likelihood* (LL), *Akaike Information Criterion* (AIC) and *Bayesian Information Criterion* (BIC). Natomiast z grupy *Bayesian Dirichlet scores* (BD) zastosowano wg. [33]: *Bayesian Dirichlet equivalent uniform* (BDeu), *modified Bayesian Dirichlet equivalent score* (mBDe), *Bayesian Dirichlet sparse score* (BDs), *locally averaged Bayesian*

Dirichlet score (BDIa) and the *K2* score. Schemat chronologiczny przeprowadzonych badań przedstawiono na rys. 2.



Rys. 2. Chronologiczny schemat przeprowadzonych analiz
 Fig. 2. Chronological diagram of the carried out analyses

5. Wyniki przeprowadzonych analiz

W badaniach zastosowano metodykę opisaną w rozdziale 3, w której założono dwa podejścia do budowy struktury *sieci przekonań Bayesa* (BBN) z danych. Obydwa podejścia należą do grupy metod *score-based structure learning* [18]. Jako pierwszy zastosowany algorytm *Hill-Climbing* (HC). Natomiast w drugim wykorzystano jego modyfikację, a mianowicie algorytm *Tabu-Search* (TS).

Zestaw danych o budynkach, opisany w rozdziale 2, został podzielony na zbiór treningowy i testowy. Kluczowe było tutaj takie rozdzielanie zbiorów, aby zachować w nich równomierny podział wartości skategoryzowanych zmiennych, ujętych w tab.1 tj.: w_{u2} , w_{u3} , w_{u7} , w_{u11} , w_{u12} , w_{u13} , w_{u17} i w_{u19} . Finalnie dokonano tego za pomocą metody *Stratified Sampling* (SS). Uzyskano tym samym zbiór treningowy o licznosci równej 105 wzorców (81,36% całkowitej liczby wzorców w bazie danych) oraz zbiór testowy, którego liczba wzorców była równa 24 (18,6%).

W przypadku wykorzystanego w analizach pakietu *bnlearn*, istnieje możliwość uwzględnienia wiedzy eksperta. Odbywa się to za pomocą wprowadzania listy reprezentującej powiązania zakazane (*Black List*) i wymuszone (*White List*). Są to ograniczenia typu mocnego (*strong constraint*), które nie podlegają modyfikacji w trakcie procesu uczenia.

W pracy zestawiono wyniki przeprowadzonych analiz, których celem było porównanie obydwu zastosowanych podejść. Ostatecznie wyłoniono 16 struktur BBN, które poddano ocenie. Wynikało to z przyjęcia dwóch algorytmów uczących i ośmiu kryteriów optymalizacyjnych (*Information-Theoretic scores* (IT) and *Bayesian Dirichlet scores* (BD) - wg. rozdziału 3).

Przed rozpoczęciem obliczeń, wprowadzono ograniczenia (*strong constraint*) powodujące wykluczenie z sieci tych powiązań, które są sprzeczne z logiką i nie są obserwowane w rzeczywistości. Łącznie utworzono zestaw 427 par między węzłami, w których zabroniono tworzenia się powiązań. Należy zaznaczyć, iż nie tworzą listy wymuszeń, pozostawiając w tej kwestii swobodę poszczególnym algorytmom uczącym.

W **pierwszym etapie** oceniono poszczególne struktury pod kątem ilościowym. Podstawą była tutaj analiza poprawności klasyfikacji oraz własności generalizacyjne wyłonionych struktur BBN. Zastosowano tutaj miarę względną w postaci procentowego udziału poprawnie sklasyfikowanych wzorców względem licznosci całego zbioru danych (odpowiednio treningowego i testowego). Wyniki te dla poszczególnych kombinacji wynikających z przyjętej metody uczenia i kryterium optymalizacyjnego zestawiono w tab. 2 ÷ 5. Pozwoliło to na ocenę poprawności klasyfikacji oraz własności generalizacyjnych utworzonych modeli. Podstawą było tutaj symulowanie odpowiedzi sieci w węzłach odpowiadających poszczególnym wskaźnikom uszkodzeń: w_{u2} , w_{u3} , w_{u7} , w_{u11} , w_{u12} , w_{u13} , w_{u17} i w_{u19} .

Tabela. 2. Poprawność klasyfikacji dla poszczególnych węzłów odpowiadających wskaźnikom uszkodzeń – zbiór treningowy – algorytm uczący HC

Table 2. Accuracy of classification for individual nodes corresponding to damage indices - training set - algorithm teaching HC

kryterium	Poziom dokładności ac^{HC-TR} [%]							
	w_{u2}	w_{u3}	w_{u7}	w_{u11}	w_{u12}	w_{u13}	w_{u17}	w_{u19}
LL	100,00	99,05	99,05	100,00	100,00	100,00	99,05	99,05
AIC	92,38	90,48	85,71	92,38	94,29	95,24	90,48	90,48
BIC	93,33	90,48	89,52	92,38	94,29	90,48	90,48	90,48
BDeu	89,52	91,43	86,67	92,38	94,29	94,29	91,43	93,33
BDs	92,38	93,33	89,52	97,14	97,14	99,05	94,29	92,38
mBDe	91,43	91,43	85,71	92,38	94,29	94,29	91,43	93,33
BDla	90,48	93,33	85,71	97,14	97,14	95,24	95,24	94,29
K2	93,33	90,48	90,48	93,33	91,43	94,29	95,24	90,48

Tabela. 3. Poprawność klasyfikacji dla poszczególnych węzłów odpowiadających wskaźnikom uszkodzeń – zbiór testowy – algorytm uczący HC

Table 3. Accuracy of classification for individual nodes corresponding to damage indices - test set - algorithm teaching HC

kryterium	Poziom dokładności ac^{HC-TS} [%]							
	wu_2	wu_3	wu_7	wu_{11}	wu_{12}	wu_{13}	wu_{17}	wu_{19}
LL	95,83	95,83	79,17	100,00	100,00	100,00	79,17	95,83
AIC	83,33	87,50	75,00	95,83	95,83	100,00	37,50	95,83
BIC	87,50	87,50	79,17	91,67	91,67	91,67	41,67	95,83
BDeu	83,33	87,50	79,17	95,83	95,83	100,00	62,50	100,00
BDs	91,67	91,67	79,17	100,00	100,00	100,00	75,00	100,00
mBDe	83,33	87,50	83,33	95,83	95,83	100,00	66,67	100,00
BDla	83,33	91,67	83,33	100,00	100,00	100,00	87,50	100,00
K2	95,83	83,33	75,00	95,83	100,00	91,67	87,50	95,83

Tabela. 4. Poprawność klasyfikacji dla poszczególnych węzłów odpowiadających wskaźnikom uszkodzeń – zbiór treningowy – algorytm uczący TS

Table 4. Accuracy of classification for individual nodes corresponding to damage indices - training set - algorithm teaching TS

kryterium	Poziom dokładności $ac^{TABU-TR}$ [%]							
	wu_2	wu_3	wu_7	wu_{11}	wu_{12}	wu_{13}	wu_{17}	wu_{19}
LL	100,00	99,05	100,00	100,00	100,00	100,00	100,00	100,00
AIC	92,38	90,48	85,71	92,38	94,29	96,19	90,48	90,48
BIC	93,33	90,48	89,52	92,38	94,29	90,48	90,48	90,48
BDeu	91,43	90,48	88,57	92,38	94,29	96,19	91,43	93,33
BDs	95,24	96,19	89,52	97,14	97,14	98,10	94,29	93,33
mBDe	91,43	89,52	86,67	93,33	94,29	94,29	90,48	93,33
BDla	91,43	93,33	86,67	97,14	97,14	95,24	95,24	93,33
K2	93,33	90,48	86,67	89,52	94,29	92,38	95,24	90,48

Tabela. 5. Poprawność klasyfikacji dla poszczególnych węzłów odpowiadających wskaźnikom uszkodzeń – zbiór testowy – algorytm uczący TS

Table 5. Accuracy of classification for individual nodes corresponding to damage indices - test set - algorithm teaching TS

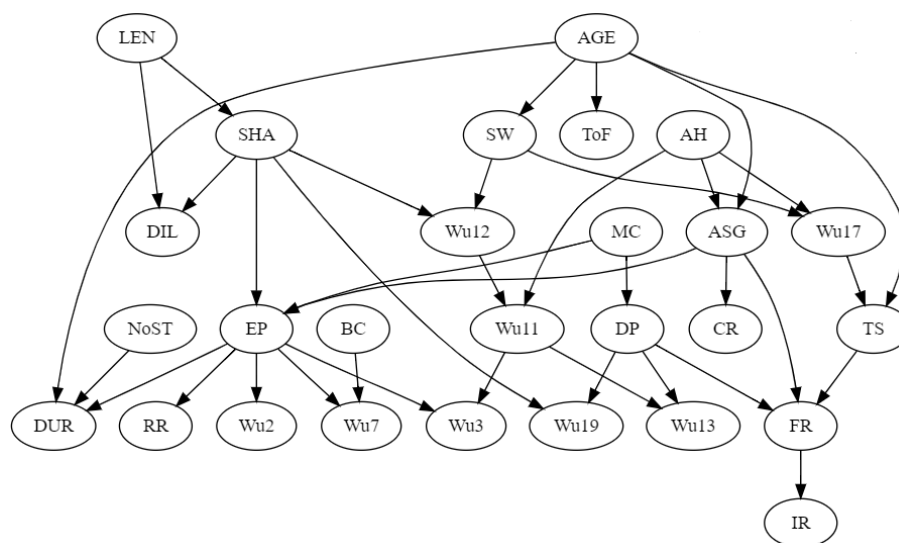
kryterium	Poziom dokładności $ac^{TABU-TS}$ [%]							
	wu_2	wu_3	wu_7	wu_{11}	wu_{12}	wu_{13}	wu_{17}	wu_{19}
LL	95,83	95,83	79,17	100,00	100,00	100,00	79,17	95,83
AIC	83,33	87,50	70,83	95,83	95,83	100,00	50,00	95,83
BIC	87,50	87,50	79,17	91,67	91,67	91,67	54,17	95,83
BDeu	87,50	87,50	87,50	95,83	95,83	100,00	66,67	100,00
BDs	91,67	91,67	83,33	100,00	100,00	100,00	66,67	100,00
mBDe	83,33	87,50	83,33	95,83	95,83	100,00	66,67	100,00
BDla	83,33	91,67	87,50	95,83	100,00	100,00	79,17	100,00
K2	95,83	91,67	79,17	91,67	91,67	91,67	87,50	95,83

W **drugim etapie** weryfikowano uzyskane struktury BBN pod kątem jakościowym. Było to podyktowane względami praktycznymi wynikającymi z możliwości późniejszego stosowania utworzonego modelu do oceny ryzyka w budownictwie. Tutaj skupiono się na wykrytych powiązaniach przyczynowo-skutkowych. Ocena była podejmowana w sposób ekspercki, wynikający z doświadczenia autorów w zakresie problematyki uszkodzeń budynków na terenach górniczych.

Na podstawie wyników analizy ilościowej, którą reprezentuje poziom poprawnie sklasyfikowanych wzorców dla zbiorów treningowych i testowych opisanych miarą: ac^{HC-TR} , ac^{HC-TS} , $ac^{TABU-TR}$, $ac^{TABU-TS}$ (tab. 2 ÷ 5), stwierdzono, że:

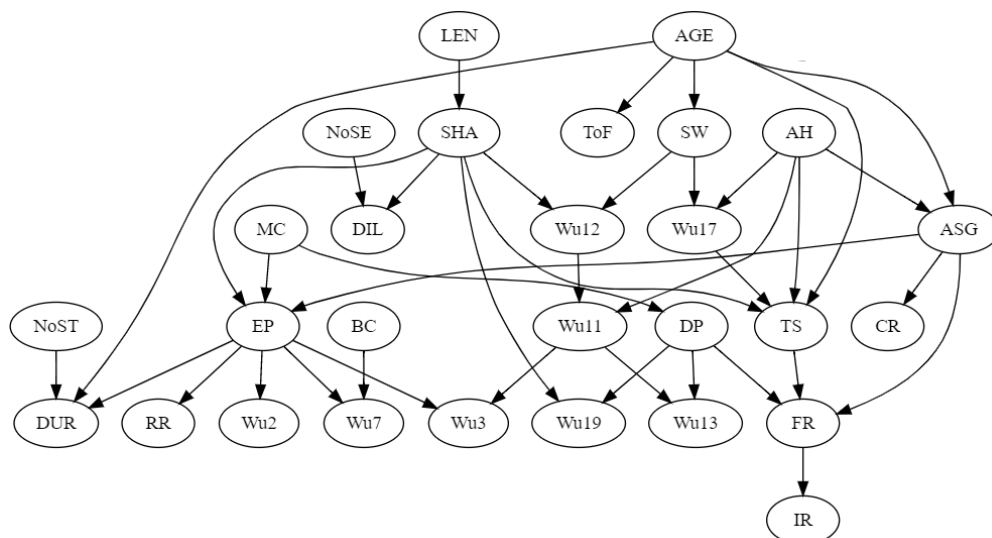
- kryterium optymalizacyjne LL prowadzi do najlepszego dopasowania wzorców treningowych (tab. 2 i 4), ale generuje największe błędy klasyfikacji dla zbiorów testowych (tab. 3 i 5). Wynika z tego, że zarówno dla metody HC, jak i TS, wyłonione struktury BBN nie posiadają dobrych własności generalizujących i prowadzą do przeuczenia modelu,
- najgorsze rezultaty uzyskano dla predykcji wskaźników uszkodzeń w_{u7} i w_{u17} – por. tab. 3 i 5. Przy czym odpowiadające wyniki w zbiorze treningowym osiągają wartość poprawnej klasyfikacji na poziomie $ac^{HC-TR} \approx ac^{HC-TS} \approx 90\%$. Różnica między wynikami dla zbioru treningowego i testowego świadczy o przeuczeniu modelu w zakresie predykcji wskaźników w_{u7} i w_{u17} .
- najlepsze rezultaty w odniesieniu do predykcji wartości wskaźników uszkodzeń w_{u7} i w_{u17} , zarówno w zbiorze treningowym, jak i testowym, uzyskano przyjmując jako kryterium optymalizacyjne miarę BDla i K2 – por. tab. 3 i 5,
- w przypadku predykcji pozostałych wskaźników uszkodzeń: w_{u2} , w_{u3} , w_{u11} , w_{u12} , w_{u13} i w_{u19} , pomijając kryterium LL, obydwie metody uczenia generują modele osiągające w symulacjach bardzo wysoki poziom trafności, zarówno w zbiorze treningowym i testującym. Nie ma tutaj znaczącego wpływu wybór poszczególnych kryteriów optymalizacyjnych.

W ramach **etapu pierwszego**, czyli oceny, ilościowej wyodrębniono łącznie 4 struktury, dla których uzyskano najlepsze wyniki predykcji dla wszystkich wskaźników uszkodzeń (łącznie ze wskaźnikami w_{u7} i w_{u17}) oraz potwierdzono ich własności generalizacyjne. Były to struktury utworzone metodą HC i TS z zastosowaniem kryteriów optymalizacyjnych BDla i K2. Wyniki w postaci graficznej zaprezentowano na rys. 3 do 6.



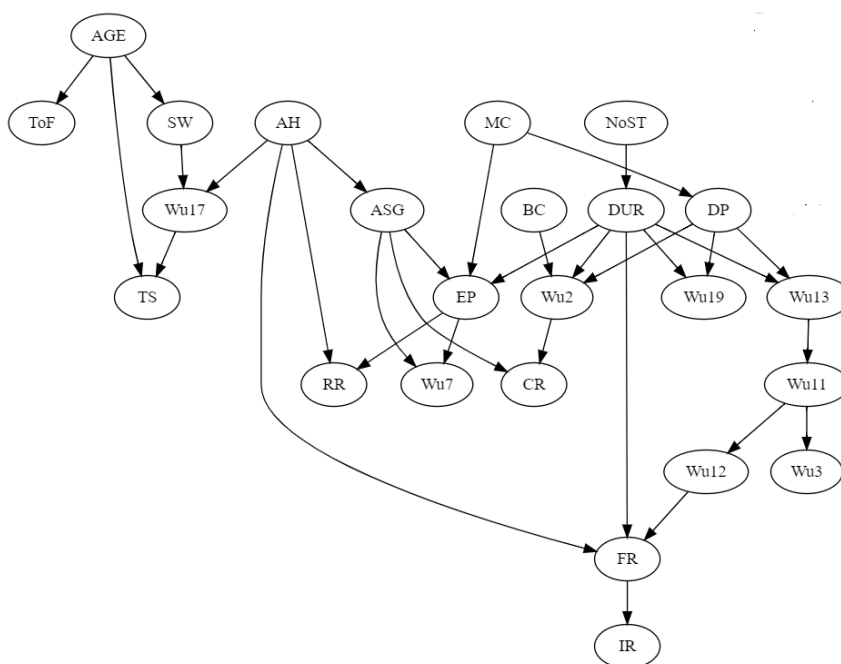
Rys. 3. Struktura BBN wyłoniona metodą HC przy zastosowanym kryterium optymalizacyjnym BDla

Fig. 3. The BBN structure extracted by the HC method using the BDla optimisation criterion



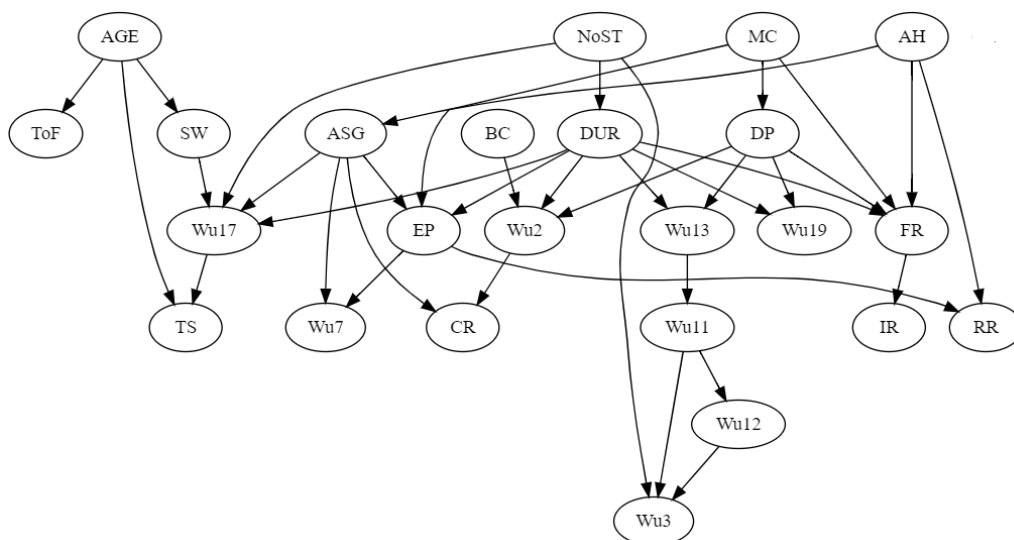
Rys. 4. Struktura BBN wyłonią metodą TS przy zastosowanym kryterium optymalizacyjnym BDla

Fig. 4. The BBN structure extracted by the TS method using the BDla optimisation criterion



Rys. 5. Struktura BBN wyłonią metodą HC przy zastosowanym kryterium optymalizacyjnym K2

Fig. 5. The BBN structure extracted by the HC method using the K2 optimisation criterion



Rys. 6. Struktura BBN wyłoniona metodą TS przy zastosowanym kryterium optymalizacyjnym K2

Fig. 6. The BBN structure extracted by the TS method using the K2 optimisation criterion

W przypadku struktury BBN uzyskanej metodą HC i kryterium BD1a (rys. 3), można zauważyć, że pominięte są dwa węzły reprezentujące zmienne: **NoSE** i **LoSE**. Analogicznie w przypadku struktury wyłonionej metodą TS i kryterium BD1e. Tutaj pominięty jest jedynie jeden węzeł wskazujący zmienną **LoSE**. Zmienna **NoSE** występuje natomiast w łącznej strukturze sieci BBN – rys. 4.

Z kolei w przypadku kryterium K2, bez względu na zastosowany algorytm uczący, poza wyłoniętymi strukturami (rys. 5 i 6) pozostają węzły **LoSE**, **NoSE**, **LEN**, **DIL**, **SHA**.

Cechy geometryczne danego obiektu (por. tab. 1), mają wpływ na jego odporność [27], a co za tym idzie, także na ryzyko powstania uszkodzeń. Dlatego wydaje się, na podstawie wieloletnich doświadczeń autorów, że bliższe rzeczywistości są modele reprezentowane przez struktury sieci BBN wyłonione przy zastosowaniu algorytmu TS oraz kryterium BD1a – por. rys. 4. W tym przypadku uzyskano najpełniejszy opis analizowanego problemu, biorąc pod uwagę zbiór aktywnych zmiennych zawartych w sieci BBN reprezentowanych przez poszczególne węzły.

6. Podsumowanie

W pracy zaprezentowane zostały wyniki analiz dotyczących budowy *sieci Bayesowskiej* (BN) do prognozowania zakresu i oceny uszkodzeń budynków poddanych wpływom eksploatacji górniczej. Z uwagi na fakt, że ustalona *sieć Bayesowska* reprezentuje wynikowe prawdopodobieństwo występowania uszkodzeń oraz wszystkich czynników warunkujących ich powstanie, może stanowić model do oceny ryzyka powstania uszkodzeń. Ostatecznie model taki został opracowany na podstawie informacji dotyczących 129 wielkogabarytowych budynków wielorodzinnych i użyteczności publicznej o konstrukcji żelbetowej prefabrykowanej, wzniesionych w systemie budownictwa uprzemysłowionego WBL. W bazie danych zebrano informacje o cechach konstrukcyjno-materiałowych, jakości utrzymania, a także dane dotyczące oceny uszkodzeń elementów oraz wpływu oddziaływań górniczych. Na tej podstawie, przyjmując dwa algorytmy uczenia struktur *sieci Bayesowskich* z danych (*Hill-Climbing* i *Tabu-Search*) oraz 8 różnych kryteriów optymalizacyjnych utworzono 16 struktur sieci Bayesa do celów predykcji zakresu i prawdopodobieństwa powstania uszkodzeń.

Uzyskane *sieci Bayesowskie* poddano ocenie ilościowej i jakościowej. Ocena ilościowa polegała na weryfikacji zgodności predykcji poszczególnych sieci z danymi ze zbiorów treningowych i testowych. Tym samym sprawdzono również stopień generalizacji nabytej wiedzy. W ramach tych analiz wyodrębniono 4 struktury, które poddano dalszej analizie jakościowej. Z kolei, w trakcie analizy jakościowej oceniano kompletność utworzonych struktur pod względem aktywnie połączonych węzłów reprezentujących poszczególne zmienne potencjalnie mogące wpływać na proces powstawania uszkodzeń. Rezultatem było zawężenie wstępnie wyselekcjonowanych sieci z etapu analizy ilościowej do jednego modelu. Model ten został utworzony z zastosowaniem algorytmu *Tabu-Search* przy zastosowaniu kryterium optymalizacyjnego BDla.

Utworzony model ryzyka w postaci *sieci Bayesowskiej*, powstały w wyniku uczenia struktury DAG na podstawie danych, posiada kilka bardzo istotnych zalet:

- Daje możliwość zarówno przewidywania prawdopodobieństwa zaistnienia uszkodzeń, jak i diagnozy przyczyn ich powstania. Z tego powodu może być wykorzystany jako narzędzie do szacowania ryzyka powstania uszkodzeń dla dużej liczby budynków zlokalizowanych na terenie górniczym.
- Umożliwia interpretację powiązań przyczynowo-skutkowych, których znajomość poszerza wiedzę nt. modelowanego zjawiska. Szczególnie w przypadku oddziaływania zmiennych dotyczących jakości utrzymania obiektu.
- Pozwala na przeprowadzenie wnioskowania względem dowolnej zmiennej zawartej w strukturze DAG reprezentującej sieć *Bayesowską*.
- Może być wykorzystywane w przypadku braku ścisłej informacji nt. stanu zmiennych w poszczególnych węzłach. Umożliwia zatem stosowanie modelu w obszarze niepewności.
- Może być łatwo aktualizowany w przypadku rejestracji nowych danych wynikających z obserwacji rzeczywistego przebiegu danego procesu.

Dodatkowo, tego rodzaju metodyka może być zastosowana w dużo szerszym zakresie niż tylko do oceny wpływów eksploatacji górniczej. Jako przykładowe można wymienić tu: monitoring stanu technicznego konstrukcji (*Structure Health Monitoring*) lub zarządzanie utrzymaniem technicznym i konserwacją obiektów budowlanych (*Maintenance Management*).

Uzyskane wyniki dają asumpt do dalszych badań w tym zakresie, zarówno w kwestii analizy budynków innego typu, jak i testowania innych metod pozwalających na uczenie struktur *sieci Bayesowskich* z danych.

Literatura

1. Abdelkader EM, Zayed T, Marzouk M. A computerized hybrid Bayesian-based approach for modelling the deterioration of concrete bridge decks. *Structure and Infrastructure Engineering* 2019; 15: 1178–99.
2. Carvalho AM. Scoring functions for learning Bayesian networks. Inesc-id Tec. Rep 2009; 12.
3. Chung H, Lee I-M, Jung J-H, Park J. Bayesian networks-based shield TBM risk management system: Methodology development and application. *KSCE Journal of Civil Engineering* 2019; 23: 452–65.
4. Dahire S, Tahir F, Jiao Y, Liu Y. Bayesian Network inference for probabilistic strength estimation of aging pipeline systems. *International Journal of Pressure Vessels and Piping* 2018; 162: 30–9.
5. Fang S-E, Tan J, Zhang X-H. Safety evaluation of truss structures using nested discrete Bayesian networks. *Structural Health Monitoring-an International Journal* 2020.

6. Fedorowicz J, Słowik L. Interpretation of the behavior of a system building object-difficult subsoil in modern numerical modeling. *Proceedings of the 11th International Conference on New Trends in Statics and Dynamics of Buildings* October. Vol3. 2013.
7. Fereshtehnejad E, Banazadeh M, Shafieezadeh A. System reliability-based seismic collapse assessment of steel moment frames using incremental dynamic analysis and Bayesian probability network. *Engineering Structures* 2016; 118: 274–86.
8. Firek K. Proposal for classification of prefabricated panel building damage intensity rate in mining areas. *Archives of Mining Sciences* 2009.
9. Gogoshin G, Boerwinkle E, Rodin AS. New algorithm and software (BNOmics) for inferring and visualizing bayesian networks from heterogeneous big biological and genetic data. *Journal of Computational Biology* 2017; 24: 340–56.
10. Golewski GL. Measurement of fracture mechanics parameters of concrete containing fly ash thanks to use of Digital Image Correlation (DIC) method. *Measurement* 2019; 135: 96–105.
11. Heckerman D, Geiger D, Chickering DM. Learning Bayesian networks: The combination of knowledge and statistical data. *Machine learning* 1995; 20: 197–243.
12. Jin Y, Zhang J, Sun L. Safety Risk Assessment of Prefabricated Building Construction Based on Bayesian Network. *IOP Conference Series: Earth and Environmental Science*. Vol371. 2019: 32052.
13. Ravikumar K, Rajiv Kannan A. Survey of Spatial Datamining Methods for Natural Disaster Management. *Middle-East Journal of Scientific Research* 2017; 25: 217–27.
14. Kębłowski MW, Grzyl B, Kristowski A, Siemaszko A. Risk Modelling with Bayesian Networks - Case Study: Construction of Tunnel under the Dead Vistula River in Gdansk. *Procedia Engineering* 2017; 196: 585–91.
15. Knyziak P, Kanoniczak M. Difficulties in Operation of Elevations in Large-Panel Buildings. *IOP Conference Series: Materials Science and Engineering*. Vol661. 2019: 12059.
16. Lee S, Choi M, Lee H-S, Park M. Bayesian Network-based Seismic Damage Estimation for Power and Potable Water Supply Systems. *Reliability Engineering & System Safety* 2020: 106796.
17. Marsili F, Croce P, Klawonn F, Vignoli A, Boschi S, Landi F. A Bayesian network for the definition of probability models for masonry mechanical parameters. *14th International Probabilistic Workshop*. 2017: 253–68.
18. Nagarajan R, Scutari M, Lèbre S. *Bayesian networks in r*. Springer 2013; 122: 125–7.
19. Niedojadło Z, Gruszczynski W. The impact of the estimation of the parameters values on the accuracy of predicting the impacts of mining exploitation. *Archives of Mining Sciences* 2015; doi 10.1515/amsc-2015-0012.
20. Nielsen JS, Sørensen JD. Computational framework for risk-based planning of inspections, maintenance, and condition monitoring using discrete Bayesian networks. *Structure and Infrastructure Engineering* 2018; 14: 1082–94.
21. Nielsen JS, Tcherniak D, Ulriksen MD. A case study on risk-based maintenance of wind turbine blades with structural health monitoring. *Structure and Infrastructure Engineering* 2020: 1–17.
22. Niewiadomski P, Hoła J. Failure process of compressed self-compacting concrete modified with nanoparticles assessed by acoustic emission method. *Automation in Construction* 2020; 112: 103111.

23. Pachla F, Tataro T. Dynamic Resistance of Residential Masonry Building with Structural Irregularities. *Seismic Behaviour and Design of Irregular and Complex Civil Structures III*. Springer, 2020: 335–47.
24. Pearl J. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Elsevier, 2014.
25. Rusek J. Application of Support Vector Machine in the analysis of the technical state of development in the LGOM mining area. *Eksploracja i Niezawodność – Maintenance and Reliability 2017*; 19 (1): 54–61, <http://dx.doi.org/10.17531/ein.2017.1.8>.
26. Rusek J, Firek K. Assessment of technical condition of prefabricated large-block building structures located in mining area using the Naive Bayes classifier. *International Multidisciplinary Scientific GeoConference Surveying Geology and Mining Ecology Management, SGEM. Vol2*. 2016.
27. Rusek J. The Point Nuisance Method as a Decision-Support System Based on Bayesian Inference Approach. *Archives of Mining Sciences 2020*; 65: 117–27.
28. Sadowski L. Non-destructive investigation of corrosion current density in steel reinforced concrete by artificial neural networks. *Archives of Civil and Mechanical Engineering 2013*; 13: 104–11.
29. Sadowski Ł, Hoła J, Czarnecki Sł. Non-destructive neural identification of the bond between concrete layers in existing elements. *Construction and Building Materials 2016*; 127: 49–58.
30. Sari DP, Rosadi D, Effendie AR, others. Application of Bayesian network model in determining the risk of building damage caused by earthquakes. *2018 International Conference on Information and Communications Technology (ICOIACT)*. 2018: 131–5.
31. Schabowicz K, Gorzelańczyk T. A nondestructive methodology for the testing of fibre cement boards by means of a non-contact ultrasound scanner. *Construction and Building Materials 2016*; 102: 200–7.
32. Scutari M. Dirichlet Bayesian network scores and the maximum relative entropy principle. *Behaviormetrika 2018*; 45: 337–62.
33. Scutari M. Learning Bayesian networks with the bnlearn R package. *arXiv preprint arXiv:0908.38172009*; doi 10.18637/jss.v035.i03.
34. Scutari M, Graafland CE, Gutiérrez JM. Who learns better Bayesian network structures: Accuracy and speed of structure learning algorithms. *International Journal of Approximate Reasoning 2019*; 115: 235–53.
35. Scutari M, Vitolo C, Tucker A. Learning Bayesian networks from big data with greedy search: computational complexity and efficient implementation. *Statistics and Computing 2019*: 1–14.
36. Siemaszko A, Jakubczyk-Gałczyńska A, Jankowski R. The Idea of Using Bayesian Networks in Forecasting Impact of Traffic-Induced Vibrations Transmitted through the Ground on Residential Buildings. *Geosciences 2019*; 9: 339.
37. Sperotto A, Molina J-L, Torresan S, Critto A, Marcomini A. Reviewing Bayesian Networks potentials for climate change impacts assessment and management: A multi-risk perspective. *Journal of environmental management 2017*; 202: 320–31.
38. Tajduś K, Misa R, Sroka A. Analysis of the surface horizontal displacement changes due to longwall panel advance. *International Journal of Rock Mechanics and Mining Sciences 2018*; doi 10.1016/j.ijrmms.2018.02.005.

39. Tran T-B, Bastidas-Arteaga E, Aoues Y et al. Reliability assessment and updating of notched timber components subjected to environmental and mechanical loading. *Engineering Structures* 2018; 166: 107–16.
40. Wang F, Li H, Dong C, Ding L. Knowledge representation using non-parametric Bayesian networks for tunneling risk analysis. *Reliability Engineering & System Safety* 2019; 191: 106529.
41. Wodyński A, Lasocki S. Assessment of mining tremor influence on the technical wear of building. *Acta Geodynamica et Geomaterialia. Ser. A and Ser. B* 2004; 50: 187–94.
42. Wu L, Wang J, Zhou J et al. Multi-scale geotechnical features of dredger fills and subsidence risk evaluation in reclaimed land using BN. *Marine Georesources & Geotechnology* 2019: 1–23.
43. Yazdani A, Shahidzadeh M-S, Takada T. Bayesian networks for disaggregation of structural reliability. *Structural Safety* 2020; 82: 101892.
44. Yuan M, Liu Y, Yan D, Liu Y. Probabilistic fatigue life prediction for concrete bridges using Bayesian inference. *Advances in Structural Engineering* 2019; 22: 765–78.