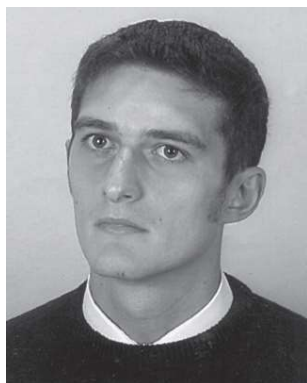


Wybrane metody eksploracji danych i uczenia maszynowego w analizie stanu uszkodzeń oraz zużycia technicznego zabudowy terenów górniczych

Selected methods of data mining and machine learning in risk analysis for developments located in mining areas



Dr inż. Karol Firek^{*)}



Dr inż. Janusz Rusek^{*)}



Prof. dr hab. inż. Aleksander Wodyński^{*)}

Treść: W referacie przedstawiono metodykę oraz wyniki badań wpływu oddziaływań eksploatacji górniczej na zabudowę powierzchni, które zostały przeprowadzone w ostatnich latach w Katedrze Geodezji Inżynierskiej i Budownictwa AGH. Obejmowały one modelowanie przebiegu zużycia technicznego budynków metodami *uczenia maszynowego* oraz analizę zakresu i intensywności ich uszkodzeń z zastosowaniem metod *eksploracji danych*. Uzyskane wyniki potwierdzają przydatność zastosowanych metod do rozwiązywania zagadnień związanych z budownictwem na terenach górniczych.

Abstract: This paper presents the methodology and results of the studies on the influence of mining impacts on developments located in mining areas, which have been performed in recent years at the Department of Engineering Surveying and Civil Engineering of AGH University of Science and Technology. The studies included modeling the course of technical wear of buildings, by the methods of machine learning, as well as the analysis of the scope and intensity of their damage with the methods of data mining. The obtained results confirm the usefulness of the methods to solve the issues related to construction in mining areas.

Słowa kluczowe:

uczenie się maszynowe, wydobywanie danych, techniczne zużycie budynku, uszkodzenie budynku, wpływ eksploatacji

Key words:

machine learning, data mining, technical wear of building, damage of building, mining effects

1. Wprowadzenie

Zużycie techniczne jest miarą stanu technicznego budynku. Dominujący wpływ na stopień zużycia obiektu ma zazwyczaj zużycie naturalne, związane z procesem starzenia się materiałów budowlanych w określonych warunkach środowiskowych i eksploatacyjnych. Trudności z ustaleniem stopnia zużycia wynikają z faktu, iż obok specyfikacji konstrukcyjno-materiałowej uwzględnić należy cechy trudne do ilościowego określenia, takie jak jakość materiałów i wykonawstwa, gospodarka remontowa oraz sposób eksploatacji obiektu. Z drugiej strony na zużycie techniczne wpływa szereg dodatkowych czynników, w sensie statystycznym losowych, do których zaliczyć należy również oddziaływania eksploatacji górniczej.

W literaturze przedmiotu podano wiele sposobów pozwalających na oszacowanie wartości zużycia technicznego (np. [15]). Stosunkowo dokładną metodą jest tzw. metoda średniej ważonej

$$s_z = \sum_{i=1}^n \frac{u_i \cdot s_{ei}}{100}, \% \quad (1.1)$$

gdzie:

- u_i – procentowy udział kosztu odtworzenia danego elementu w strukturze kosztu odtworzenia obiektu, pełniący rolę wagi,
- s_{ei} – stopień zużycia danego elementu, w %,
- n – liczba ocenianych elementów w obiekcie.

Wyznaczanie zużycia technicznego metodą średniej ważonej odbywa się *post factum*, co stanowi poważne utrudnienie w jej zastosowaniu do prognozowania skutków oddziaływań górniczych na zużycie techniczne budynków. Trudności te spowodowały, iż podjęto próbę budowy takiego modelu przebiegu zużycia technicznego budynków, którego dziedzinę stanowią potencjalne przyczyny (wskaźniki deformacji lub charakterystyki wstrząsów górniczych), a nie skutki w poszczególnych elementach budynku, jak to ma miejsce w metodzie średniej ważonej. Wyniki badań przeprowadzonych w tym zakresie przedstawiono w rozdziale 2.

^{*)} AGH w Krakowie

Na stan techniczny budynku, oprócz zużycia naturalnego, wpływa intensywność, czyli zakres i częstość uszkodzeń wszystkich jego elementów, zarówno konstrukcyjnych, wykończeniowych, jak i wyposażenia (instalacji). Wspólną cechą tych uszkodzeń jest fakt, iż przyczyn ich powstania szukać należy głównie w czynnikach losowych, takich jak skutki błędów projektowych i wykonawczych, przeciążenia elementów konstrukcyjnych, drgania komunikacyjne lub od maszyn i urządzeń, zmiany stosunków wodnych w podłożu fundamentowym, ruchy masowe, itp. Należą do nich również wpływy górnicze, zarówno w postaci wstrząsów, jak i ciągłych deformacji powierzchni.

Na podstawie prowadzonych w ostatnich latach analiz przedstawiono propozycję klasyfikacji intensywności uszkodzeń poszczególnych elementów budynków o konstrukcji murowanej, jak również budynków wzniesionych w systemach uprzemysłowionych (wielki blok, wielka płyta) oraz budynków typu halowego. W rozdziale 3 przedstawiono wyniki dotychczasowych badań przeprowadzonych w oparciu o wskaźnik intensywności uszkodzeń ustalonych dla poszczególnych typów budynków oraz zasygnalizowano zakres badań aktualnie prowadzonych w tym obszarze.

2. Modelowanie przebiegu zużycia technicznego budynków na terenach górniczych metodami uczenia maszynowego (*machine learning - ML*)

Badania prowadzone przed kilku laty w Katedrze Geodezji Inżynierskiej i Budownictwa AGH wykazały istotny w sensie statystycznym wpływ na zużycie techniczne budynków zarówno deformacji ciągłych, jak i wstrząsów górniczych (np. [15]). W badaniach tych stosowano klasyczne metody statystyki matematycznej. Poszerzenie analiz o modele regresji nieliniowej pozwoliło wnioskować o występowaniu nieliniowych związków pomiędzy czynnikami górniczymi a przebiegiem stopnia zużycia technicznego budynków [3]. Dało to asumpt do podjęcia badań w celu ustalenia wielowymiarowych relacji wiążących przyczyny ze zużyciem technicznym.

Biorąc również pod uwagę niepewność co do wartości szacowanych w czasie inwentaryzacji czynników (niekiedy ustalanych w oparciu o niejednoznaczne sformułowania lingwistyczne typu: niski, znaczny, zadowalający), uznano, że najefektywniejszym rozwiązaniem będzie stworzenie modelu, którego działanie określone zostanie przez reguły zapis przyczynowo-skutkowy z uwzględnieniem niepewności (np. [11], [12]). Rozpoznanie literaturowe pozwoliło na zawężenie dziedziny poszukiwań do rodziny metod zaliczanych do tzw. *uczenia maszynowego (machine learning - ML)*.

Metody *ML* to w ogólności algorytmy numeryczne pozwalające na budowę złożonych systemów w procesie adaptacji (uczenia się). Każdy system jest skonstruowany w celu rozwiązywania postawionego problemu. Proces adaptacji pozwala na stworzenie struktury wiedzy oraz efektywniejsze wykonywanie danego zadania na nowych danych, niejednokrotnie różnych od danych uczących. Dlatego

przyjmuje się, iż możliwość uogólniania nabytej wiedzy w procesie uczenia jest cechą nadrzędną przy budowie systemów należących do tej grupy.

Do metod *ML* można zaliczyć m.in. sztuczne sieci neuronowe, adaptacyjne systemy wnioskowania rozmytego, metodę wektorów podpierających (SVM), a także sieci przekonania Bayesa.

2.1. Metoda SVM w ujęciu regresyjnym ε -SVR

Metoda *SVM (Support Vector Machine)* jest uszczegółowieniem sztucznych sieci neuronowych z radialnymi funkcjami bazowymi (np. [1], [10], [14]). W odróżnieniu od sieci neuronowych metoda budowy takiego systemu jest ukierunkowana przede wszystkim na osiągnięcie wysokiego poziomu generalizacji. Poziom uogólniania jest charakteryzowany przez szerokość tzw. pasma tolerancji, które jest jednym z parametrów poddawanych strojeniu w procesie adaptacji. W zależności od przeznaczenia takiego systemu oraz rodzaju przyjętych funkcji jąder, strojeniu podlegają: dwa parametry dla zadań klasyfikacji oraz trzy parametry dla zadań regresji [8]. Z uwagi na relatywnie małą liczbę parametrów proces uczenia może być prowadzony przy użyciu efektywnych metod optymalizacji.

Podstawę badań stanowiła baza danych, w której zebrano informacje o konstrukcji i stanie technicznym 1726 budynków zlokalizowanych na obszarze Górnośląskiego Zagłębia Węglowego (820 obiektów) i Legnicko-Głogowskiego Okręgu Miedziowego (906 obiektów). Zebrano również szczegółowe dane o potencjalnych czynnikach wpływających na ich zużycie techniczne, w tym o oddziaływaniach górniczych.

W omawianych badaniach przyjęto radialne funkcje jąder systemu oraz metodę optymalizacji parametrów wykorzystującą algorytm genetyczny jako bezgradientową metodę optymalizacji [13]. Następnie utworzony model poddano analizie wrażliwości na drodze wielokrotnych symulacji mających na celu ocenę monotoniczności uzyskanej wielowymiarowej hiperpowierzchni realizującej przebieg stopnia zużycia technicznego budynków. W rezultacie uzyskano model przebiegu zużycia technicznego budynków oraz określenie stopnia wpływu poszczególnych czynników górniczych i ogólnobudowlanych na kształtowanie się aproksymowanej powierzchni. W szczególności wyznaczono wkłady poszczególnych czynników górniczych i budowlanych na przebieg modelowanego stopnia zużycia technicznego budynków (tab. 1). Miara M^s oddaje tempo z jakim, względem poszczególnych czynników, aproksymowana funkcja wznosi się lub opada.

Z przedstawionych w tabeli 1. danych wynika, że spośród analizowanych czynników górniczych największy wpływ na wzrost zużycia technicznego badanej grupy budynków ma zmienna opisująca poziome odkształcenia rozciągające. Wykazano również wpływ zabezpieczeń profilaktycznych budynku, które powodują obniżenie jego wartości stopnia zużycia postępujące wraz ze wzrostem zakresu tych zabezpieczeń.

Tak przeprowadzone badania pozwoliły na zidentyfikowanie wkładu, zarówno ilościowego, jak i jakościowego,

Tabela 1. Zestawienie wartości miar wrażliwości M^s dla grupy budynków z obszarów GZW i LGOM wg [14]

Table 1. Summary of values of sensitivity measures M^s for the group of buildings in the Upper Silesian Coal Basin and Legnica-Głogów Copper Belt area, acc. to [13]

Miary wrażliwości M^s dla poszczególnych konfiguracji zmiennych	Wiek	Poziome odkształcenia rozciągające ε^+	Poziome odkształcenia ściskające ε^-	Poziome odkształcenia na kierunkach głównych ε^{\max}	Wskaźnik zabezpieczeń
	+0,581	+0,219			-0,040
	+0,495		+0,133		-0,010
	+0,513			+0,153	-0,013

z jakim na przebieg zjawiska oddziałują czynniki górnicze i budowlane. Dodatkowo, dysponując modelem oraz wynikami dotyczącymi wpływu poszczególnych zmiennych na przebieg modelowanego procesu, możliwe było przeniesienie uzyskanych informacji na grunt formalizmów logiki rozmytej i rozmytego wnioskowania (por. rozdz. 2.2).

Podobną analizę przeprowadzono również dla grupy budynków o konstrukcji wielkopłytywowej [16].

2.2. Metody bazujące na logice rozmytej

Zastosowanie formalizmów logiki rozmytej pozwala na budowę systemu decyzyjnego postaci „if then” ujmującego niepewność w stosunku do danych wejściowych [8], [9]. Takie ujęcie problemu pozwala na zapis reguł decyzyjnych z uwzględnieniem tolerancji na ograniczony zasób informacji wynikających z danych i umożliwia wykorzystanie oszacowań podawanych dla poszczególnych zmiennych w formie rozmytych kategorii, którymi mogą być np. sformułowania lingwistyczne.

W omawianych badaniach bazą informacji dla budowy systemu rozmytego był wstępnie utworzony model stopnia zużycia technicznego, wykorzystujący metodę SVM w podejściu regresyjnym (rozdz. 2.1). W kolejnych etapach przyjętej procedury budowy modelu wykorzystano systemy wnioskowania rozmytego: adaptacyjny system *Wanga-Mendela* oraz ekspertowy system *Mamdani* [13].

Ustalanie reguł dla systemu wnioskowania rozmytego przebiegało zgodnie ze schematem podanym na rys. 1.

Zastosowany 3-etapowy program badań umożliwił budowę systemu wnioskowania rozmytego *Mamdani* jako modelu przebiegu zużycia technicznego budynków na terenach górniczych oraz pozwolił na ocenę wpływu poszczególnych czynników warunkujących jego przebieg.

Ostatecznie uzyskano system, którego działanie dla przykładowego zestawu zmiennych wejściowych przedstawiono na rys. 2 (np. [12]). Schematy a), b), c) i d) na tym rysunku pokazują ustalone na drodze przyjętej procedury budowy modelu, zbiory rozmyte dla zmiennych wejściowych wraz ze stopniami zapłonu przy projekcji danych wartości liczbowych.

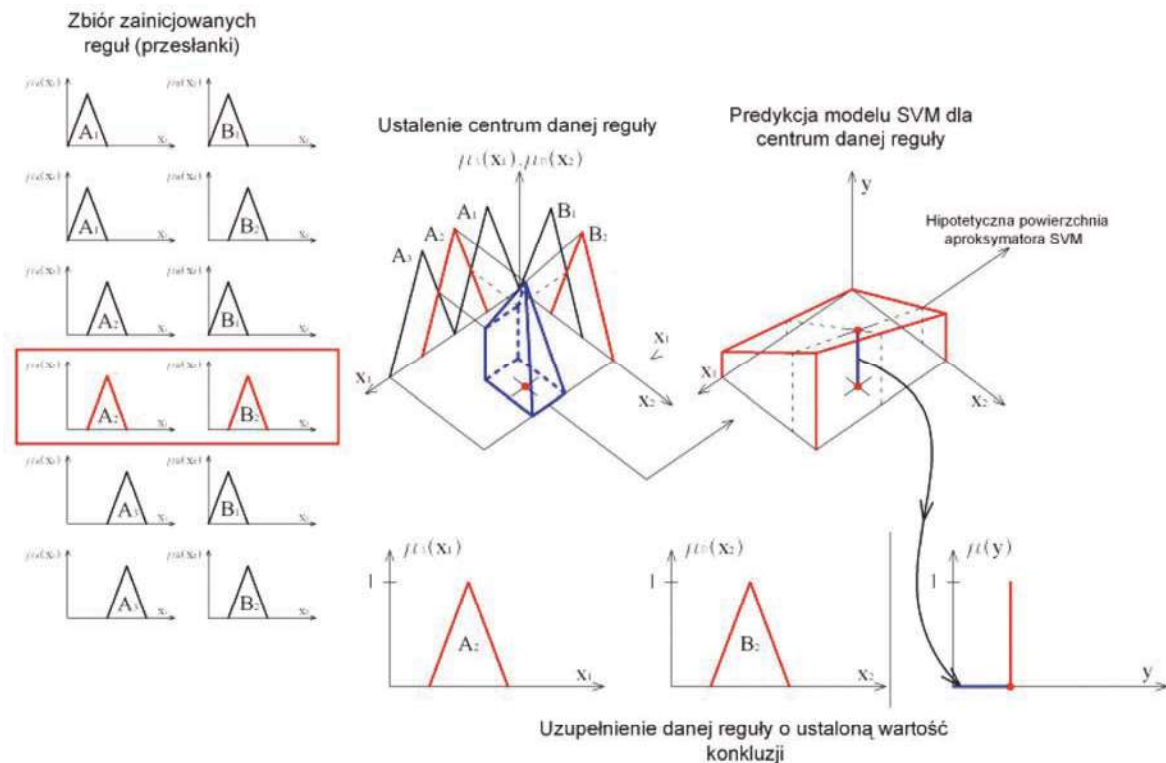
Z kolei wynikowy schemat e) przedstawia wynikowe połączenie zaktwowanych zbiorów rozmytych oraz wartość stopnia zużycia ($s_z = 44,5\%$) odpowiadającą położeniu środka ciężkości dla tego zbioru.

Przeprowadzone badania pozwoliły na budowę regulowanego systemu wnioskowania rozmytego, który oprócz aproksymacji wartości stopnia zużycia technicznego budynku pełni funkcję klasyfikatora stanu technicznego podawanego w niecisłej postaci sformułowań lingwistycznych [13]. Powyższa procedura może również wspomagać rozwiązanie problemu przewidywania stopnia uszkodzeń proponowanego w pracy [4].

3. Badanie zakresu i intensywności uszkodzeń budynków

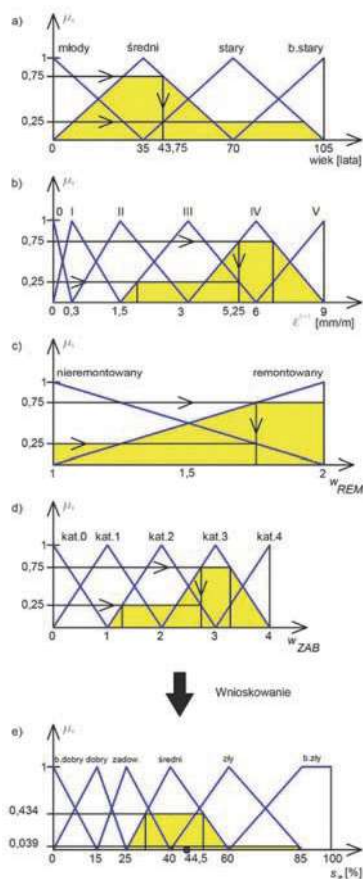
Na stan techniczny budynku, oprócz zużycia naturalnego, wpływa intensywność, czyli zakres i częstość uszkodzeń wszystkich jego elementów.

Podjęcie próby określenia uniwersalnego wskaźnika intensywności uszkodzeń dla budynków wznoszonych w różnych technologiach wynikało z faktu, iż: może on służyć do sprawnej i syntetycznej oceny zakresu i intensywności uszkodzeń licznych grup budynków w ramach inwentaryzacji ich stanu technicznego, w tym również do porównań intensywności uszkodzeń grup budynków i ich elementów. Wskaźnik ten może również posłużyć do analizy relacji między obserwowanym zakresem uszkodzeń budynków



Rys. 1. Schemat ustalania konkluzji dla systemu rozmytego *Wanga-Mendela* z wykorzystaniem wspomagającego modelu ϵ -SVR wg [14]

Fig. 1. Scheme for determining conclusions for the *Wang-Mendel* fuzzy system using the supporting model ϵ -SVR according to [13]



Rys. 2. Poziomy aktywacji zbiorów rozmytych dla zmiennych wejściowych i aproksymowanej zmiennej wyjściowej systemu wg [12]

Fig. 2. Level of activation of fuzzy sets for input variables and the approximated output variable of the system, acc. [12]

a oddziaływaniami eksploatacji górniczej. W efekcie może być wykorzystany do prognozowania kosztów związanych z usuwaniem szkód górniczych.

W pracy [4], a następnie [5], [17] i [18] przedstawiono propozycję klasyfikacji intensywności uszkodzeń poszczególnych elementów budynków o konstrukcji murowanej, wykonanych w systemach przemysłowych (wielki blok, wielka płyta), a także obiektów szkieletowych typu halowego. Przy ustalaniu szczegółowych kryteriów oceny intensywności uszkodzeń dla poszczególnych elementów budynku korzystano z własnych doświadczeń oraz ustaleń zawartych w piśmiennictwie traktującym o diagnostyce stanu technicznego budynków.

3.1. Metodyka ustalania wskaźnika intensywności uszkodzeń elementów budynków

Do oceny zakresu i częstości występowania uszkodzeń poszczególnych elementów konstrukcyjnych i niekonstrukcyjnych budynku zaproponowany został wskaźnik w_{ui} zdefiniowany w 6-stopniowej skali (tab. 2 i 3). Intensywność uszkodzenia poszczególnych elementów każdego z badanych budynków została ustalona według szczegółowych kryteriów zestawionych w pracach [4], [5], [17], [18].

Przy określaniu szczegółowych kryteriów oceny intensywności uszkodzeń dla poszczególnych elementów korzystano z literatury traktującej o diagnostyce stanu technicznego

budynków. W przyjętej skali, wskaźnik $w_u = 0$ oznacza, że uszkodzenia nie występują, $w_u = 1$ – nieznaczne uszkodzenia, $w_u = 2$ – umiarkowane uszkodzenia, itd. W praktyce zwykle znajdujemy się w zakresie wskaźnika w_u od 0 do 3. W niektórych przypadkach, przy intensywnych uszkodzeniach elementów konstrukcyjnych lub bardzo intensywnych uszkodzeniach elementów drugorzędnych może wystąpić konieczność wykonania szczegółowej ekspertyzy w celu sprawdzenia bezpieczeństwa budynku.

Tabela 2. Definicja wskaźnika intensywności uszkodzeń w_u wg [4]
Table 2. Definition of the index of damage intensity w_u , acc. [4]

Wskaźnik intensywności uszkodzeń w_u	Definicja	Opis	Zakres [%]
0	Nie występują	uszkodzenia nie występują lub są niezauważalne	0
1	Nieznaczne	uszkodzenia mało znaczące, drobne, występujące pojedynczo	(0-10)
2	Umiarkowane	uszkodzenia umiarkowane, występujące lokalnie, miejscami	(10-30)
3	Intensywne	uszkodzenia intensywne, znaczne, rozległe, występujące lokalnie lub licznie	(30-50)
4 (i 5)	Bardzo intensywne	uszkodzenia bardzo intensywne, znaczne, rozległe, występujące bardzo licznie (aż do zniszczenia)	>50

Tabela 3. Wskaźniki intensywności uszkodzeń elementów budynku
Table 3. Indexes of intensity of damage of the building

Oznaczenie	Opis wskaźnika
Elementy stroju nośnego	
w_{u1}	intensywność uszkodzeń fundamentu
w_{u2}	intensywność uszkodzeń ścian nośnych piwnic lub fundamentowych
w_{u3}	intensywność uszkodzeń ścian nośnych nadziemna wewn. i zewn. (w tym nadproży i murów podokiennych)
w_{u4}	intensywność uszkodzeń murów ogniowych
w_{u5}	intensywność uszkodzeń elementów szkieletowego stroju nośnego (słupy, rygle)
w_{u6}	intensywność uszkodzeń stropów nad piwnicami
w_{u7}	intensywność uszkodzeń stropów wyżej, stropodachu (przekrycia)
w_{u8}	intensywność uszkodzeń schodów
w_{u9}	intensywność uszkodzeń balkonów i loggii (oraz okapów i gzymsów)
w_{u10}	intensywność uszkodzeń konstrukcji dachu
Elementy drugorzędne (wykończeniowe)	
w_{u11}	intensywność uszkodzeń ścian działowych
w_{u12}	intensywność uszkodzeń tynków wewnętrznych i okładzin ściennych
w_{u13}	intensywność uszkodzeń podłogi (warstw podłogowych)
w_{u14}	intensywność uszkodzeń ścian kominowych
w_{u15}	intensywność uszkodzeń stężeń (elementów zapewniających sztywność przestrzenną)
w_{u16}	intensywność uszkodzeń ścian wypełniających (lub osłonowych)
w_{u17}	intensywność uszkodzeń elewacji (warstw elewacyjnych)
w_{u18}	intensywność uszkodzeń izolacji przeciwwilgociowych
w_{u19}	intensywność uszkodzeń pokrycia dachowego
w_{u20}	intensywność uszkodzeń obróbek blacharskich, rynien i rur spustowych
w_{u21}	intensywność uszkodzeń stolarki
w_{u22}	intensywność uszkodzeń elementów zewnętrznych (wejścia do budynków oraz podesty, tarasy opaski itp.)

3.2. Analiza intensywności uszkodzeń metodami eksploracji danych (*data mining* - DM)

Zgromadzone w Katedrze Geodezji i Budownictwa AGH bazy danych o stanie technicznym i potencjalnych przyczynach uszkodzeń budynków zlokalizowanych na terenach górniczych zostały rozszerzone o wskaźniki intensywności uszkodzeń ustalone dla poszczególnych elementów budynków. Wstępne analizy przeprowadzone metodami klasycznej statystyki pozwoliły m.in. na porównanie intensywności uszkodzeń pomiędzy elementami budynków wyodrębnionych z uwagi na technologię budowy lub poziom oddziaływań górniczych. Ponadto dokonano próby oceny wkładu wskaźników uszkodzeń poszczególnych elementów budynku w jego stopień zużycia [5]. W oparciu o wyniki wstępnych analiz, aktualnie prowadzone są badania metodami eksploracji danych (*data mining* - DM).

Metody DM polegają na ekstrakcji wiedzy z rozległych zasobów baz danych. Pozwala ona między innymi na wyłonienie asocjacji między zmiennymi, redukcji wymiarowości danych itd. Celem prowadzonych badań była optymalizacja wskaźnika intensywności uszkodzeń dla budynków zlokalizowanych na terenach górniczych oraz badanie jego części składowych w aspekcie bezpieczeństwa i uciążliwości użytkowania. Wykorzystano przy tym:

- metodę *CART* (*Classification and Regression Tree*) oraz *losowy las* (*Random Forest*) (np. [2], [7]),
- metodę *PCA* (*Principal Component Analysis*) oraz *PLSR* (*Partial Least Squares Regression*) - np. [6], [7].

Ad. a) **Drzewa decyzyjne** są uznana metodą eksploracji danych, pozwalającą na rozwiązywanie problemów zarówno klasyfikacyjnych, jak i regresyjnych. Oprócz aproksymacji funkcji wielu zmiennych, metoda ta umożliwia analizę struktury utworzonego systemu oraz ocenę udziału, tj. istotności poszczególnych zmiennych wejściowych. Ponadto drzewa decyzyjne w podejściu regresyjnym, w odróżnieniu od np. tradycyjnej metody regresji wielorakiej, pozwalają na przedstawienie przebiegu aproksymowanej funkcji z uwzględnieniem jej zmienności lokalnej, realizując aproksymację w wydzielonych partycjach przestrzeni zmiennych wejściowych.

Każde drzewo decyzyjne (klasyfikacyjne bądź regresyjne) ma postać spójnego grafu acyklicznego, który reprezentuje proces podziału zbioru uczącego na jednorodne podzbiory. W korzeniu takiego drzewa skupiony jest cały zbiór danych uczących, a pozostałymi elementami są węzły wewnętrzne (miejsca, gdzie dokonywany jest podział danego podzbiory) oraz liście (miejsca końcowe, w których podział kolejno następujących po sobie podzbiorów już nie zachodzi).

Zaletą metody drzew decyzyjnych jest to, że mimo niejawniej reprezentacji ostatecznego podziału przestrzeni wielowymiarowej, struktura drzewa jest w pełni przejrzysta i pozwala na interpretację powiązań przyczynowo-skutkowych w opisie modelu. Dodatkowo, zarówno procedura budowy drzew decyzyjnych, jak również analiza struktury istniejącego drzewa pozwala ustalić wkład poszczególnych zmiennych wejściowych na wartość modelowanego procesu (np. [2]).

W toku przeprowadzonych badań [17] wykonano analizy z wykorzystaniem metody *CART* (*Classification and Regression Tree*) oraz *losowego lasu* (*Random Forest*). Pozwoliło to na systematyzowanie wkładów poszczególnych wskaźników uszkodzeń pod kątem ich znaczenia dla opisu stopnia zużycia technicznego grupy 94 budynków typu halowego usytuowanych na terenie górniczym (tab. 3).

Uwzględniając specyfikę budynków halowych, w badaniach ustalono łączne wskaźniki uszkodzeń wewnętrznych elementów wykończeniowych (ścian działowych, tynków oraz okładzin ściennych i podłogowych - $w_{u11,12,13}$) oraz izo-

lacji przeciwwilgociowych, pokrycia dachowego, obróbek blacharskich, rynien i rur spustowych ($w_{u18,19,20}$).

Tabela 3. Rangi wskaźników intensywności uszkodzeń poszczególnych elementów w stopniu zużycia technicznego budynku wg [17]

Table 3. Ranks of indexes of intensity of damage of the particular elements in the degree of technical wear of the building, acc. to [17]

Metoda	Rangi wskaźników intensywności uszkodzeń							
	w_{u5}	w_{u7}	w_{u15}	w_{u16}	w_{u17}	$w_{u11,12,13}$	$w_{u18,19,20}$	w_{u22}
<i>CART</i>	6	4	8	2	1	3	5	7
<i>losowych lasów</i>	5	6	8	1	3	2	4	7

Z tabeli 3. wynika, że największy udział w zużyciu technicznym budynków stwierdzono w przypadku uszkodzeń ścian wypełniających i osłonowych, warstw elewacyjnych oraz wewnętrznych elementów wykończeniowych, reprezentowanych w analizie przez wskaźniki w_{u16} , w_{u17} i $w_{u11,12,13}$.

Z kolei najmniej znaczący wkład w stopniu zużycia mają uszkodzenia elementów stężających (w_{u15}) oraz zewnętrznych (w_{u22}).

Ad. b) **Metoda PCA** (*Principal Component Analysis*) polega na wyznaczeniu wartości i wektorów własnych macierzy kowariancji. Tym samym otrzymuje się nowe, nieskorelowane wzajemnie kombinacje liniowe zmiennych wejściowych. Dla każdej z nowych zmiennych należy dążyć do możliwie największej wariancji [9].

Z kolei w **metodzie PLSR** (*Partial Least Squares Regression*) wyłanianie wektorów własnych w przestrzeni zmiennych wejściowych jest jednym z etapów budowy wielowymiarowego modelu regresji. Procedura ekstrakcji wartości własnych przebiega jednak odmiennie niż w przypadku *PCA*. W metodzie *PLSR* każdy etap iteracyjnej procedury wyłaniania składowych głównych jest nierozdzielnie związany ze zmienną zależną, którą w ramach badań był stopień zużycia technicznego budynków. Fakt ten jest bardzo istotny w sytuacji, gdy celem jest wyłonienie uogólnionego czynnika w postaci najlepiej korespondującej ze zmienną zależną.

W efekcie dotychczas przeprowadzonych analiz, których bazę stanowiły dane o 129 budynkach o konstrukcji wieloblokowej stwierdzono, że metoda *PLSR* jest bardziej efektywna na poziomie wstępnej analizy danych niż metoda *PCA* [18]. Jest to efektem iteracyjnego wyodrębniania składowych głównych w przestrzeni zmiennych wejściowych, które następuje z permanentnym zachowywaniem związku ze zmienną zależną. Świadczą o tym między innymi wyniki analizy wpływu składowych głównych na wyjaśnienie zmienności danej zależnej przedstawione w tabeli 4.

Tabela 4. Zestawienie wpływu poszczególnych składowych głównych uzyskanych metodami PLSR i PCA na zmienność danej objaśnianej s_z wg [18]

Table 4. Summary of impact of particular main components obtained by PLSR and PCA methods on the variability of output data s_z , acc. to [18]

Numer składowej głównej	<i>PCA</i>	<i>PLSR</i>
	Wartości współczynnika determinacji R^2	
PC_1	2,48	8,16
PC_2	6,25	0,62
PC_3	2,03	0,01
PC_4	0,00	0,44
PC_5	0,04	0,75
PC_6	0,03	0,49
PC_7	1,11	0,01

Z rezultatów badań przedstawionych w tabeli 4 wynika, że metoda *PLSR* jest bardziej efektywna na poziomie wstępnej analizy danych niż metoda *PCA*. W metodzie *PLSR* pierwsza wyłoniona składowa główna przyczynia się do wyjaśnienia powyżej 8% zmienności zawartej w pierwotnym zbiorze zmiennej wyjściowej s_z . Porównywalny rząd wielkości wyjaśnionej zmienności w metodzie *PCA* przypada dla trzech pierwszych składowych głównych. Mając na uwadze przyjęty cel, jakim było wyodrębnienie jednej, optymalnej dla opisu zużycia technicznego budynków reprezentacji wskaźnika uszkodzeń widać, że metoda *PLSR* pozwala na jego realizację. Wskazuje bowiem tylko jedną składową główną, która zdecydowanie odbiega ilościowo od składowych pozostałych.

Ostatecznie dla pierwszej składowej głównej (PC_1 – tab. 4) uzyskano wskaźnik uszkodzeń dla badanej grupy budynków, opisany w postaci liniowej kombinacji intensywności uszkodzeń elementów składowych, wygenerowanej metodą *PLSR* w postaci równania (3.1):

$$w_u = 0,057w_{u2} + 0,058w_{u3} + 0,064w_{u7} + 0,039w_{u11} + 0,015w_{u12} + 0,037w_{u13} + 0,044w_{u17} + 7,03 \quad (3.1)$$

4. Podsumowanie i wnioski

- Przedstawiona w rozdziale 2 metodyka badań była podstawą do utworzenia wielowymiarowych modeli przebiegu zużycia technicznego budynków na terenach górniczych. Pozwoliło to na jakościową i ilościową ocenę wpływu czynników górniczych i budowlanych na stan techniczny badanej zabudowy.
- Opisane w rozdziale 3 badania pozwoliły na ocenę udziału wyspecyfikowanych wskaźników uszkodzenia elementów budynków w wartości ich stopnia zużycia technicznego. Wykazano, że podejście wykorzystujące *drzewa decyzyjne* może okazać się przydatne w początkowej fazie analizy bazy danych. Pozwala ono na utworzenie wstępnego modelu oraz wnioskowanie o udziałach poszczególnych czynników wejściowych w zmienności zależnej danej wyjściowej. Ponadto stwierdzono, że metoda *PLSR* jest bardziej efektywna na poziomie wstępnej analizy danych o intensywności uszkodzeń budynków, niż metoda *PCA*. Wygenerowane metodą *PLSR* wstępne przybliżenie uogólnionego wskaźnika uszkodzeń budynków o konstrukcji wieloblokowej może być wykorzystane do oceny ich stanu technicznego, bądź oceny zakresu szkód górniczych.
- Opisane wyniki wstępnych analiz potwierdzają potrzebę poszukiwań uniwersalnego wskaźnika intensywności uszkodzeń dla budynków wznoszonych na terenach górniczych oraz pozwalają z optymizmem planować jego dalsze wykorzystanie do analizy zagrożenia zabudowy terenów górniczych.
- W planach badań uzasadnione wydaje się wykorzystanie formalizmów probabilistycznego wnioskowania warunkowego i wdrożenie *sieci Bayesowskich* lub specyficznego rodzaju sieci neuronowych *PNN (Probabilistic Neural Networks)*. Powinny one pozwolić na połączenie niepewności formalnej z prawdopodobieństwem występowania poszczególnych zmiennych mających wpływ na przebieg zużycia technicznego budynków. Wykorzystanie podejścia probabilistycznego jakiego dają *sieci Bayesowskie* pozwolić może na powiązanie wyników dotychczasowych badań z dziedzina niezawodności konstrukcji, w której od dawna funkcjonują techniki oparte na ściśle zdefiniowanych rozkładach prawdopodobieństwa.

Referat opracowano w ramach badań statutowych AGH nr 11.11.150.005.

Literatura

1. *Burges C.J.C.*: A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery*, 2, 121–167 (1998), 1998 Kluwer Academic Publishers, Boston 1998.
2. *Breiman, L., Friedman, J.H., Olshen, R.A., and Stone, C.I.*: Classification and regression trees. Belmont, Calif.: Wadsworth 1984.
3. *Firek K.*: Badanie wpływu czynników górniczych i budowlanych na zużycie techniczne tradycyjnej zabudowy terenu górniczego LGOM. Praca doktorska, niepublikowana. Akademia Górniczo-Hutnicza, Kraków 2005.
4. *Firek K.*: Proposal for classification of prefabricated panel building damage intensity rate in mining areas (*Propozycja klasyfikacji intensywności uszkodzeń budynków wielkopłytych na terenach górniczych*). X Dni Miernictwa Górniczego i Ochrony Terenów Górniczych, Kraków 2009. *Archives of Mining Sciences*, 2009, Vol. 54, No 3, Kraków.
5. *Firek K.*: Ocena udziału najczęściej występujących typów uszkodzeń w stopniu zużycia technicznego budynków murowanych usytuowanych na terenach górniczych (*Evaluating the contribution of the most common types of damage on the degree of technical wear of masonry buildings located in mining areas*). *Geomatics and Environmental Engineering, Quarterly*, AGH University of Science and Technology Press, 2013, Vol. 7, No. 2, Cracow.
6. *Geladi P., Kowalski R.B.*: Partial least-squares regression: a tutorial. *Journal: Analytica Chimica Acta-ANAL CHIMACTA*, 1986, vol. 185, no. 1.
7. *Hastie T., Tibshirani R., Friedman J.*: The Elements of Statistical Learning: Data Mining, Inference, and Prediction. February 2009. Springer Series in Statistics 2009.
8. *Łęski J.*: Systemy neuronowo-rozmyte. WNT, Warszawa 2008.
9. *Oowski S.*: Sieci neuronowe do przetwarzania informacji. Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa 2006.
10. *Schölkopf B., Smola A.J.*: Learning with Kernels. MIT Press, Cambridge, Massachusetts 2002.
11. *Rusek J.*: Creating model of technical wear of building in mining area, with utilization of regressive SVM approach (*Budowa modelu przebiegu zużycia technicznego budynków na terenach górniczych z wykorzystaniem metody wektorów podpierających „support vector machine” w ujęciu regresyjnym*). X Dni Miernictwa Górniczego i Ochrony Terenów Górniczych, Kraków 2009. *Archives of Mining Sciences*, 2009, Vol. 54, No 3, Kraków.
12. *Rusek J., Wodyński A.*: Creating a model of technical wear of buildings in mining areas with the use of fuzzy inference systems (Modelowanie przebiegu zużycia technicznego budynków na terenach górniczych z wykorzystaniem systemów wnioskowania rozmytego). 56 Konferencja Naukowa „Problemy naukowo-badawcze budownictwa”, Kielce-Krynica 2010.
13. *Rusek J.*: Modelowanie stopnia zużycia technicznego budynków na terenach górniczych z wykorzystaniem wybranych metod sztucznej inteligencji. Praca doktorska, niepublikowana. Akademia Górniczo-Hutnicza. Kraków 2010.
14. *Vapnik V.*: Statistical learning theory. John Wiley & Sons, New York 1998.
15. *Wodyński A.*: Zużycie techniczne budynków na terenach górniczych (*The process of technical wear of buildings in mining areas*). Uczelniane Wydawnictwa Naukowo-Dydaktyczne, Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie, Kraków 2007.
16. *Wodyński A., Firek K., Rusek J.*: Metoda wektorów podpierających (SVM) w modelowaniu zużycia technicznego budynków wielkopłytych na terenach górniczych (*Support Vector Method (SVM) in technical wear modelling of prefabricated buildings (using large plate) on mining areas*). VIII Konferencja Naukowo-Techniczna „Ochrona środowiska na terenach górniczych, Szczyrk-Katowice, 2010. „Przeгляд Górniczy”, 2010, Nr 10.