LUCRĂRILE SIMPOZIONULUI "SISTEME INFORMAȚIONALE GEOGRAFICE", NR. 10.
Anal. Şt. Univ. „Al.I.Cuza" Iaşi, Tom L, s.IIc. Geografie, 2004

GEO-MONITORIZAREA UNEI REGIUNI

Laura Ciocoiu, Cristian Paraschiv, Mihai Ioniță, Dragoş Barbu

Introducere

Geografia este o disciplină interactivă, aşa că datele geografice obligatoriu deschid o perspectivă întinsă şi interese, de la mediul social la cel fizic şi la toate punctele dintre ele.

- Complexitate asociată cu volumul de date.
- Complexitate asociată cu domeniul în sine.

Semnalele relevante şi interesante din continutul datelor sunt uneori în întregime ascunse de forme puternice care mai întâi trebuie scoase. Multe din aceste complexități au ca origine spațiul atât cel actual cât şi cel temporar care se asociază cu varietatea de game rezultată dintr-o varietate de cauze.

Complexitate cauzată de varietatea locală. Sistemele Pamântului sunt conectate atît de complex, încât este dificil să izolezi o analiză realizată într-o anumită parte a sistemului de efectele altor aspecte nemodelate.

Complexitate cauzată de strângerea datelor şi exemplelor. Deşi datele sunt variabile şi au un volum mare, este încă întâlnit cazul în care se recurge la înlocuire în detrimentul fenomenelor de interes, mai degrabă decât măsuri directe.

Dificultăți în formalizarea domeniului geografic. Una din principalele dificultăți cu activitatea de descoperire a datelor din domeniul geografic este însăşi comlexitatea domeniului de lucru.

Datele geo-spațiale conțin informații geometrice sau topologice, în comparație cu noile tipuri de date cum sunt factorii de mediu sau explorarea resursele naturale.

1. Procesarea cunoştințelor geografice

Tehnologia descoperirii cunoştințelor (Knowledge discovery) înglobează dezvoltarea unei noi generații de management la bazelor de date şi de sisteme informaționale prin abilitățile sale de a extreage noi informații inglobate în baze de date mari şi pentru a formula cunoştințe noi.

În particular, înțelesul datelor se concentrează spre creerea de algoritmi pentru extragerea de noi date statistice, modelarea rețelelor neurale, vizualizarea în vederea clasificării datelor şi identificării formelor. Descoperirea cunoştințelor ţinteste spre a activa un sistem de informații în vederea transformării informațiilor în cunoştințe prin testarea de ipoteze şi formularea de teorii.

Înțelesul caracteristic al datelor se referă la cazul în care datele sunt prea mari sau mult prea complexe pentru a permite amaliza manuală ori analiza facută de om prin simplă întrebare.

Înțelesul datelor este alcătuit din doi mari paşi:

- O analiză inițială a datelor, în cursul cărora relevante sunt caracteristicile de nivel înalt sau atributele sunt extrase din datele de nivel scăzut.

- Recunoaşterea de forme, în care este recunoscută forma din date folosind anumite caracteristici ale acesteia.

Analiza inițială a datelor este primul pas. Pentru a asigura un succes al procesului de înțelegere a datelor, este important ca caracteristicile extrase din date să fie relevante pentru rezolvarea problemei și să fie reprezentative pentu date.

În funcție de tipul datelor care sunt înțelese, pașii de analiza inițială constau în câteva sub-teme: datele sunt mari, statistic referindu-se la debit, se simplifică și se lucrează doar cu câteva instanțe. În cazul prognozei de recoltă, unde datele provin din mai multe surse, alegem fuziunea datelor cu scopul de a da voie să fie exploatate toate datele disponibile pentru această problemă.

La sfârșitul acestui prim pas, există o caracteristică vectorială pentru fiecare instanță de date. Sunt reduse numărul caracteristicilor folosind selecția de caracteristici(în cazul prognozei de recoltă) și reducerea dimensiunilor în cazul prognozei de debit.

Recunoașterea de forme după această pocesare inițială, datele sunt gata pentu recunoșterea de forme prin folosirea algoritmilor cum sunt rețele neurale și baze de cunoștințe utilizând agenți inteligenți). Aceste forme sunt ulterior afișate utilizatorilor pentru validare. Înțelegerea datelor este un proces interactiv și de repetare. Scoaterea oricăriu pas din domeniu bazelor de cunoștințe rezultă într-o repetare rafinată a oricărei sub-teme.

2. Metode de realizare a conceptului de PROGOZĂ

Acest concept a devenit o realitate odată cu realizarea marilor industrii informatice. Pornește de la o bază cu date științifice, pe care ulterior analizâdu-le, se ajunge la transformarea lor în concepte de preconizare și prognoză. Folosite în domenii diferite pot atinge standarde mai ridicate în dezvoltara economică umană și industrială.

În partea de geografie s-a pornit de la prognoza meteorologică care este definiția cea mai simplă a procedeului. Prin preluarea de date statistice de pe perioade mari de timp împreună cu folosirea hărților sinoptice se poate realiza o prognoză pe o perioadă de timp de la 1 zi pană la o săptămână, în care exactitatea poate varia. Cu cât perioada de timp este mai îndepărtată de prezent, veridicitatea datelor prognozate scade deoarece condițiile climatice pot varia foarte ușor într-un timp foarte scurt. Pornind de la aceleași considerente, se ajunge și la o interpretare de date geografice și o prognonoză pe urmatoarele ramuri:

2.1. Prognoză de debit pe un râu.

Pentru a se realiza o prognoză de debit în vederea cunoașterii eventualelor evenimente care se produc în unitatea de timp pe o rețea hidrografică, analiza, prelucrarea și interpretarea datelor se fac prin mijloace informatice dar și cu ajutorul unor informații. Baza de cunoștințe variază în funcție de necesitate și de temă. Pentru acest proiect au fost preluate date atât numerice (date statistice de debite și precipitații, preluate de la Institutul Național de Meteorologie și Hidrologie), date generale și actuale ale zonei analizate, cunoștințe fizico-geografice ale terenului din unitatea hidrografică dar și din imediata vecinătate a acesteia.

Cunoașterea reliefului este primordială în vederea obținerii de rezultate optime, astfel că hărțile topografice neactuaizate nu mai reprezintă o sursă de date suficientă în analiză spațiului dar combinate cu imagini satelitare și aerofotograme și importate în programe de anliză și prelucrare (Erdas Imagine, Arcview) duc la obținerea rezultatelor dorite.

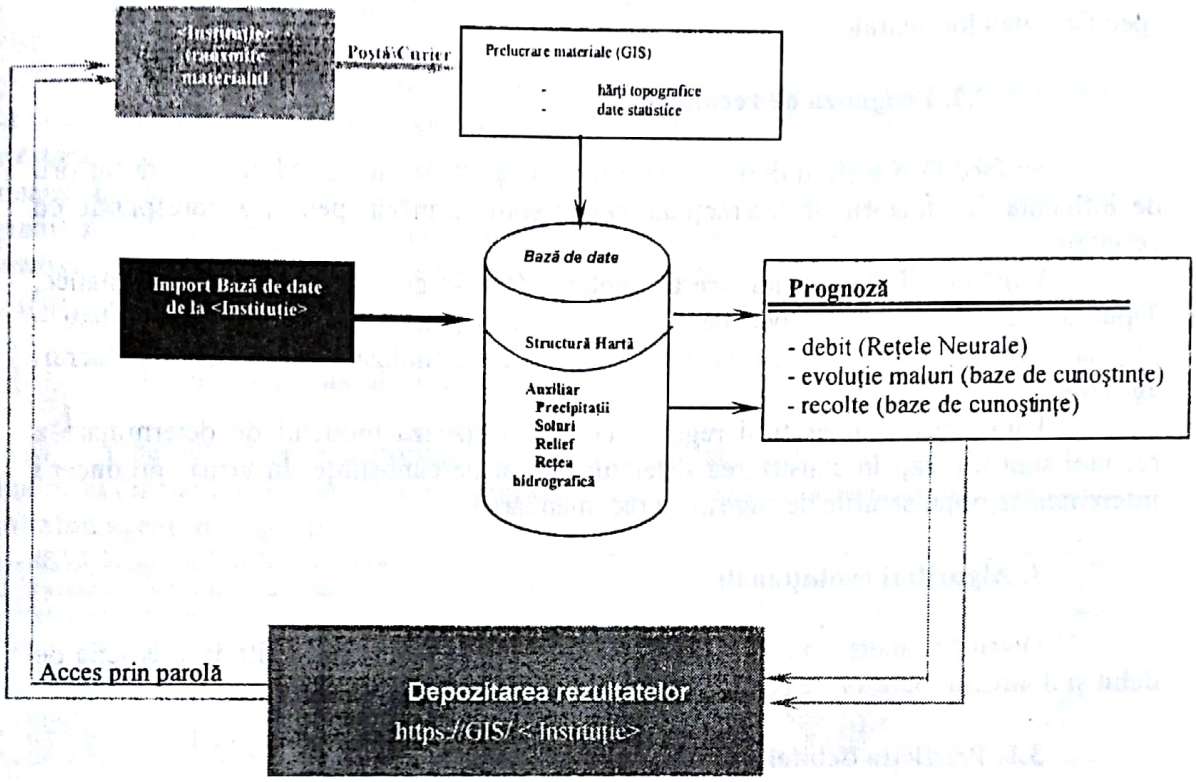


Figura 1 Schema fluxului informațional în sistemul de Geomonitorizare a unei regiuni

La realizarea acestei prognoze trebuie analizați şi factorii care acționează pe rețeaua hidrografică.

Factori de referință în stabilirea şi determinarea modificărilor pe rețeaua hidrografică

Factori naturali	Factori antropici	
	Direcți	Indirecti
Cantitatea de precipitații cazută în unitatea de timp	Defrişări pe maluri(rupturi, alunecări)	
Debitul înregistrat la postul hidrologic cel mai apropiat(analiză pe valori lunare)	Construcții de amenajare	Infiltrații deversăriă
Infiltrații afluenți	Modificări de curgere ale cursurilor permanente şi temporare	Defrişări pe versanți (maluri) rețelei hidrografice
Poziția geografică		
Structura albiei majore		
Structura de talveg, structura geologică	Modificări ale cantității de apă din alcatuire(deversări, barări, secări)	
Structura geologică totală pe rețea		
Fenomene geomorfologice actuale(alunecări de teren, densitățe, adancime fragmentăre)		
Modificări climatice puternice(îngheț-dezgheț, diferențe mari de temperatură de la zi la noapte, vânturi, inversiuni termice)		

După analiza tuturor factorilor enumerați mai sus şi unificarea datelor rezultate cu cele obțimute din analiza materialelor fotografice şi de hartă se ajunge la exprimarea

datelor de intrare necesare utilizării algoritmilor *feed forward şi backpropagation* specifici reţelelor neurale.

2.2. Prognoza de recoltare

Se face în principiu după aceleaşi metode şi etape, numai că aici diferă factorii de influenţă sau factorii de interacţiune care trebuie urmăriţi pentru a corespunde cu cerinţele.

Parametrii de prognozare a recoltei : Pozitia geografică, Condiţii climatice, Tipul de sol, Irigaţii şi metode de irigare, Sămânţă plantată, Perioada de plantare – recoltare, Date statistice ale zonei, Suprafaţa de teren analizată, Tratamente şi lucrari agricole.

Parametrii enumeraţi şi regulile ce caracterizează modelul de determinare a recoltei sunt utilizaţi în construirea sistemului bazat pe cunoştinţe. În urma producerii inferenţei, se obţin seturile de sugestii şi recomandări.

3. Algoritmi evoluţionali

Discutăm acum două tipuri de algoritmi: reţele neurale folosite în predicţia de debit şi domeniul bazelor de cunoştinţe folosiţi în prognoza de recoltă.

3.1. Predicţia debitelor fluviale utilizând reţelele neurale

Predicţia fenomenelor naturale legate de scurgerea râurilor - debitele de diferite asigurări, inundaţiile, fenomenele de secare - constituie una din prioritătile hidrologiei în momentul de faţă. Pentru realizarea unor predicţii adecvate este necesară cunoaşterea modului de manifestare a fenomenului respectiv pe o perioadă îndelungată de timp, de regulă mai mare de 30 de ani.

Această lucrare prezintă o aplicaţie a reţelelor neurale la predicţia debitelor medii lunare lichide bazată pe algoritmul *Backpropagation* pe baza măsurătorilor efectuate la postul hidrometric Borzii Vineţi, situat pe râul Jiu, în Sud-Vestul României, în Depresiunea Petroşani. Bazinul hidrologic corespunzător acestei staţii are o arie de 1222 Km² şi o înălţime medie de 1135 m. Datele folosite sunt din perioada 1950 –1994.

Bazinul hidrografic Jiu, din zona Borzii Vineţi, are o importanţă deosebită deoarece se colectează apa dintr-o regiune problematică, mono-industrială, care se confruntă cu probleme particulare din punct de vedere al proviziilor de apă pentru populaţie şi al industriei carbonifere. În aval de această secţiune, râul Jiu traversează un alt depozit de cărbune; de-altfel traversează cel mai mare bazin carbonifer din România. Acesta alimentează cu apa platforma industrială din oraşul Craiova cât şi trei din cele mai mari termocentrale din România. Astfel, cunoscând dinainte capacitatea debitului a râului Jiu va favoriza un management mai bun al resurselor de apă din aceast bazin hidrografic.

Principalele obiective a geomonitorizării unei regiuni şi a vremii cuprind:

- Prognozarea recoltei bazată pe vreme
- Modelarea debitelor fluviale

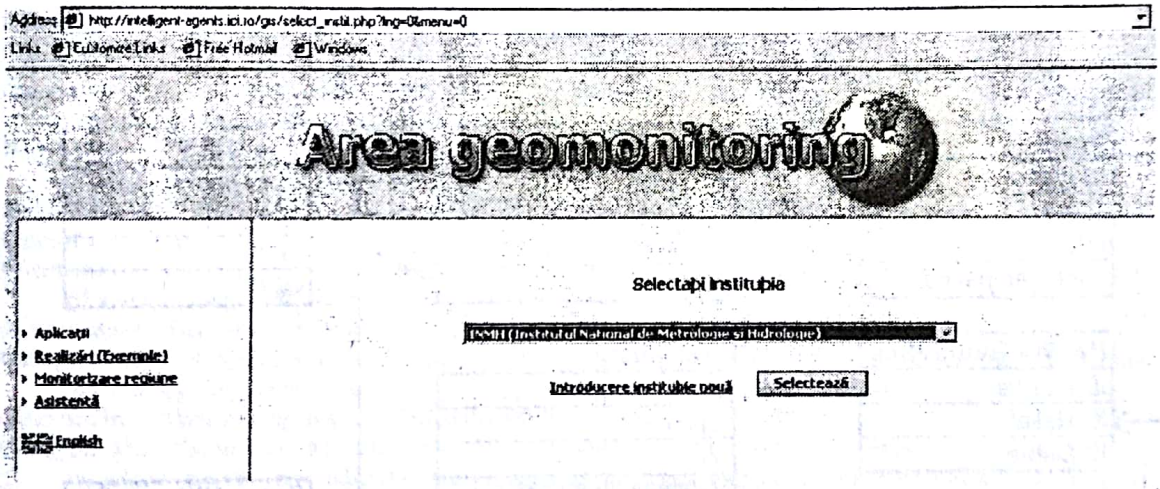
Prognozarea recoltei se realizează pe baza datelor geografice cum ar folosirea modelării bazelor de cunoştinţe (o clasificare ierarhică a claselor şi definiţiilor metodelor, descompunerea FBS, cooperarea între agenţii).

3.2. Predicția recoltei

Procesul de geomonitorizare a regiunii folosește sisteme de baze de cunoștințe ca o resursă pentru asistența specialiștilor și persoanelor în realizarea obiectivelor acestora. Designul modelului procesului reprezintă, de fapt, transferarea experienței umane în monitorizare într-un *model interactiv*. Cunoștințele despre procesul de geomonitorizare a unei regiuni (spre exemplu: interpretarea hărților, metode statistice, strategii, etc.) sunt reprezentate ca modele care fac referință la componentele observabile și importanța acestora.

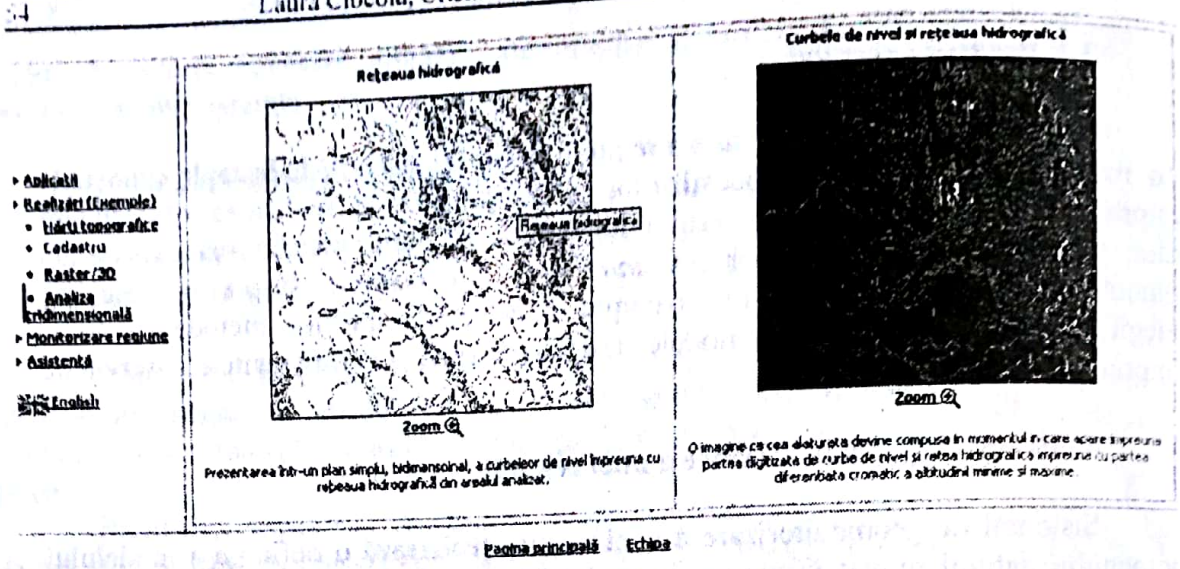
4. Sistem de geomonitorizare a unei regiuni

Sistemul de geomonitorizare a unei regiuni realizează o definiție a modelului fenomenului natural folosit în conceptia rețelelelor neurale și a bazelor de cunoștințe utilizând agenți inteligenți.



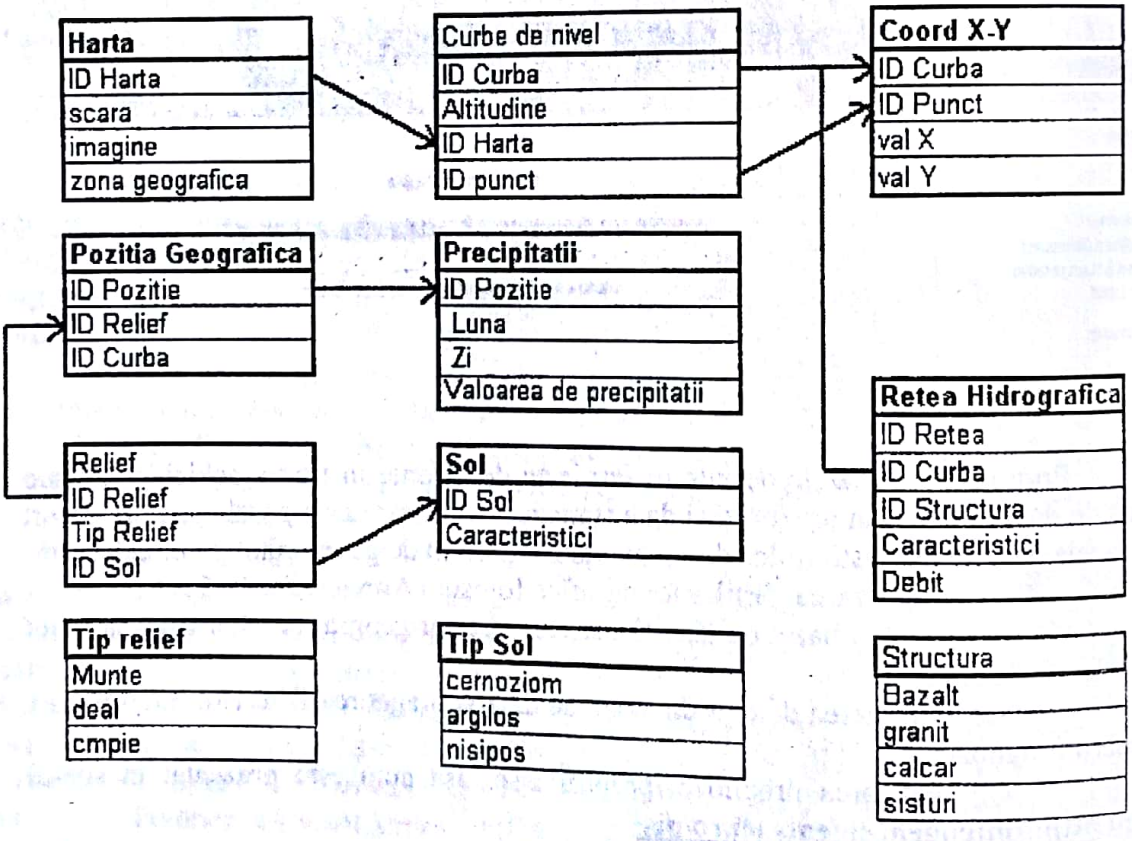
Pornind de la datele de intrare furnizate de diferite instituții, achiziționate sub formă de documente primare (hărți și date statistice) transmise prin poștă sau prin import din bazele de date ale instituțiilor de specialitate, sistemul de geomonitorizare realizează

- prelucrarea hărților topografice folosind Arcview sau Erdas
- crearea bazei de date (înscrierea datelor geografice referitoare la relief și datele statistice)
- extragerea datelor din baza de date și pregătirea fișierelor intermediare necesării prognozei
- realizarea prognozei propriu zise, așa cum este prezentat în site-ul nostru **http://inteligent-agents.ici.ro/gis/**



Structura bazei de date geografice este următoarea:

Baza de date geografica



Concluzii

Eforturile colaborative în inteligența artificială, statisticile şi bazele de date au fost făcute cu scopul de a modela fenomenul şi de a fi un suport de decizie în creere.

Baza geografică de date digitale este o simplă reprezentare a conoştințelor geografice la nivelul unor geometrii de bază, topologic şi constrângerea măsurătorilor.

Bazele de cunoştinţe GIS încearcă să creeze nu nivel mai ridicat de cunoştinţe geografice în baze geografice de date digitale pentru analiza complexă a fenomenului. Descoperirea cunoştinţelor geografice este o sursă pentru bazele de cunoştinţe GIS şi analiză inteligentă.

Următoarele proiecte vor dezvolta aplicaţii referitoare la :

- Interpretarea hărţilor şi extragerea automată a informaţiilor şi forme,
- Definiţia caracteristicilor de mediu (tehnici de învăţare inductive şi reţele neurale pentru clasificarea harţilor pe tipuri de soluri, vegetaţie etc.),
- Creerea de baze de date geografice

Acest articol a fost realizat de echipa din Institutul Naţional de Cercetare şi Dezvoltare în Informatică, pentru a fi înglobat la proiectul **INTAS** număr 397 „Tehnologia prelucrării datelor şi procesare a imaginilor: teorie şi aplicaţii" coordonat de Universitatea de Tehnologie din Lappeenranta, Finlanda. Vrem să mulţumim dnei **dr.mat.Angela Ioniţă, Director Adjunct al RACAI**, pentru munca sa şi pentru ajutorul dat în toată această perioadă.

BIBLIOGRAFIE

Brachman, R.J., Anand, T., 1996, "The Process of Knowledge Discovery in Databases", in Advances in Knowledge Discovery and data Mining, U.M. Fayyad et al. (eds.) AAAI Press/ The MIT Press

Buttenfield, B.P., 1998, Looking Forward: Geographic Information Services and Libraries in the Future. *Cartography and GIS*, 25(3);

Buttenfield, B., Gahegan, M., Miller, H., Yuan, M., "*Geospatial Data Mining and Knowledge Discovery*";

Cantu-Paz, Erik and Chandrika Kamath, "*On the use of evolutionary algorithms in data mining*";

Cheeseman, P. and Stutz, J., 1996, Bayesian Classification: Theory and results. In Eds. Fayyad, U., Piatetsky-Shapiro, G, Smyth, P. and Uthurusamy, R. (eds.) *Advances in Knowledge Discovery and Data Mining*, (Cambridge, MA: AAAI/MIT Press);

Gahegan, Mark: National Academies white paper. *Intersection of Geospatial Information and Information Technology*. September, 2001

Gahegan, M. , "Data mining and knowledge discovery in the geographical domain ";

Gahegan, M., Wachowicz, M., Harrower, M. and Rhyne, T. M., 2001, The integration of geographic visualization with knowledge discovery in databases and geocomputation. *Cartography and Geographic Information Systems*, special issue on the ICA research agenda.