

Stanisław Gruszczyński*

GEOINFORMATYCZNE NARZĘDZIA W BADANIU GLEB

1. Wprowadzenie

Skutkiem powszechnej informatyzacji jest, między innymi, przybywanie kolejnych dziedzin działalności, w których gromadzenie i analiza danych służących do podejmowaniu decyzji odbywa się z udziałem przetwarzania cyfrowego. Nawet tak tradycyjne dyscypliny, jak gleboznawstwo, prowadzenie gospodarstwa rolnego lub leśnego, wykorzystują w procesach decyzyjnych technologie informacyjne. Główną przyczyną rozpowszechnienia się tego podejścia jest wygoda gromadzenia, przechowywania i udostępniania danych przestrzennych, jak też wykonanych na ich podstawie analiz.

Dyrektywa unijna dotycząca infrastruktury danych przestrzennych (Model INSPIRE) zaleca włączanie do gromadzonych danych cyfrowych informacji dotyczących gleb i podglebia, charakteryzowanych poprzez głębokość, teksturę, strukturę i zawartość cząstek organicznych, kamienistość, w miarę potrzeby – średni spadek. Zaleca się integrację danych dotyczących wszystkich rodzajów użytkowania (obecnie oddzielna dokumentacja kartograficzna w Polsce obejmuje tereny rolne i obszary lasów państwowych), a także realizację klasyfikacji bonitacyjnej na terenach leśnych należących do państwa.

Inne okoliczności, związane z zamiarami zahamowania niekorzystnych tendencji rozwoju cywilizacyjnego, spowodowały dostrzegalną zmianę podejścia do gleb jako obiektów przyrodniczych podlegających dokumentowaniu. Tradycyjnie gleby traktuje się jako składnik gospodarstwa wiejskiego, są zatem głównie badane i klasyfikowane pod kątem wypełnienia specyficznych funkcji tych rodzajów działalności. Obecnie jednak dostrzega się także inne funkcje gleb, takie jak: akumulacja związków chemicznych emitowanych w wyniku działań gospodarczych, pośredniczenie w zanieczyszczaniu i oczyszczaniu wód, pośred-

* Akademia Górniczo-Hutnicza, Wydział Geodezji Górniczej i Inżynierii Środowiska

niczenie w kształtowaniu poziomu ryzyka środowiskowego, wskaźnik stopnia antropogenicznych przekształceń środowiska.

Nic dziwnego zatem, że obecnie formułowane cele dokumentowania gleb obejmują cztery zadania [25]:

- 1) ich opis i charakterystykę,
- 2) wszechstronną klasyfikację zgodnie z obowiązującymi standardami,
- 3) ustalenie granic jednostek glebowych,
- 4) predykcję ich zachowania pod wpływem różnego typu przekształceń.

Zwłaszcza ostatnie z wymienionych zadań wymaga zastosowania zaawansowanych narzędzi prognozowania.

Przekształcenia gleb można rozpatrywać w dwu aspektach. Pierwszy z nich wynika z naturalnych dążeń do poprawiania ich wartości użytkowej poprzez rozmaite działania melioracyjne. Mają one na celu usunięcie przyczyn niektórych wad gleb, na przykład nadmiernej podmokłości gruntów. Akcje tego typu są podejmowane obecnie z mniejszą intensywnością, aczkolwiek w latach finalizowania kartograficznego dokumentowania gleb rolniczych Polski przewidywano [28] dezaktualizowanie, w wyniku tego typu działań, około 10÷20% dokumentacji w ciągu kilkudziesięciu lat. Rozwiązanie tego problemu uzależniano od skali dezaktualizacji poszczególnych arkuszy; problem ten może być obecnie postawiony w innym świetle, głównie w wyniku wykorzystania elastycznych metod GIS.

Drugi aspekt przekształceń wynika z niezamierzonych następstw oddziaływań antropogenicznych, głównie przemysłowych. Ten typ przekształceń charakteryzuje się na ogół znacznym stopniem koncentracji (ogranicza się do rejonów przemysłowych) i znacznie większym natężeniem negatywnych skutków. Oznacza to, że leży on poza główną sferą zainteresowań specjalistów z dziedziny dokumentowania gleb, choć lokalnie może stanowić istotny problem. Dostrzegając go, ale też akceptując pojawienie nowych możliwości analitycznych, uważa się, że systemy dokumentowania gleb reprezentują obecnie raczej bazę wiedzy [4, 22], podlegającą różnorodnym analizom, głównie przestrzennym, służącym wyprowadzaniu użytecznych wniosków. W kontekście tej dominującej tendencji naturalnym kandydatem do roli podstawowego instrumentu analizy aktualnego stanu i predykcji zachowania się gleb w przyszłości jest zbiór algorytmów służących do poszukiwania wiedzy w bazach danych KDD (*Knowledge Discovery in Databases*). Systemy GIS służące do budowania obszernych baz danych przestrzennych są dobrym polem zastosowań technik KDD; dzięki wykorzystaniu algorytmów eksploracji danych DM (*Data Mining*), techniki KDD są narzędziem wyszukiwania regularności i wyjątków w obszernych zbiorach informacji [23].

2. Techniki KDD

Rozwój technik KDD i DM jest naturalną konsekwencją powszechnego i nasilającego się rozwoju rozmaitych baz danych. Bazy danych początkowo stanowiły zbiory dokumentujące: aktywność klientów banków, sprzedaż artykułów w sieciach handlowych, sprzedaż

usług linii lotniczych, zmiany warunków meteorologicznych, przebieg zjawisk przyrodniczych itp. Zainteresowanie GIS spowodowało powstanie baz danych wzbogaconych o informację przestrzenną, która obok informacji czasowej stała się istotnym czynnikiem różnicującym dane. Zachowując dokumentacyjny sens baz danych, zwrócono uwagę, że są one nośnikiem cennej dla użytkowników wiedzy, która wymaga jednak konkretyzacji poprzez ekstrakcję użytecznych prawidłowości (regularności, wzorców) charakteryzujących związki między danymi. Ich poznanie pozwala, z pewną dozą ryzyka, na monitorowanie bieżących zachowań obserwowanych systemów, prognozowanie ich zachowań pod wpływem różnych bodźców oraz wykrywanie symptomów odbiegających od normy, które być może sygnalizują pojawienie się niepożądanych oddziaływań w systemie (na przykład: nietypowe zachowanie się klienta banku, być może wskazujące na przejęcie jego uprawnień przez osobę obcą, wykrycie określonej treści jednego z tysięcy zobrazowań satelitarnych wskazującej na pojawienie się niepokojącej sytuacji meteorologicznej, wykrycie wystąpienia splotu okoliczności sprzyjających powstaniu sytuacji smogowej itp.).

KDD obejmuje kilka następujących po sobie faz [23]:

1. Przygotowanie danych, składające się z kilku etapów, zależnie od okoliczności: gromadzenie, łączenie, czyszczenie (usuwanie danych niepewnych lub niepełnych), selekcja, preprocesing, itp.
2. Określenie funkcji i algorytmu eksploracji danych, zakreślające ramy poszukiwania odpowiedniej do postulowanego celu metody eksploracji.
3. Zgłębianie danych polegające na poszukiwaniu najlepszego w danych warunkach zbioru reguł określających wzorce zależności zawarte w bazie danych.
4. Interpretacja i wizualizacja wykrytych reguł.
5. Aplikacja wiedzy zawartej w modelu polegająca na skonstruowaniu najlepszej w danych warunkach, użytecznej formy realizacji zadania związanego z identyfikacją wykrytych wzorców.

Kluczowym zagadnieniem jest wybór optymalnej, w kontekście celu badania, metody zgłębiania (eksploracji) danych. Można zaliczyć do nich wiele różnych algorytmów, stosowanych w połączeniach lub indywidualnie. Należą do nich między innymi:

- modele statystyczne,
- modele probabilistyczne,
- drzewa decyzyjne i drzewa reguł,
- modele bazujące na oprogramowaniu logiki indukcyjnej,
- modele bazujące na przykładach,
- sieci neuronowe,
- modele bazujące na wnioskowaniu rozmytym,
- modele bazujące na zbiorach przybliżonych,
- algorytmy genetyczne,
- modele hybrydowe z zakresu „softcomputingu”.

Łatwo zauważyć, że wymienione metody analizy i syntezy mają zastosowanie również w innych niż KDD dyscyplinach wiedzy. Niektóre z nich należą do dziedzin sztucznej inteligencji i uczenia maszynowego, niektóre z nich (na przykład sieci neuronowe czy algorytmy genetyczne) stanowią też samodzielne obszary badań.

3. Zastosowanie KDD w dokumentowaniu gleb

Niezwykle szeroki zakres zastosowań metod eksploracji danych i wyszukiwania na tej podstawie użytecznej wiedzy uzyskuje, w odniesieniu do dokumentowania gleb, znaczenie związane z ich naturalną zmiennością przestrzenną i czasową [31, 32]. Zapewne możliwe jest wskazanie wielu ewentualnych zastosowań wspomnianych metod, jednakże można zwrócić uwagę na trzy następujące, charakterystyczne dla problematyki oceny zasobów glebowych zagadnienia:

- 1) klasyfikację gruntów,
- 2) opis przestrzennego rozkładu zanieczyszczeń gleb,
- 3) prognozowanie stanu gleb w zależności od przewidywanych przekształceń środowiska.

3.1. Klasyfikacja gruntów

Gleby podlegają licznym podziałom mającym na celu możliwie dokładną charakterystykę ich właściwości. Podział typologiczny służy sklasyfikowaniu gleb pod względem przestrzennego zróżnicowania głównych procesów genetycznych. Podział bonitacyjny obrazuje przestrzenne zróżnicowanie jakości gleb w odniesieniu do określonego użytkowania. Można wyobrazić sobie wiele innych podziałów, służących do pełniejszego charakteryzowania gruntów, na przykład podział na siedliska leśne lub kompleksy rolniczej przydatności. Główną cechą charakteryzującą wszelkiego rodzaju klasyfikacje gleb jest konieczność transformacji ciągłej zmienności cech na jej obraz dyskretny, tworzony dla celów sporządzenia map tematycznych. W podziale typologicznym, gdzie diagnoza odbywa się na podstawie badania cech poszczególnych horyzontów (a także takich cech, jak uziarnienie, barwa, struktura itp.), miąższości poziomów najprawdopodobniej wykazują pewną zmienność przestrzenną, przez co należy się liczyć z lokalnymi odstępstwami od idealnego obrazu określonej jednostki typologicznej. Podobnie w klasyfikacji bonitacyjnej wyodrębnione kontury, odnoszone do idealnego wzorca klasy, są do niej mniej lub bardziej podobne. Tak samo jest z innymi rodzajami klasyfikacji. W tradycyjnej kartografii cechy i jednostki są wyraźnie rozgraniczone. Elastyczność systemów cyfrowych dopuszcza inne podejścia, pełniej oddające zmienność gruntów. Akceptowanym modelem zmienności przestrzennej gleb może być obraz bazujący na wnioskowaniu rozmytym. Metodycznym

aspektami rozmytej klasyfikacji typologicznej gleb zajmowano się w [1]. W niektórych krajach wnioskowanie rozmyte jest obowiązującym trybem klasyfikacji typologicznej gleb (np. Australia).

Obowiązująca w Polsce dedukcyjna klasyfikacja bonitacyjna gleb, bazująca na obszernym zbiorze wzorców, zawodzi w przypadkach konieczności zaklasyfikowania gruntu spoza zbioru dostępnych wzorców, na przykład poddanego rekultywacji, bez istotnych cech naturalnego procesu glebotwórczego. W takim przypadku konieczne jest odwołanie się do techniki KDD jako narzędzia odpowiedniego do sformułowania reguł klasyfikacji na podstawie bazy danych reprezentatywnej dla gleb charakteryzujących się różnymi cechami litologicznymi i morfologicznymi. Projekt badawczy z takim zadaniem jest obecnie realizowany na Wydziale Geodezji Górniczej i Inżynierii Środowiska AGH.

W zakresie metodycznym polega on na zbudowaniu odpowiedniego zbioru reguł klasyfikacji, bazującego na mierzalnych cechach glebowych w celu identyfikacji położenia gruntu w szeregu bonitacyjnym, przy braku bezpośredniego odniesienia do systemu wzorców klasyfikacji gleb. W pierwszym etapie zebrano dane laboratoryjne dotyczące oznaczeń cech gleb oraz zarchiwizowane dane dotyczące obiektów poddawanych rekultywacji (bezglebowych). Celem badań jest zaproponowanie systemu klasyfikacji gruntów pozbawionych cech gleb naturalnych. Ze względu na praktyczne uwarunkowania system klasyfikacji musi mieć charakter wnioskowania rozmytego.

3.2. Ocena zanieczyszczenia gleb

Model przestrzenny stanu sanitarnego gleb, w celu spełnienia wymogów prawnych [26], powinien odzwierciedlać koncentrację normowanych zanieczyszczeń w trójwymiarowej przestrzeni środowiska glebowego. Odpowiednie przepisy wydzielają dwa typy terenów: odpowiadające standardom (umożliwiające określone użytkowanie gruntów) oraz nieodpowiadające standardom (wymagające rekultywacji lub remediacji). Standardy normują koncentrację zanieczyszczeń w dwu horyzontach głębokości: 0÷0,3 m oraz 0,3÷15,0 m.

Intencją prawodawcy było zapewne spowodowanie jednoznacznego rozgraniczenia terenów „czystych” i „zanieczyszczonych”. Jest to jednak utrudnione ze względu na:

- obecność mniej lub bardziej wyraźnego, nieliniowego trendu w poziomym rozkładzie zanieczyszczeń wokół ich źródeł;
- obecność nieliniowego trendu w pionowym rozkładzie zawartości zanieczyszczeń w profilu glebowym;
- znaczną zmienność zjawiska, wnikającą z rozdrobnienia własności oraz form użytkowania gleb w naszych realiach struktury własności gruntów.

Oczywistym warunkiem opisu stanu sanitarnego gleb jest dostatecznie reprezentatywna baza oznaczeń zawartości zanieczyszczeń w glebach. Kluczowym problemem jest zaproponowanie dobrego algorytmu rozszerzającego punktowe oznaczenia zawartości zanieczyszczeń na przestrzeń pomiędzy węzłami siatki obserwacyjnej.

W warunkach znacznej, przestrzennej zmienności i nieregularności rozkładu zanieczyszczeń metody interpolacji mogą być zawodne z powodu obciążenia danych dużym szumem losowym. Należy zwrócić uwagę, że na podstawie badania w punkcie wnioskujemy o pewnym szerszym terenie, warto byłoby zatem poznać przynajmniej stopień wiarygodności naszego oszacowania zanieczyszczenia. Naturalnymi kandydatami do ustalenia pełnego modelu zmienności przestrzennej są zatem algorytmy regresji lub klasyfikacji, dostatecznie elastyczne dla wyraźnej nieliniowości zjawiska, a równocześnie zapewniające sprawność obliczeniową przydatną w systemie GIS.

Należy zwrócić uwagę, że w naszych warunkach geologicznych zanieczyszczenie gleb, w sensie ujmowanym w standardach, może być praktycznie wyłącznie skutkiem wpływu jakiegoś rodzaju emisji. Sporadycznym odstępstwem od tej reguły mogą być obszary wychodni dolomitów kruszczośnych oraz niektóre osady fliszowe zasobniejsze w kadm; lokalizacja takich terenów jest jednak dobrze rozpoznana. Poza emisją dwutlenku i trójtlenku siarki, która jest zjawiskiem powszechnym, pozostałe źródła emisji mają charakter punktowy lub liniowy. Biorąc pod uwagę sformułowania standardów czystości ziemi, informacja o stanie zanieczyszczenia powinna prezentować jego koncentrację w glebach w określonym położeniu przestrzennym. Ze względu na nieliniowy, zazwyczaj, trend koncentracji w profilu glebowym konieczne jest podanie zakresu głębokości obowiązywania tego oszacowania. Najprostszy model rozmieszczenia zanieczyszczeń w glebach powinien zatem obejmować wektor złożony z co najmniej czterech zmiennych objaśniających: dwu współrzędnych poziomych oraz dwu wartości głębokości, między którymi mieści się analizowana warstwa gleby. Na podstawie tego wektora położenia x system powinien dostarczyć oszacowania reprezentatywnej dla niego wartości t , będącej rozsądnym przybliżeniem koncentracji danego zanieczyszczenia w glebie. Wartość t może być ciągła, stanowiąc przybliżenie koncentracji zanieczyszczenia w jednostkach fizycznych (np. w ppm lub mg/kg gleby), lub też dyskretna, wskazując na przynależność warstwy w tym położeniu do jednej z klas czystości (np. NORMA – koncentracja mieści się w standardzie, NNORMA – koncentracja przekraczająca zawartość standardową). Można rozważać wiele różnych modeli umożliwiających budowę użytecznego algorytmu dającego odpowiednie oszacowanie. Dobrym rozwiązaniem może być na przykład algorytm MDN (*Mixture Density Network*) [2, 3, 9, 10] będący połączeniem modelu mieszanych rozkładów (na przykład sumowania rozkładów normalnych) z optymalizacją ich parametrów, na podstawie danych obserwacyjnych, za pomocą perceptronu wielowarstwowego (sieci MLP). Algorytm ten dostarcza oszacowania $p(t | x)$ warunkowego prawdopodobieństwa wartości t przy wektorze objaśniającym x . Oszacowanie tego rodzaju jest szczególnie cenne, kiedy zależność nie jest homoscedastyczna lub rozkład warunkowy jest wielomodalny. Jest to okoliczność bardzo istotna w procedurze oceny stopnia zanieczyszczenia gleb w rejonach silnie obciążonych, ponieważ lokalna wariancja zawartości zanieczyszczeń wzrasta wraz ze wzrostem ich koncentracji.

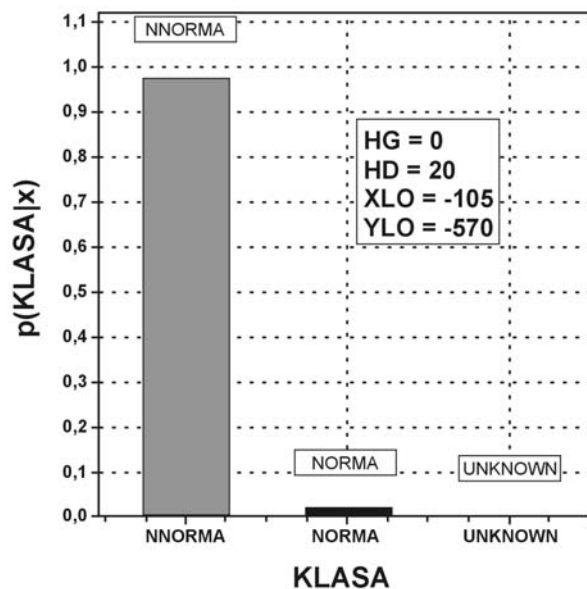
Równie elastyczny i efektywny model rozkładu zanieczyszczeń może być opracowany na bazie alternatywnych algorytmów adaptacyjnych. Należy tu wskazać na przykład algo-

rytm FSM (*Feature Space Mapping*), opracowany na Uniwersytecie w Toruniu [6, 7, 18] i udostępniony w pakiecie o nazwie *Ghost-Miner* firmy Fujitsu Poland (FQS).

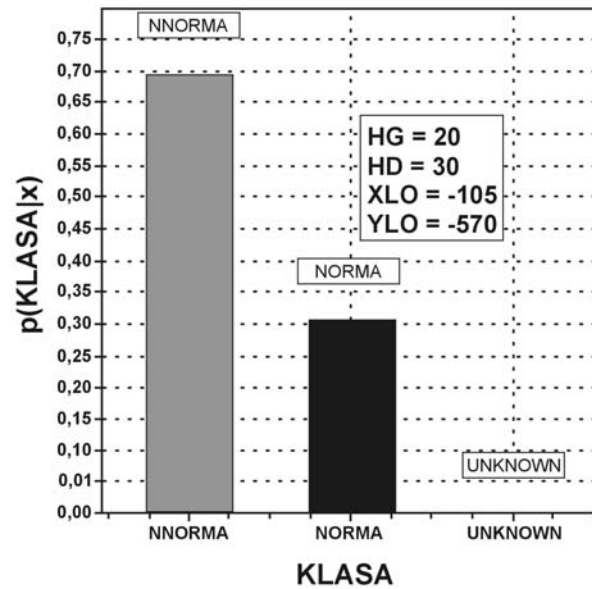
Rysunki 1–4 obrazują przydatność analityczną tego algorytmu. Daje on możliwość oszacowania prawdopodobieństwa warunkowego spełnienia lub niespełnienia standardu czystości, w tym przypadku ze względu na zawartość chromu w glebach, w określonym położeniu. Zmiennymi objaśniającymi są: współrzędna x w układzie lokalnym (XLO , źródło emisji ma współrzędne $[0, 0]$), współrzędna y w układzie lokalnym (YLO), głębokość „stropu” analizowanej warstwy gleby (HG) i „spągu” analizowanej warstwy gleby (HD).

Wydaje się, że model dość dobrze oddaje naturę zjawiska: ścisłe powiązanie między koncentracją zanieczyszczenia a głębokością warstwy glebowej. Jego skuteczność jest bardzo dobra – ponad 98% poprawnych wskazań. Algorytm pozwala na oszacowanie bayesowskiego prawdopodobieństwa *a posteriori* w różnej konfiguracji danych objaśniających. Jego silną stroną jest także kontrola lokalizacji punktu w zakresie danych bazy. Przekroczenie tego zakresu jest sygnalizowane przypisaniem oszacowania do klasy „nieznana”.

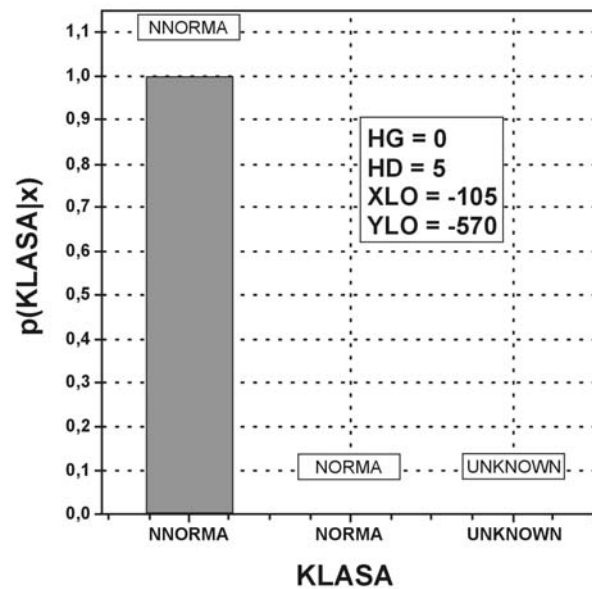
Algorytm FSM, przy pewnych dodatkowych założeniach dotyczących kształtu funkcji transferu, daje możliwość wyekstrahowania z danych reguł logicznych klasyfikacji. Jest to realizacja procedur KDD przy użyciu samoorganizującej się sieci neuronowej (tab. 1). Ekstrakcja reguł ułatwia konstrukcję rozmytego systemu klasyfikacji, choć przejście to nie jest bezpośrednie i wymaga dodatkowej kalibracji.



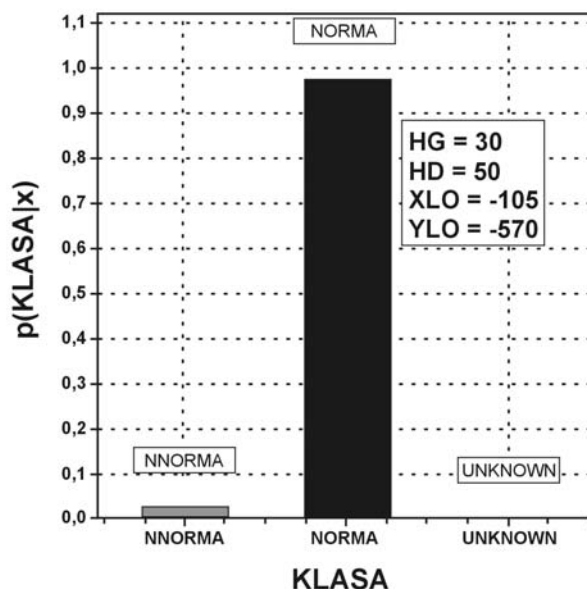
Rys. 1. Oszacowanie warunkowego prawdopodobieństwa przynależności warstwy glebowej 0÷20 cm, w określonym położeniu względem źródła emisji, do klas standardów czystości ziemi za pomocą klasyfikatora FSM



Rys. 2. Oszacowanie warunkowego prawdopodobieństwa przynależności warstwy glebowej 20÷30 cm, w określonym położeniu względem źródła emisji, do klas standardów czystości ziemi za pomocą klasyfikatora FSM



Rys. 3. Oszacowanie warunkowego prawdopodobieństwa przynależności warstwy glebowej 0÷5 cm, w określonym położeniu względem źródła emisji, do klas standardów czystości ziemi za pomocą klasyfikatora FSM



Rys. 4. Oszacowanie warunkowego prawdopodobieństwa przynależności warstwy glebowej 30÷50 cm, w określonym położeniu względem źródła emisji, do klas standardów czystości ziemi za pomocą klasyfikatora FSM

Tabela 1. Podział reguł klasyfikacji gleb na: spełniające standard (NORMA) i niespełniające go (NNORMA), w otoczeniu zakładów chemicznych, ze względu na koncentrację Cr. Reguły wyekstrahowane z danych obserwacyjnych przy użyciu pakietu *GhostMiner*

Lp.	Reguła klasyfikacji gleb
1	if HD in [15.3005,151.137] and YLO > 347.5 then class NORMA
2	if HD in [3.79087,30] and YLO in [-1147.5,495] then class NNORMA
3	if HD in [8.8764,33.0608] and YLO < -2021.99 then class NORMA
4	if HD > 49.394 then class NORMA
5	if XLO in [1242.31,2750.49] and YLO in [-2119.78,-998.269] then class NORMA
6	if HD > 29.489 and XLO < -44.5 then class NORMA
7	if HD < 19.712 and XLO < 1530 then class NNORMA

3.3. Prognozowanie następstw przekształceń gleb

Gleby są obiektami dynamicznymi, choć tempo ich zmian w normalnych okolicznościach jest bardzo powolne. Przyczynami ich sukcesywnych przekształceń są naturalne procesy glebotwórcze, wpływające na dostosowywanie się gleb do istniejących warunków,

działania związane z celowym dostosowaniem gleb do wymagań użytkownika (melioracje), uprawa rolnicza lub leśna. W rejonach przemysłowych często pojawiają się okoliczności modyfikujące naturalne tendencje rozwojowe gleb. Szczególnie częste są one na obszarach górniczych [11–14, 19, 27, 29]. Ponieważ przekształcenia te dotyczą czynników morfologicznych i hydrologicznych, stosunkowo łatwo (przy odpowiedniej liczbie danych) zbudować model klasyfikacji, wykorzystując jeden z dostępnych algorytmów adaptacyjnych. Według wyników dostępnych badań odpowiednimi mogą tu być sieci MLP (*Multi-layer Perceptron*), RBF (*Radial Basis Function*) czy też PNN (*Probabilistic Neural Network*). Systematycznie rozwijane techniki KDD dają obecnie możliwość pogłębionej analizy danych poprzez wykorzystanie techniki zespołów klasyfikatorów i modelowanie prawdopodobieństwa przynależności do odpowiednich klas bonitacyjnych przy modyfikacji zmiennych objaśniających. Zespoły (komitety) modeli [16–18, 20, 21, 24, 30] dają znacząco lepsze rezultaty niż pojedyncze modele, głównie z powodu niwelowania słabych stron niektórych algorytmów w punktach przestrzeni przez algorytmy radzące sobie lepiej. Równocześnie, analogicznie jak w sieciach probabilistycznych, zespoły modeli pozwalają na szacowanie bayesowskiego prawdopodobieństwa przynależności do klas w różnych konfiguracjach. Istnieje zatem możliwość szczegółowej analizy zachowań się gleb w warunkach zmieniających się czynników zewnętrznych. Pomocne może być tutaj określenie podobieństwa warunków dla różnych konfiguracji morfologii i litologii gleb. Zastosowanie zespołu klasyfikatorów złożonego z sieci FSM, zbioru sieci klasy IncNet oraz drzewa decyzyjnego SSV wyodrębniającego 125 logicznych reguł klasyfikacji [18] poprawiło skuteczność identyfikacji klas, w stosunku do najlepszego, pojedynczego klasyfikatora, z ok. 87 do ponad 90%. Rozszerzenie zespołu o sieci z nielokalnymi funkcjami transferu poprawia ten wynik. Zespół klasyfikatorów złożony z sieci liniowej, sieci MLP, RBF i PNN, których wartości aktywacji są następnie przetwarzane przez klasyfikator typu FSM, uzyskuje skuteczność identyfikacji około 95%. Konstrukcja tego typu szczególnie przydatna może być w analizach wpływu zmian poszczególnych parametrów morfologicznych (rzędnej, nachylenia, głębokości do wody gruntowej) na przeklasyfikowanie gleb. Może to być istotnym czynnikiem kształtowania decyzji o sposobie eksploatacji górniczej zmierzającej do minimalizowania szkód w użytkowaniu przyrodniczym.

4. Podsumowanie

1. Transformacja informacji analogowych na format cyfrowy udostępnia badaniom środowiskowym nowe narzędzia analizy umożliwiające nie tylko wygodniejszą i szybszą wizualizację, ale także pogłębione studia zjawisk i skutków ich zakłóceń. Łatwo zauważyć, że stopniowo umacnia się tendencja do wykorzystywania algorytmów adaptacyjnych nie tylko do modelowania zjawisk i zależności nieliniowych, ale także jako istotnego źródła informacji dotyczących wewnętrznych zależności w wielkich zbiorach danych.

2. Obowiązuje wymóg poszerzenia celów dokumentowania gleb o prognozowanie ich zachowań pod wpływem różnorodnych oddziaływań, w tym antropopresji. Fakt dedukcyjnego systemu klasyfikacji gleb, obowiązującego w Polsce, znacznie utrudnia przewidywanie stanu gleb w różnych, mniej typowych okolicznościach. Koniecznym uzupełnieniem diagnostyki bonitacyjnej mogą być rozmyte systemy klasyfikacji gruntów rekultywowanych, niemających odpowiedników wśród wzorców obowiązującego systemu.
3. Ważnym obszarem zastosowań narzędzi geoinformatycznych jest ocena ryzyka środowiskowego związanego z akumulacją zanieczyszczeń w glebach. W warunkach rozdrobnionej własności gruntów istnieje dominująca tendencja do znacznego zróżnicowania koncentracji zanieczyszczeń, nawet przy niewielkich odległościach punktów obserwacji. Sprawia to trudność w poprawnym zaklasyfikowaniu gruntów do standardów czystości i wyznaczaniu zasięgów zanieczyszczenia przemysłowego. Użyteczne w takich przypadkach są algorytmy adaptacyjne wyposażone w instrumenty oceny prawdopodobieństwa poszczególnych stanów gleb.
4. Systematyczny rozwój technik eksploracji danych dostarcza coraz nowszych narzędzi do badania współzależności wewnątrz baz danych przestrzennych. Perspektywnym kierunkiem rozwoju tych algorytmów jest ich zastosowanie w formie zespołów klasyfikatorów i modeli regresyjnych. Ich wprowadzenie ułatwia oszacowanie wiarygodności prognoz skutków przekształceń środowiska.

Literatura

- [1] Ameskamp M.: *Three-Dimensional Rule-Based Continuous Soil Modelling*. Christian-Albrechts-Universität Kiel, Institut für Informatik und Praktische Mathematik 1997 (praca doktorska)
- [2] Bishop C.M.: *Mixture density networks*. Raport instytutowy NCRG/94/004, dostępne na stronie: citeseer.ist.psu.edu/bishop94mixture.html, 1994
- [3] Bishop C.M.: *Neural Networks for Pattern Recognition*. Oxford University Press 1995
- [4] Bui L.N.: *Soil survey as a knowledge system*. *Geoderma*, 120, 2004, 17–26
- [5] Cornford D., Nabney L.T., Bishop C.M.: *Neural network based wind vector retrieval from satellite scatterometer data*. *Neural Computing and Application*, 8, 1999, 206–217
- [6] Duch W., Dierksen G.H.F.: *Feature space mapping as a universal adaptive system*. *Computer Physics Communications*, 87 (16), 1994, 341–371
- [7] Duch W., Grabczewski K.: *Heterogeneous adaptive systems*. Dostępne na stronie: citeseer.ist.psu.edu/627984.html, 2002
- [8] Duch W., Setiono R., Żurada J. M.: *Computational intelligence methods for rule-based data understanding*. *Proceedings of the IEEE*, 92 (5), 2004, 771–805
- [9] Evans D.J.: *Mixture density network training by computation in parameter space*. Raport instytutowy NCRG/98/016, Neural Computing Research Grup, Aston University, Birmingham B4 7ET, dostępne na stronie: <http://www.ncrg.aston.ac.uk>, 1998

- [10] Goldberg P.W., Williams C.K.I., Bishop C.M.: *Regression with input-dependent noise: A gaussian process treatment*. Advances in Neural Information Processing Systems, vol. 10, The MIT Press 1998, dostępne na stronie: cite-seer.ist.psu.edu/article/goldberg98regression.html
- [11] Gruszczyński S.: *Ocena zagrożenia gleb w rejonach górniczych za pomocą sztucznych sieci neuronowych*. Geoinformatica Polonica, 1 (1), 1999, 45–62
- [12] Gruszczyński S.: *Symulacja skutków przekształceń gleb na terenach górniczych za pomocą klasyfikatorów neuronowych*. Kraków, UWND AGH 2000
- [13] Gruszczyński S.: *Szkody górnicze na terenach użytkowanych przyrodniczo – niektóre problemy dokumentowania, prognozy i oceny*. Prace Naukowe GIG, 2002, 167–172
- [14] Gruszczyński S.: *Raport merytoryczny z realizacji projektu badawczego Komitetu Badań Naukowych 9 T 12E 016 18 pt. „Prognozowanie przyrodniczych skutków odtwarzania układu hydrologiczno-glebowego w otoczeniu likwidowanych zakładów górniczych”*. Kraków, AGH 2003
- [15] Gruszczyński S., Żuławski C.: *Dokumentowanie szkód wywołanych osuszeniem gleb*. Zeszyty Naukowe AGH, Sozologia i Sozotechnika (37), 1993
- [16] Husmeier D.: *Modelling conditional probability densities with neural networks*. Dostępne na stronie: citeseer.ist.psu.edu/husmeier97modelling.html, 1997
- [17] Husmeier D.: *Modelling conditional probabilities with network committees: how overfitting can be useful*. Dostępne na stronie: cite-seer.ist.psu.edu/article/husmeier98modelling.html, 1998
- [18] Jankowski N.: *Ontogeniczne sieci neuronowe. O sieciach zmieniających swoją strukturę*. Warszawa, Exit 2004
- [19] Krajewski R., Skawina T., Żuławski C.: *Hydrogeologiczno-glebowa metoda szacowania szkód w użytkach rolnych, wywołanych osuszającą działalnością górniczą*. Ochrona Terenów Górniczych, 1 (9), 1969
- [20] Kuncheva L.: *Switching between selection and fusion in combining classifiers: An experiment*. IEEE Transactions On Systems Man And Cybernetics, Part B-cybernetics, 32 (2), 2002, 146–156, dostępne na stronie: cite-seer.ist.psu.edu/kuncheva02switching.html
- [21] Kuncheva L., Bezdek J., Duin R.: *Decision templates for multiple classifier fusion: an experimental comparison*. Dostępne na stronie: cite-seer.ist.psu.edu/article/kuncheva99decision.html, 1999
- [22] McBratney A.B., Minasny B., Cattle S.R., Vervoort R.W.: *From pedotransfer functions to soil inference systems*. Geoderma, 109 (1–2), 2002, 41–73
- [23] Pal S.K., Mitra P.: *Pattern Recognition Algorithms for Data Mining. Scalability, Knowledge Discovery and Soft Granular Computing*. Boca Raton, London, New York, Washington, D.C., Chapman and Hall/CRC Press Company 2004
- [24] Principe J., Euliano N.R., Lefebvre W.: *Neural and Adaptive Systems: Fundamentals Through Simulations*. New York, John Wiley & Sons, Inc. 2000
- [25] Rossiter D.G.: *Methodology for soil resource inventories. 2-nd revised version*. Raport inżynierski, International Institute for Aerospace Survey & Earth Sciences (ITC), Lecture Notes & Reference, Enschede, Netherlands, 2000

- [26] Rozporządzenie Ministra Środowiska z dnia 9 września 2002 r. w sprawie standardów jakości gleby oraz standardów jakości ziemi. Warszawa, 2002, Dz.U. 02.165.1359 z dnia 4.10.2002 r.
- [27] Skawina T., Trafas M., Żuławski C.: *Klasyfikacja stosunków wodnych gleb dla potrzeb oceny i prognozowania szkód górniczych*. Wydawnictwa PTG 1972
- [28] Strzemiński M.: *Racjonalne użytkowanie ziemi w polskiej kartografii gleboznawczej*. P (13). Warszawa, Państwowe Wydawnictwo Rolnicze i Leśne i Instytut Uprawy Nawożenia i Gleboznawstwa 1966
- [29] Trafas M. : *Wpływ poeksploatacyjnych deformacji powierzchni na zmiany w glebach i sposobie ich użytkowania*. Kraków, AGH 1969 (praca doktorska)
- [30] Wolpert D.H.: *Stacked generalization*. Raport instytutowy LA-UR-90-3460, Los Alamos, NM, dostępne na stronie: citeseer.ist.psu.edu/wolpert92stacked.html, 1990
- [31] Zhu A.-X.: *Mapping soil landscape as spatial continua: The neural network approach*. Water Resources Research, 36 (3), 2000, 633–677
- [32] Zhu A.-X., Hudson B., Burt J., Lubich K., Simonson D.: *Soil mapping using GIS, expert knowledge, and fuzzy logic*. Soil Science Society of America Journal, 65 (1), 2001, 1463–1472