ORIGINAL PAPER

Classification and regression tree theory application for assessment of building damage caused by surface deformation

Agnieszka Malinowska

Abstract A framework of applying the classification and regression tree theory (CART) for assessing the concrete building damage, caused by surface deformation, is proposed. The prognosis methods used for approximated building hazard estimation caused by continuous deformation are unsatisfactory. Variable local soil condition, changing intensity of the continuous deformation and variable resistance of the concrete buildings require the prognosis method adapted to the local condition. Terrains intensely induced by surface deformation are build-up with hundreds of building, so the method of their hazard estimation needs to be approximated and relatively fast. Therefore, promising might be addressing problems of reliable building damage risk assessment by application of classification and regression tree. The presented method based on the classification and regression tree theory enables to establish the most significant risk factors causing the building damage. Chosen risk factors underlie foundation for the concrete building damage prognosis method, which was caused by the surface continuous deformation. The established method enabled to assess the severity of building damage and was adapted to the local condition. High accuracy of shown approach is validated based on the independent data set of the buildings from the similar region. The research presented introduces the CART to determination of the risk of building damage with the emphasis on the grade of the building damage. Since presented method bases on the observations of the damages from the previous subsidence, the method might be applied to any local condition, where the previous subsidence is known.

Keywords Subsidence · Building damage · Classification and regression tree theory · Hazard estimation

A. Malinowska (✉)
AGH University of Science and Technology, Al. Mickiewicza 30, 30059 Kraków, Poland
e-mail: amalin@agh.edu.pl

1 Introduction

The correct evaluation of the class of building's potential damage, induced by surface deformation, has been analyzed over the World for years. The main problem lies in its scale, i.e., necessity to quickly evaluate the potential risk for a big group of buildings. Methods used in Poland nowadays are both time-consuming and very inaccurate. The approximate estimation of vulnerability of buildings in Poland is based on a common theoretical assumption that the more numerous are the risk factors, the more accurately the vulnerability of the buildings can be estimated (e.g., Polish point method). The building damage risk is assessed by comparing the vulnerability of buildings in hazarded areas, associated with maximum strains and tilts (GIG 2000). The final result of the presently applied method is information that the buildings can be damaged (or to be more precise, that their usability level has dropped down). It should be emphasized that the method for assessing damage risk for buildings presently used in Poland was created in the 1950s (Knothe 1953; Przybyła and Świądrowski 1968). The criteria of negative aspect of continuous deformations of terrain assumed in this method were based on experiences dating back to the 1950s. This method was created for evaluating the impact of hard coal exploitation on continuous deformations of terrain. Later in time, this method was adopted for evaluating risks generated by exploitation of other raw minerals. The character and dynamics of deformation of terrain are different for copper ores deposits or salt domes. The variable dynamics and range of continuous deformations, and different construction characteristic typical of various counties definitely have influence on the scale of potential damage. This certainly negatively affects the efficiency of the currently applied method of assessing damage risk for buildings in areas subjected to continuous deformations. Moreover, with this method, one cannot evaluate the degree to which the building was damaged, which is a considerable drawback, hindering the process of managing damage in areas staying under the influence of continuous deformation. A number of other methods which are satisfactorily adaptable to the local conditions, thanks to which the potential risk can be assessed for the buildings, are used worldwide (Bhattacharya and Singh 1984; Boscarding and Cording 1989; Burland and Wroth 1975; Mine Subsidence Engineering Consultants 2007; National Coal Board 1975; Yu et al. 1988). It is worth noting that these methods tend to minimize the negative factors. Methods based on two or three variables enable quick evaluation of the risk. World's investigations focused on the evaluation of building's damage due to continuous deformations and seismic tremors are aimed at creating a classification thanks to which the class of building damage can be predicted (Askan and Yucemen 2010; Grünthal et al. 2006; Grünthal and Wahlström 2006; Pei and van de Lindt 2009; Straub and Kiureghian 2008; Yucemen et al. 2004). Analyses made by the author reveal that the methods used worldwide are equally or sometimes even more accurate than the ones used in Poland (Malinowska 2013a, b). Moreover, the application of these methods is also much simpler. The accuracy of selected methods mainly depends on the proper selection of risk factors and scaling dependences between the risk and damage class.

Basing on the World's and Polish experience, the author undertook an attempt at creating a method employing minimum number of risk factors and which could be used for efficient and reasonably accurate evaluation of the building damage class. Attention was mainly paid to working out a classification algorithm for defining the degree to which the building was damaged by continuous deformations in view of local conditions. The algorithm presented in this paper can be used for adapting this method to the assessment of building damage class in mining areas in changeable exploitation conditions. The analyses were made with the geographic information system (GIS) and a classification and regression theory (CART).

2 A new method for assessing building's damage risk

The investigation was done in followed stages (Fig. 1).

At the first stage, two research areas (training and testing) were selected. At second stage, information about the buildings in that area, as well as potential and actual risk factors was acquired (paragraph 2.1, Fig. 1 No. 1–2).

At the next stage, all factors, which may contribute to the building damage, were analyzed. On this basis, the significance of factors could be determined in view of their influence on building's potential damage (paragraph 2.2, Fig. 1 No. 3–5).

At the following stage, attempts were made to determine boundary conditions of risk factors using the interactive trees CART. A detailed characteristic was made for determining the damage class to particular elements of the building. This created bases for defining risk factors, for which the building could be damaged slightly, moderately or very severely. The basic aim of this stage was finding detailed criteria of risk factor boundary values, thanks to which classes of damage class could be discriminated (paragraph 2.3, Fig. 1 No. 6–7).

At the fourth stage, the efficiency of the classification was verified. The new method was used for an independent group of buildings sited in the neighboring area. The obtained results were compared with the efficiency of methods presently used in Poland (Fig. 1 No. 8–9). The results of estimation of the damage risk were confronted with actual damage observed in buildings (paragraph 2.4, Fig. 1 No. 10–11).

2.1 Characteristic of research area

The empirical data were collected in a hard coal mine area in the Upper Silesia. Two groups of excavation-induced buildings under the continuous surface deformations were selected [training set ZB1 (Figs. 2, 3), testing set ZB2 (Figs. 4, 5)].

One of the major assumptions of the presented surveys was that the structure of buildings sited in the neighboring area (within suburban areas) was similar as far as their design and vulnerability were concerned. In the selected areas, there were single-family dwelling houses, detached houses, single- to two-story objects typical of an industrial–agricultural settlement. About 80 % of all objects were residential houses made of brick, mostly with a cellar. The remaining objects were farm or warehouses. Most of the buildings (ca. 80 %) were erected before 1985.

Mining done in those areas caused surface subsidence. The essential information about mining has been presented in Tables 1 and 2.

The surveyed areas were so selected as to make terrain deformations the principal factor responsible for damage. The spatial distribution of deformations in that area could not be determined on the basis of the results of measurements. The frequency of the monitoring was too low and the density of observation points insufficient to determine the spatial distribution of extreme indices. Terrain deformations were assessed with the use of Knothe's functional method of predicting continuous deformations (Knothe 1953). Measurements realized in the surveyed areas enabled one to only define assumed prediction parameters more accurately. Buildings covered the continuous deformation impact area of small and considerable strains, sometimes reaching up to 8 mm/m).

The research works were focused on acquiring information about:

- Construction and vulnerability characteristic of buildings preceding the exploitation,
- Maximum values of deformation indices observed under each of the buildings,
- Observed exploitation-induced damage.
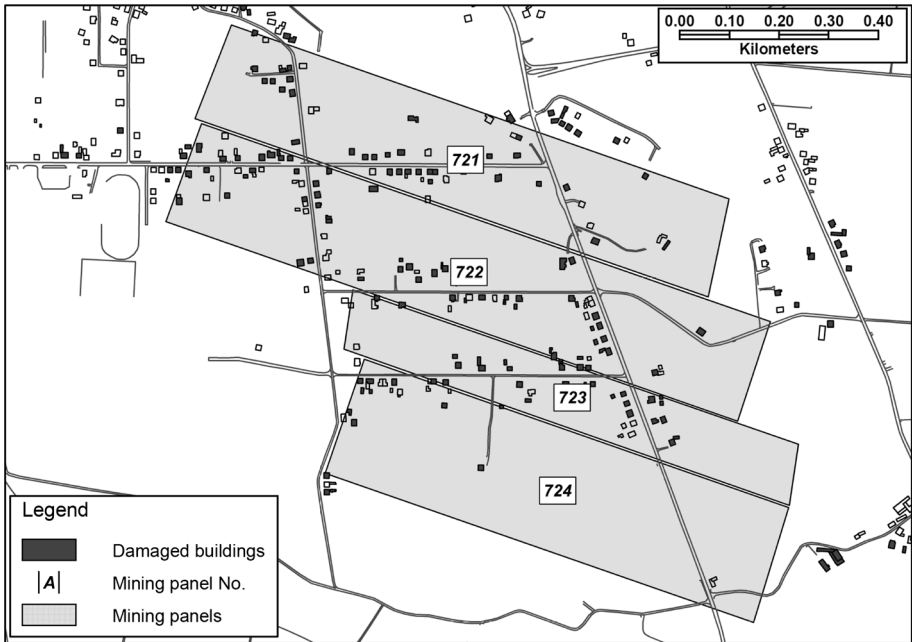
Fig. 1 Flowchart of model definition algorithm
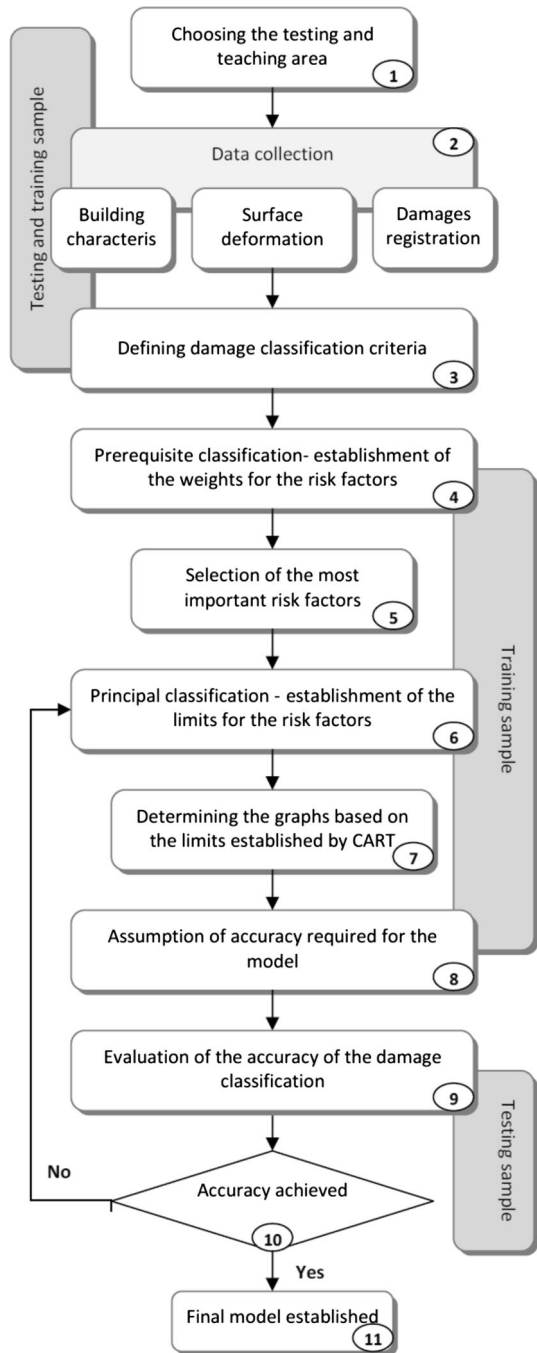
Fig. 2 Research training area ZB1-mining panels localization
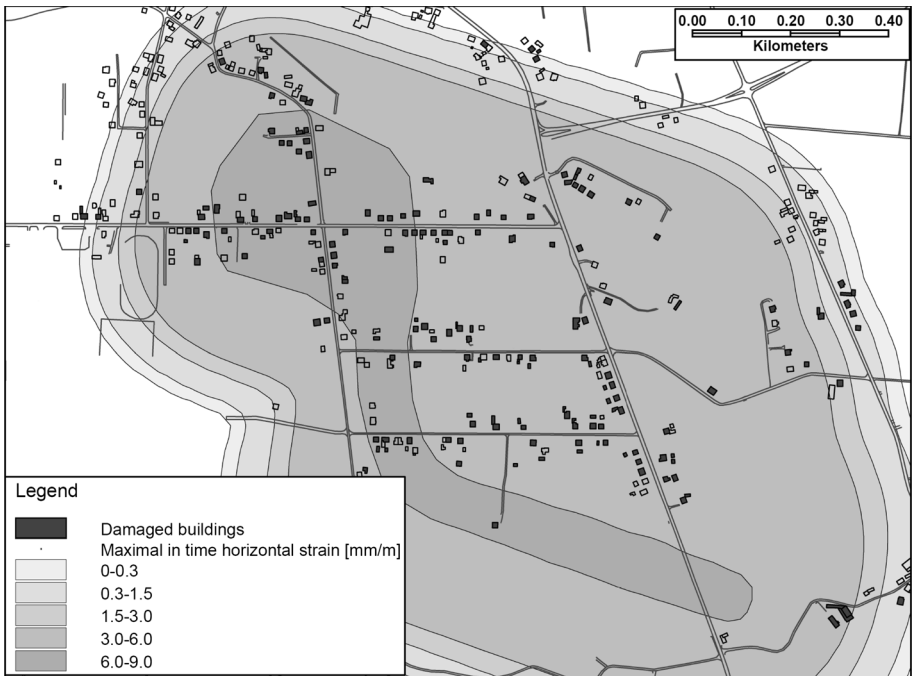

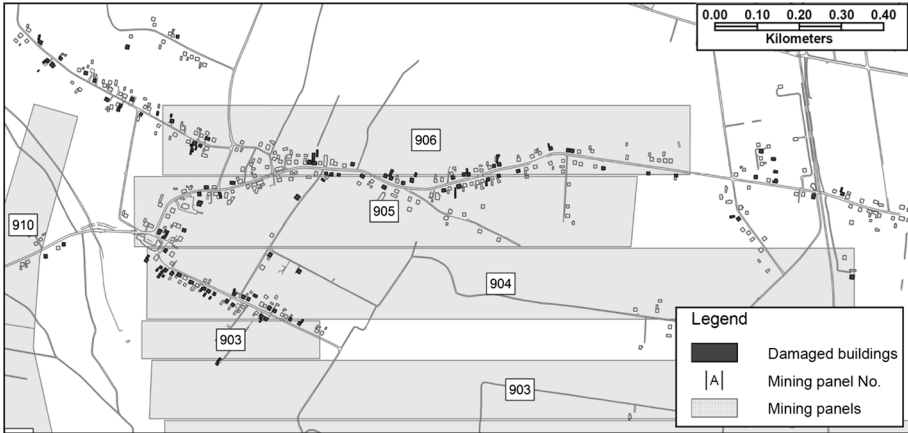**Fig. 3** Research training area ZB1-horizontal strain distribution

Fig. 4 Research training area ZB2-mining panels localization
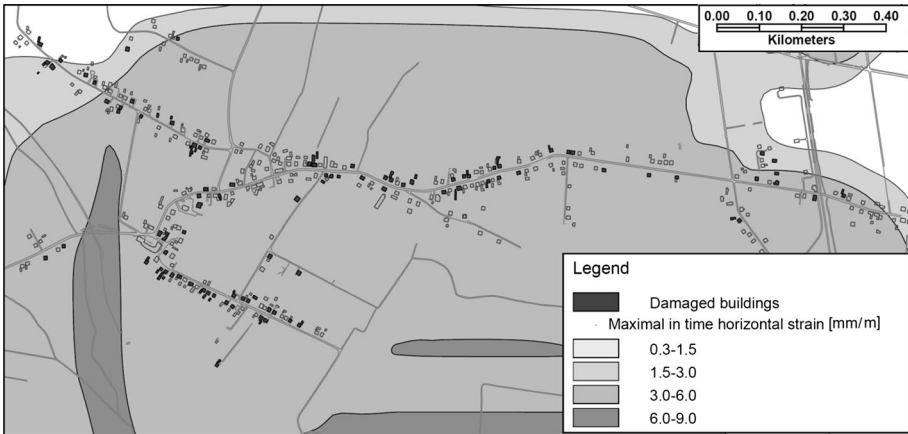

Fig. 5 Research training area ZB2-horizontal strain distribution

2.2 Knowledge discovery with CART

2.2.1 CART background

When the qualitative and quantitative analysis of variables is required and when large sets of data are involved, then data mining technique is a very good solution. A number of classification trees are used for the CHAID, CART, QUEST practice. One of the most frequent techniques of exploitation and classification of data is the CART method (Breiman et al. 1984). Its popularity results from, e.g., its basic assumptions regarding data, on the basis of which the classification is made. Variables used for modeling do not need to have a normal distribution (Rothwell et al. 2008). Moreover, the classification algorithm used in the CART is resistant to the missing data. The analyses prove that when the missing data do not exceed 5 % of the set, satisfactory results of CART classification can

Table 1 Basic information about mining in area ZB1

Mining panel no.	Mining panel depth (m)	Mining panel thickness (m)	Speed of mining (m/day)	Date of mining start	Date of mining end
721	574	3.0	2.7	VII 2002	V 2003
722	568	3.0	6.1	V 2003	X 2003
723	535	3.0	4.9	XII 2003	XI 2004
724	521	3.0	2.9	XI 2004	XI 2005

be expected (Yong 2006). Thanks to its transparence and high accuracy, the method is used in various branches of science, e.g., medicine (Ture et al. 2009), photogrammetry (Yang et al. 2003), economics (Chang 2011), environmental protection (Smeti et al. 2009), food science (Hazir et al. 2012) and chemistry (Chudzinska and Baralkiewicz 2011). The CART classification is made in a few stages. The first one lies in assuming preliminary criteria of division, stopping and construction of the tree. In the analyses conducted by the author, the goodness of fit for a given class was estimated with the Gini Measure. Gini Measure helps one estimate how frequently a random element will be misclassified. This index is calculated as a sum of random probabilities of a given element to its misclassification ratio. For elements assuming values from the interval 1 to m, the Gini index at node t is calculated from the below dependence (Breiman et al. 1984):

$$I(t) = 1 - \sum_{i=1}^{m}(P_i(j/t))^2 \qquad (1)$$

where $I(t)$ Gini index, $P_i(j/t)$ is the relative frequency of class j at node t, t number of node.

Classification created at the first stage of research very frequently leads to generating a tree with too many branches. The complexity of the tree may be caused by considerable dispersion of data. The size of the tree also depends on given stop parameters, i.e., minimum population in the successive nodes, minimum population of children, maximum number of levels and maximum number of nodes. Many years' experience has shown that mainly the population of elements in the terminal nodes is decisive for the classification of a small number of cases (Breiman et al. 1984). It should be emphasized that the increased size of the tree is not proportionate to higher accuracy of classification. The generated trees tend to lower the actual ability to make a correct classification by overfit of the training set. Generating biggest possible trees is a common practice. Then follows the stage of cutting. This procedure lies in decreasing the number of leaves, simultaneously tending to maximize the accuracy of classification. The last stage is the selection of a tree with best classification potential. A tree with the smallest number of misclassifications is considered to be optimal. The accuracy of the created model is realized with the use of cross-validation.

$$RE(d) = \frac{1}{N}\sum_{i=1}^{N}(y_i - d(x_i))^2 \qquad (2)$$

where y_i points in testing set (real variable), x_i points in testing set (variable classified with d model), N number of cases in a testing set.

The set of training data is divided into training and testing subsets. The prediction abilities of a tree are evaluated on the basis of the testing set data. The commonly applied practice of using CART trees advises choosing a tree with the minimum number of terminal nodes and simultaneously lowest cross-validation value. The measure of the

Table 2 Basic information about mining in area ZB2

Mining panel no.	Mining panel depth (m)	Mining panel thickness (m)	Speed of mining (m/day)	Date of mining start	Date of mining end
903	705	4	6.5	IX 2006	I 2007
904	700	4.2	6.2	VII 2007	XII 2008
905	695	4.2	5.8	II 2009	III 2010
906	690	4.2	5.1	VIII 2010	VII 2011

cross-validation $R_\alpha(T)$ is a linear dependence between the complexity of tree and the cost of misclassifications (Breiman et al. 1984).

$$R_\alpha(T) = R(T) + \alpha\widetilde{T} \Leftrightarrow \alpha = \frac{R_\alpha(T) - R(T)}{\left|\widetilde{T}\right|} \tag{3}$$

where $R_\alpha(T)$ cost-complexity measure, $R(T)$ cost of misclassifications, $\left|\widetilde{T}\right|$ complexity of tree measures as the number of terminal nodes in the tree, α parameter of tree complexity (assumes values from 0—for maximal tree to 1 for minimal tree).

An optimal tree has minimal complexity and at the same time lowest possible classification error (Fig. 6). Presented graph is an example of relation between complexity of the tree and accuracy of classification. If the tree is to complex, it is usually overfitted. However, to large generalization of the tree (low complexity) leads to low accuracy of classification. This dependence has been taken into account during estimation of optimal tree in the research.

2.2.2 Preparing data for classification

Classification was made on a set of training data ZB1 in a town X. The population was 141, and the buildings had a detailed characteristic of damage of their particular elements. An attempt was made to make boundary values for risk factors on the basis of registered damage class more precise. The aim of this operation was determining the boundary values of risk factors, for which the building could be damaged to some extent. Basing on the registered damage, seven types of damage in buildings were finally distinguished (Table 3). Four Boolean-type variables: falling off tiles, woodworks damage, falling off plaster and chimney ducts damage. These variables were ascribed true (damaged) and false (undamaged) values. Damage could be specified in more detail for three variables. These were: partition wall damage, load-bearing wall damage and ceiling damage. These variables could be characterized as: no damage, small, moderate and high. The intensity of load-bearing walls, partition walls and ceiling damaging was determined on the basis of the aperture of fissures, their distribution and intensity of occurrence. The direction of fractures was also taken into account in the evaluation.

The variables were classified, and four damage classes were distinguished (Table 4). Criteria for evaluation of the final damage categories have been established based on Mine Subsidence Engineering Consultants classification and Burland damage classification. Basis for the classes was information about crack width limits (Burland and Wroth 1975; Mine Subsidence Engineering Consultants 2007). The damaging of useful elements, i.e., sewage systems, chimney ducts, falling off tiles, woodworks, was classified as functional damage.

An output modeling-dependent variable was obtained as a result of the classification of thus observed building damage. The quality predictors which have the biggest influence on

the damage class were: year in which the building was erected and maximum horizontal strains.

2.2.3 Determining significance of risk factors

Another stage was finding independent risk factors deciding about damage occurrence in a building and evaluation of its significance. The variables, which may have influence on the potential damage class, can be divided into two groups. The first of them is related to the building and its vulnerability to horizontal strains. This may result from the construction properties, age of the object, way in which it was used, its geometry and foundation. The second group of risk factors can be associated with horizontal strains of terrain surface under the influence of underground exploitation (or other phenomena generating surface deformations). World's investigations reveal that maximal horizontal strains have the biggest influence on damage risk of space objects. Factors that are independent variables were assumed for the analyses. The factors, which were determined on the basis of building's characteristic, were refuted at the stage of preliminary data selection. Therefore, the following variables were not analyzed: technical wearing (directly related to the age of the object) and vulnerability category (or number of vulnerability points, determined on the basis of, e.g., length of the building, compaction of structure, existing protection measures, functions). Eventually, six independent variables were chosen for further analyses (Fig. 7). The risk caused by continuous deformations of terrain surface was attributed to the maximal horizontal strains, which assumed values from 0 to 9 mm/m. The remaining factors were also related to the building's characteristic. Three of the factors were of quantitative character (year of construction, length of the building, number of stories) and two of qualitative nature (protections, function).

The significance of particular predictors was analyzed in view of a dependent variable, i.e., building damage, which was attributed Boolean values ('true' or 'false'). The significance analysis of predictors was conducted for a training research area ZB1, with the population of 372 buildings, among which 141 objects were damaged. No grading of damage was introduced. The group of damaged buildings covered all kinds of damage: from small to severe. According to the ranking, maximal horizontal strain and the year of construction were the most important predictors of building damage (Table 5). The weights were established based on classification and regression tree. The four remaining variables occupy a lower position in the ranking.

It should be noted that the weight of predicted horizontal strain factor is definitely highest as compared to other risk factors. This is important in view of methods presently applied for building risk evaluation, where the risk factors, i.e., vulnerability of buildings and surface deformations, are assumed to have the same significance. The analyses made by the author reveal that the vulnerability of buildings, estimated with the approximated point method, has a definitely weaker influence on potential damage than horizontal strains (Table 6).

This proved the assumption that new risk factors have to be defined.

2.2.4 Classification

Classification was made for 141 buildings, which were damaged to a varying degree (C1–C4). The assumed predictors were: year of construction and maximal horizontal strains. The categorized dependent variable was the damage class, as discussed in above section.

Fig. 6 Finding optimal tree
(Lewis 2000)
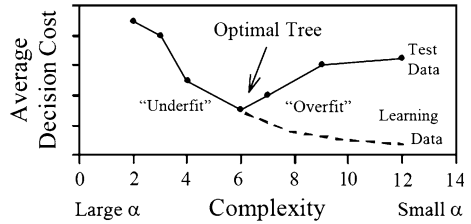
During the first stage of research, the biggest possible tree was generated. It had 135 terminal nodes and 15 levels (Fig. 8). The tree was generated without limitations in the sense of stop parameters.

The stop parameters were successively given, as according to Breiman et al. (1984) they are can be used for obtaining an optimal tree. These were:

- Minimal number in the successive nodes 5,
- Minimal population of child 5,
- Maximal number of levels 10,
- Maximal number of nodes 1,000.

A tree of 19 terminal nodes and lowest cross-validation value was obtained for such parameters. In view of the assumed minimal cross-validation value (plus standard deviation) and simultaneously minimization of the number of terminal nodes, the tree of 19 nodes fulfilled these conditions (Table 7).

The procedure of cutting the tree was realized in stages. The least error of correct classifications was the division criterion. Finally, a 15-node tree was selected. However, the cross-validation value was higher than that of a 19-node tree. Nonetheless, this tree was called optimal in the sense of classification correctness.

Divisions introduced for the ultimate classification tree were presented in the successive stages, following consecutive nodes. The division criterion in the first node was based on the maximal horizontal strain variable. The division in this node enabled one to determine the boundary value of horizontal deformations, for which small and bigger than small damage was observed. The division in the first node was made for horizontal strain values of 1.0 mm/m (Fig. 9a). Next division was made in a child node ID 3, again based on maximal horizontal strain values. The division in this node will provide criterion of the successive intensity degree: damage (Fig. 9b). Buildings, which were affected by horizontal strains >3.0 mm/m, were damaged to moderate or high degree. Buildings exposed to horizontal strains <3.0 mm/m were damaged to a small or moderate degree.

The division in the child node ID 5 resulted from the influence of horizontal strain values. The horizontal strain of 5.5 mm/m was the boundary value found at this stage of classification (Fig. 10a). High damage was observed for horizontal strains >5.5 when classifying a child node ID. The last division was based on the predictor of the year of construction (Fig. 10b). This is the only classification division based on this predictor. The classification significance of this variable was slightly smaller than for maximal horizontal strains. Therefore, the classification division based on this variable was made only once. The year of construction is very important for the intensity of damage observed in buildings under the influence of horizontal strains >5.5 mm/m. Moderate-class damage can be expected for buildings constructed after 1950. In older buildings (constructed before 1950), high and moderate damage can be expected.

Table 3 Description of damaged elements and damage classes

Damaged element	Damage class
Load-bearing wall	No/small/moderate/high
Partition wall	No/small/moderate/high
Ceiling	No/small/moderate/high
Falling off plaster	True/false
Woodworks	True/false
Falling off tiles	True/false
Chimney duct	True/false

Table 4 Criteria of building damage classification

Damage class	Load-bearing wall damage	Partition wall damage	Ceiling damage	Functional damage
C1	–	Small	–	False
C2	Small	Small/moderate	Small	False
C3	Moderate	Moderate/large	Moderate	True
C4	Large	Large	Large	True

Finally, four damage intensity zones were distinguished between maximal horizontal strain and year of construction variables. Unlike maximal horizontal strain, the year of construction decides about the damage class to a small extent. The final classification presented in a tree form is presented below (Fig. 11).

The presented classification tree enabled one to estimate the building damage class on the basis of horizontal strain and year of construction. It should be emphasized that the tree had high classification accuracy parameters for the training set ZB1. The tree had to be verified on an independent set of data to provide an evaluation of real classification abilities.

3 Accuracy evaluation of presented method

The proposed method was assessed for its accuracy on an independent set of data about residential houses ZB2 (Figs. 4, 5). Basic assumption for the method accuracy was to achieve better results in buildings hazard estimation than it is achieved with a use of Polish point method. The efficiency analyses of this method required determining class of damage observed in buildings, as damage data had a descriptive character. Among 518 buildings, 127 were damaged.

The maximal horizontal strains in the subsoil were determined for 127 objects. Besides, the information about the year of construction was acquired. Each building was ascribed a class of damage on the basis of observed damage. The building risk was determined on the basis of collected data with the newly defined method.

The *post-factum* evaluation of building damage was made with the use of a method commonly applied in Poland. It was assumed that a small damage (C0) is observed when
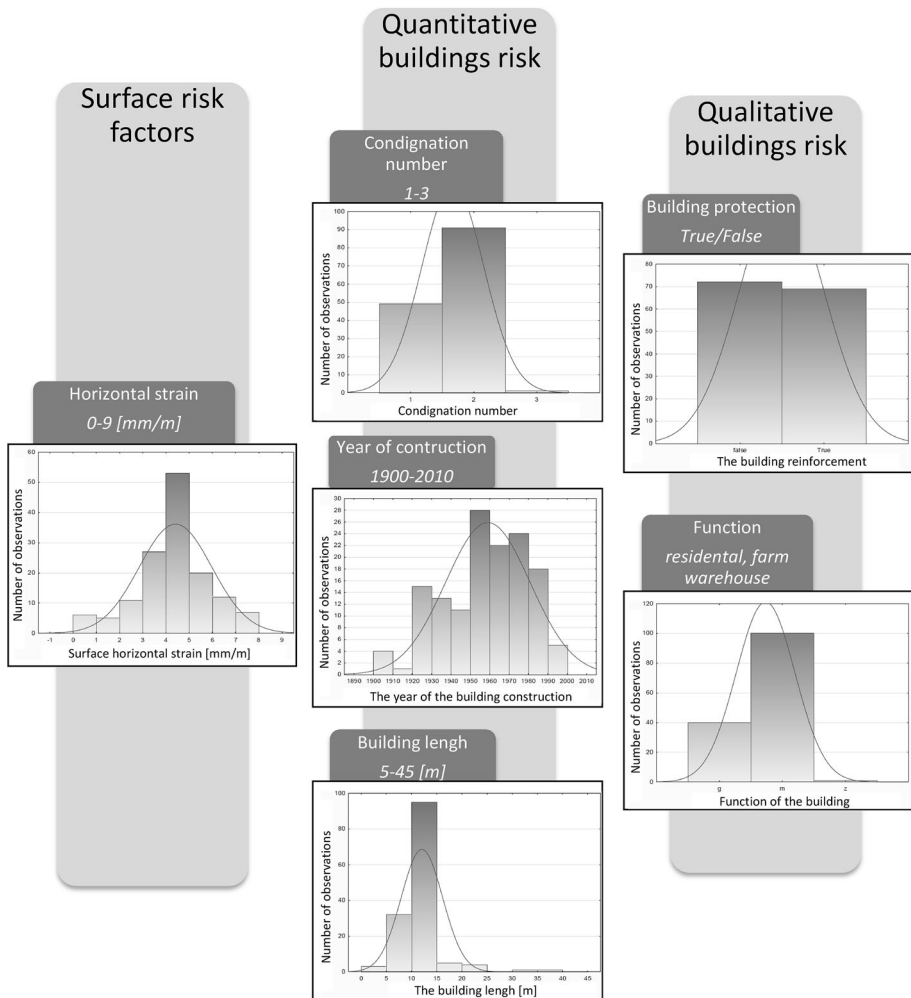
Fig. 7 Risk factors

the vulnerability category is the same as the terrain category. When the terrain category is bigger than the vulnerability class by: one—small damage is expected (C1), two—moderate damage is expected (C2), three—high damage is expected (C3). High and very high damage risk was not observed in the analyzed population (C4). The estimation results were compared with real damage in buildings (Fig. 12).

Damage classified as very low and low was observed in 32 buildings. The prediction model and the Polish method revealed low congruence of predictions and real values C1 (16 % of hits in the Polish method, 0 % of hits for the new classification). In the case of moderate damage, they were observed in 46 buildings. The damage class predicted with the use of the Polish method and new classification produced congruence of 2 % (Polish method) and 17 % of hits (new method). In the case of moderate damage class (C3), a very high congruence was noted for the new method (81 %) between the observed and estimated damage degree. In the case of estimation made with the traditional (Polish) method,

Table 5 Significance of risk factors

Variables	Weight
Predicted horizontal strain	1.00
Year of construction	0.68
Number of stories	0.60
Maximal length of wall	0.59
Function	0.33
Protection	0.14

Table 6 Significance of risk factors in method presently applied in Poland

Variables	Weight
Predicted horizontal strain	1.00
Points of building vulnerability	0.41
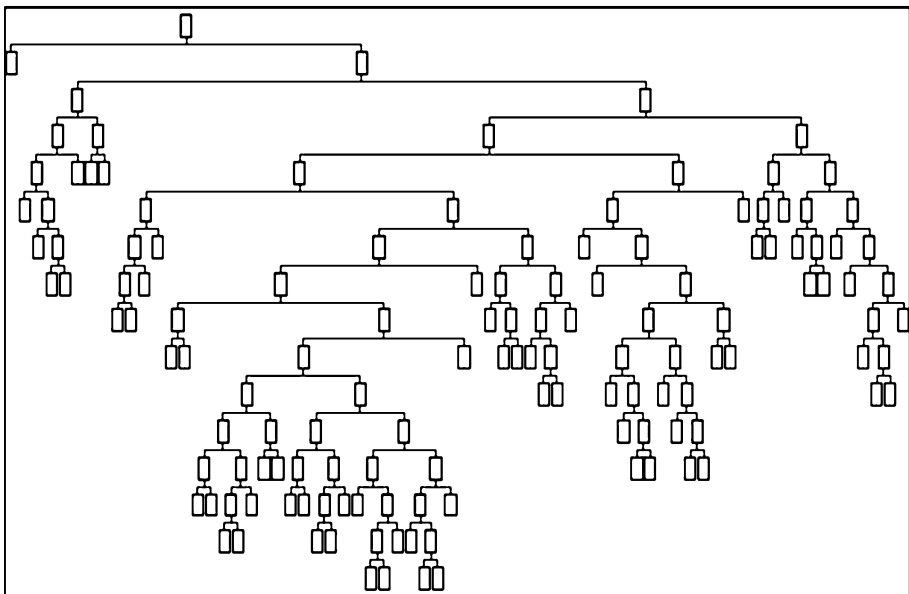
Fig. 8 Maximal classification tree

zero congruence was obtained for this class. The significantly higher estimation accuracy can be observed with the new classification than with the Polish classification point method (Fig. 12).

The evaluation of damage risk was made only for hits of each damage class. When assessing the damage class, it is also very important to know the number of damage risk overestimations and underestimations. The misclassification matrix is also used for assessing the cost of risk (Table 8).

Table 7 Number of nodes in classification tree versus cross-validation value

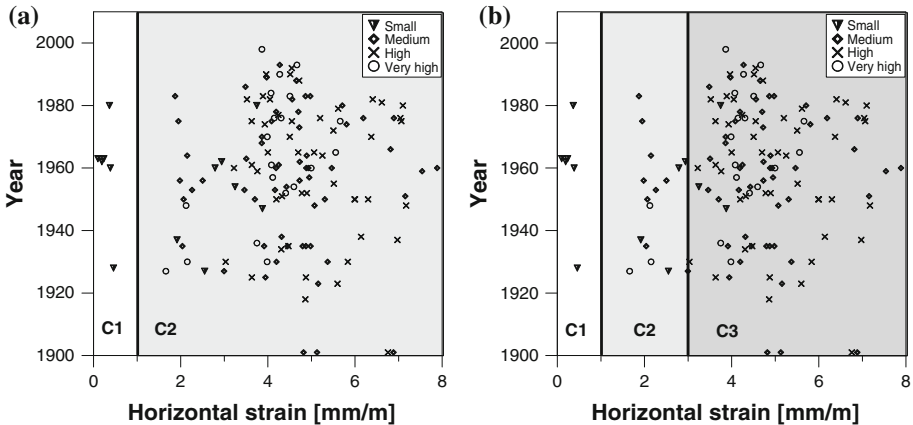| Number of terminal nodes $\lvert \widetilde{T} \rvert$	Cost of misclassifications $R(T)$	Standard deviation
135	0.6489	0.041704
91	0.6641	0.041265
19	0.5878	0.043007
15	0.6288	0.042051

Fig. 9 Classification divisions for the first (**a**) and third node (**b**)

The spatial juxtaposition of estimation results with the new method clearly reveals a tendency to overestimate the damage class with respect to the real values. Overestimation was observed particularly for low class damage (C1 and C2). The estimation of severe damage is highly accurate. The comparison of the results of damage evaluation with the new method and the Polish method reveals that the latter method leads to considerable underestimation of analyzed values. This has a great significance especially for severely damaged buildings. According to evaluations made with the Polish method, buildings which underwent severe damage C3 were expected to be subject of small and very small damage. The evaluation of both methods in view of cost of misclassification shows that considerably bigger risks and costs will be generated due to damage risk underestimations than overestimations. Accordingly, the Polish method has very low prediction efficiency as compared to the counterpart method.

The proposed method is simpler, and before all, more efficient than the Polish point method.

4 Summing up

This research focuses on a procedure employed for adapting the existing and inefficient methods of damage risk evaluation to local conditions. The proposed methodics can be quickly implemented to any densely developed area staying under the influence of continuous deformations. One of the most important assumptions is the access to data, on the
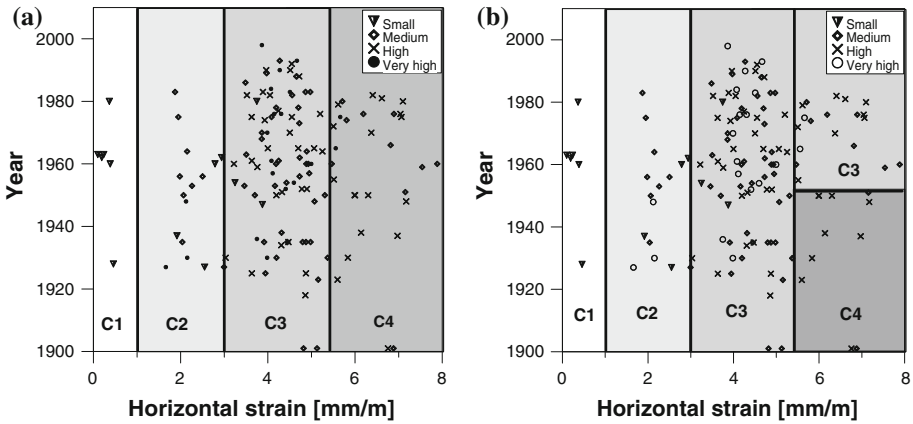
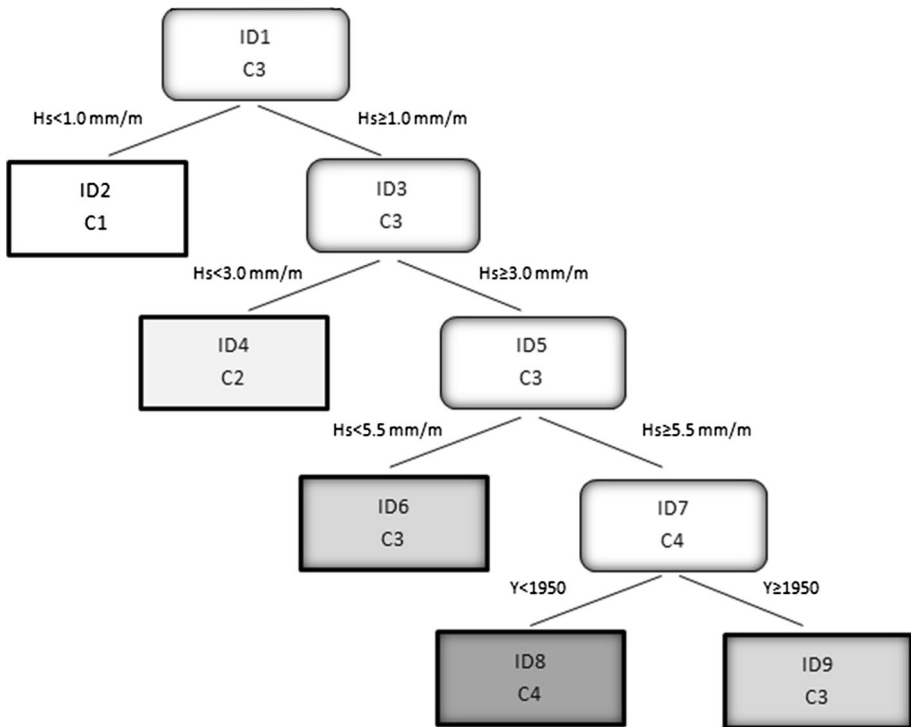**Fig. 10** Classification divisions in the fifth and seventh nodes

Fig. 11 Ultimate classification tree

basis of which the model can be adapted. Therefore, the analyzed areas should not have been subjected to continuous deformation before. In the case of highly invested areas where no cases of terrain deformation were observed, no local data for modeling can be acquired. Alternatively, attempts can be made to collect data from areas where buildings have a similar vulnerability and construction characteristic. Moreover, those buildings
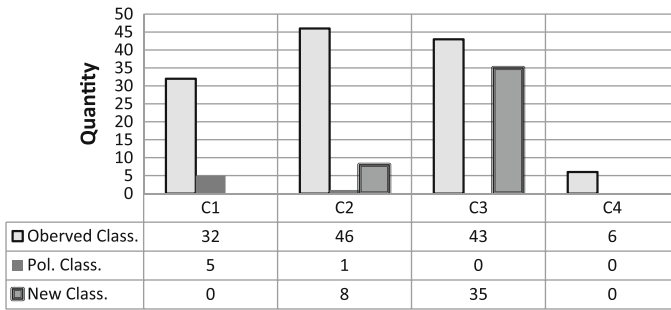
	C1	C2	C3	C4
□ Oberved Class.	32	46	43	6
■ Pol. Class.	5	1	0	0
▣ New Class.	0	8	35	0

Fig. 12 Comparison of the accuracy of developed method

Table 8 Decision cost matrix new method (a) and Polish method (b)

	Observed damage class			
	C1	C2	C3	C4
(a) Predicted damage class new method				
C0	0	0	0	0
C1	0	0	2	0
C2	5	8	4	1
C3	26	36	35	5
C4	1	2	2	0
(b) Predicted damage class Polish method				
C0	27	36	32	0
C1	5	8	9	3
C2	0	1	2	1
C3	0	1	0	0
C4	0	0	0	0

should undergo damage under the impact of surface deformation of representative intensity. This type of set can be used for classification, thanks to which boundary criteria of damage class can be distinguished. The advantage of this solution lies in its high adaptability. Low reliability of the point method results from, e.g., qualitative diversity of regions (buildings and exploitation), for which it is applied. Another thing is the subjective approach of researchers who use this method and their professional competences.

The presented algorithm can be used for determining boundary criteria of building damage class in any mining activity area. The new method might be an alternative to currently used method for the building damage estimation on the mining areas. Presented solution shall contribute to an incensement of the accuracy of the results of building damage risk evaluation. The methodic presented in this paper is also applicable to areas under the impact of continuous deformations generated by other underground activity than hard coal mining. A similar classification could be presented, e.g., for fluids or water production, materials which create considerable hazard for intensely developed areas.

It should be emphasized that the presented graph is adequate to areas where it was implemented (or areas of similar characteristic). Owing to its simplicity, the presented method can be used for an arbitrarily selected area.

References

Askan A, Yucemen MS (2010) Probabilistic methods for the estimation of potential seismic damage: application to reinforced concrete buildings in Turkey. Struct Saf 32(4):262–271

Bhattacharya S, Singh MM (1984) Proposed criteria for subsidence damage to buildings. Rock mechanics in productivity and protection. In: 25th symposium on rock mechanics

Boscarding MD, Cording EG (1989) Building response to excavation-induced settlement. J Geotech Eng 115(1):1–21

Breiman L, Friedman JH, Olshen RA, Stone CJ (1984) Classification and regression trees. Wadsworth International Group, Belmont

Burland JB, Wroth CP (1975) Settlement of building and associated damage. Building Research Establishment Current Paper, Watford

Chang T-S (2011) A comparative study of artificial neural networks, and decision trees for digital game content stocks price prediction. Expert Syst Appl 38(12):14846–14851

Chudzinska M, Baralkiewicz B (2011) Application of ICP-MS method of determination of 15 elements in honey with chemometric approach for the verification of their authenticity. Food Chem Toxicol 49(11): 2741–2749

GIG (2000) Instrukcja No. 12: Zasady oceny możliwości prowadzenia podziemnej eksploatacji z uwagi na ochronę obiektów budowlanych. Wydawnictwo GIG, Katowice

Grünthal G, Wahlström R (2006) New generation of probabilistic seismic hazard assessment for the area Cologne/Aachen considering the uncertainties of the input data. Nat Hazards 38:159–176

Grünthal G, Thieken AH, Schwarz J, Radtke K, Smolka A, Merz B (2006) Comparative risk assessment for the city of Cologne, Germany—storms, floods, earthquakes. Nat Hazards 38(1–2):21–44. doi:10.1007/s11069-005-8598-0

Hazir MHM, Shariff ARM, Amiruddin MD, Ramli AR, IqbalSaripan M (2012) Oil palm bunch ripeness classification using fluorescence technique. J Food Eng. doi:10.1016/j.jfoodeng.2012.07.008

Knothe S (1953) Równanie profilu ostatecznie wykształconej niecki osiadania. Archiwum Górnictwa i Hutnictwa, t.1, z.1, Warszawa

Lewis RJ (2000) An introduction to classification and regression tree (CART) analysis. In: Presented at the 2000 annual meeting of the society for academic emergency medicine in San Francisco

Malinowska A (2013a) Analysis of methods used for assessing damage risk of buildings under the influence of underground exploitation in the light of World's experience: part 1. Arch Min Sci 58(3):843–853. doi:10.2478/amsc-2013-0058

Malinowska A (2013b) Accuracy estimation of the approximated methods used for assessing risk of buildings damage under the influence of underground exploitation in the light of World's and Polish experience: part 2. Arch Min Sci 58(3):855–865. doi:10.2478/amsc-2013-0059

Mine Subsidence Engineering Consultants (2007) Mine subsidence damage to building structure. www.minesubsidence.com

National Coal Board (1975) Subsidence engineer's handbook. National Coal Board, London

Pei S, van de Lindt JW (2009) Methodology for earthquake-induced loss estimation: an application to woodframe building. Struct Saf 31(1):31–42

Przybyła H, Świądrowski W (1968) Określenie kategorii odporności istniejących obiektów budownictwa powszechnego na wpływy eksploatacji górniczej, Ochrona Terenów Górniczych

Rothwell JJ, Futter MN, Dise NB (2008) A classification and regression tree model of controls on dissolved inorganic nitrogen leaching from European forests. Environ Pollut. doi:10.1016/j.envpol.2008.01.007

Smeti EM, Thanasoulias NC, Lytras ES, Tzoumerkas PC, Golfinopoulosc SK (2009) Treated water quality assurance and description of distribution networks by multivariate chemometrics. Water Res 43(18):4676–4684

Straub D, Kiureghian AD (2008) Improved seismic fragility modeling from empirical data. Struct Saf 30:320–336

Ture M, Tokatli F, Kurt I (2009) Using Kaplan–Meier analysis together with decision tree methods (C&Rt, Chaid, Quest, C4.5 and Id3) in determining recurrence-free survival of breast cancer patients, Part 1. Expert Syst Appl 36(2):2017–2026. doi:10.1016/j.eswa.2007.12.002

Yang C-C, Prasher SO, Enright P, Madramootoo Ch, Burgess M, GoelP K, Callum I (2003) Application of decision tree technology for image classification using remote sensing data. Agric Syst 76:1101–1117

Yong L (2006) Predicting materials properties and behavior using classification and regression trees. Mater Sci Eng A 433:261–268. doi:10.1016/j.msea.2006.06.100

Yu Z, Karmis M, Jarosz A, Haycocks C (1988) Development of damage criteria for buildings affected by mining subsidence. In: 6th annual workshop generic mineral technology centre mine system design and ground control

Yucemen MS, Ozcebe G, Pay AC (2004) Prediction of potential damage due to severe earthquakes. Struct Saf 26(3):349–366