

REZUMATUL TEZEI DE DOCTORAT

Contribuții la studiul utilizării și acoperirii terenului în evaluarea susceptibilității la inundații prin metode de teledetecție aerospațială

Autor: Ing. Ioana PAVEL (VIZIREANU)

Email: vizireanu.yoana@gmail.com, *tel:* +40 746 957 328

Conducător de doctorat: Col.(r) Prof. Univ. Dr. Ing. **Dan-Gabriel RĂDUCANU**

Prezenta teză de doctorat abordează evaluarea susceptibilității la inundații la scară locală prin integrarea datelor geospațiale de înaltă rezoluție și a metodelor moderne de învățare automată, în vederea analizei proceselor hidrogeomorfologice din mediul rural. Demersul se înscrie în contextul intensificării fenomenelor hidrometeorologice extreme și al transformărilor accelerate ale utilizării și acoperirii terenului, care influențează direct procesele de infiltrație, scurgere de suprafață și acumulare a apei, făcând necesară identificarea și cuantificarea predispoziției la inundare.

Cercetarea se desfășoară în comuna Mihăești, județul Argeș, situată în bazinul hidrografic al râului Târgului, un areal reprezentativ pentru condițiile geomorfologice și funcționale specifice sectorului subcarpatic din județul Argeș. Alegerea zonei de studiu are în vedere variabilitatea morfologică, structura heterogenă a utilizării terenului și prezența unor sectoare favorabile acumulărilor locale de apă, aspecte care permit investigarea detaliată a relațiilor dintre factorii de mediu și susceptibilitatea la inundare.

Obiectivul general al cercetării îl reprezintă dezvoltarea și aplicarea unei metodologii integrate pentru evaluarea susceptibilității la inundații la scară locală, bazată pe utilizarea datelor de teledetecție de înaltă rezoluție și a algoritmilor de învățare automată, în corelație cu metode de interpretare a modelelor predictive. În acest scop, sunt urmărite **obiective specifice** precum elaborarea unei baze de date geospațiale multisenzor, derivarea variabilelor morfometrice și tematice relevante, definirea variabilei dependente prin metode hidrogeomorfologice automate, implementarea și compararea unor algoritmi de clasificare, evaluarea performanței predictive, analiza interpretabilității rezultatelor, validarea spațială independentă și integrarea rezultatelor în analiza mediului construit.

Lucrarea pornește de la limitările abordărilor bazate pe modele altimetrice generalizate sau pe date de rezoluție medie, care nu surprind adecvat heterogenitatea spațială fină a factorilor determinanți ai acumulării apei. În acest context, teza propune utilizarea unui set integrat de date geospațiale de înaltă rezoluție, care include date LiDAR aeropurtate (densitate medie de aproximativ 8 puncte/m²), model digital al terenului derivat din date LiDAR la rezoluție de 1 m, imagini multispectrale Sentinel-2 (10 m), ortofotoplanuri de înaltă rezoluție (0,2 m) și date vectoriale tematice.

Prin integrarea acestor surse se elaborează o bază de date geospațială multisenzor care include 8.982.391 unități spațiale elementare, fiecare caracterizată printr-un set coerent de variabile explicative, incluzând altitudine, pantă, curbură, aspect, distanță față de rețeaua hidrografică, clase de utilizare a terenului și indicele NDVI. Variabila dependentă se definește prin analiza hidrogeomorfologică de tip Bluespot, utilizată pentru identificarea automată a depresiunilor topografice cu potențial de stagnare a apei pe baza microreliefului derivat din modelul digital al terenului. Zonele susceptibile astfel identificate reprezintă 5,53% din suprafața analizată, evidențiind un dezechilibru semnificativ al claselor și justificând utilizarea unor tehnici adecvate de modelare.

În etapa de modelare se implementează și se compară trei algoritmi de clasificare: Random Forest, XGBoost și regresia logistică. Antrenarea modelelor se realizează pe un eșantion stratificat de 1.000.000 observații, iar evaluarea finală pe un set de test independent de 300.000 observații. Dezechilibrul claselor, caracterizat printr-un raport aproximativ de 1:17, se tratează prin tehnica SMOTE, iar optimizarea hiperparametrilor se realizează prin validare încrucișată. Robustețea rezultatelor se analizează prin utilizarea a două strategii de împărțire a datelor, respectiv 70–30, în care 70% din date sunt utilizate pentru antrenare și 30% pentru testare, și 80–10–10, în care 80% din date sunt utilizate pentru antrenare, 10% pentru validare și 10% pentru testare. Compararea acestor strategii evidențiază coerența rezultatelor și stabilitatea performanței modelului, diferențele observate fiind minore și fără impact semnificativ asupra performanței globale.

Evaluarea performanței predictive evidențiază superioritatea modelului Random Forest, care obține valori ridicate ale indicatorilor statistici ($AUC-ROC = 0,868$; acuratețe = 83,9%; precizie = 20,5%; rată de detecție = 66,5%; scor F1 = 0,313; MCC = 0,306), confirmând capacitatea acestuia de a modela relațiile complexe și non-liniare dintre variabilele de mediu și predispoziția la inundație. Pe baza modelului optim se generează harta susceptibilității la inundații, clasificată în cinci clase, care evidențiază predominanța nivelurilor scăzute și concentrarea claselor ridicate în zonele joase și în proximitatea rețelei hidrografice. Aceste rezultate confirmă că integrarea datelor LiDAR de înaltă rezoluție cu modele de tip ansamblu, precum Random Forest și XGBoost, oferă un cadru robust și interpretabil pentru evaluarea susceptibilității la inundații la scară locală.

Un aspect distinctiv al cercetării îl constituie corelarea nivelului ridicat de detaliu spațial al datelor de intrare cu cerința de explicabilitate a rezultatelor obținute prin modele de învățare automată. Spre deosebire de abordările bazate exclusiv pe performanța statistică a clasificatorilor, demersul propus urmărește obținerea unei hărți predictive însoțite de interpretarea relațiilor dintre variabilele de mediu și susceptibilitate, consolidând valoarea aplicativă a rezultatelor.

Analiza interpretabilității modelelor reprezintă un element central al cercetării, realizată prin metoda SHAP, care permite cuantificarea contribuției fiecărei variabile explicative la rezultatul modelului și evidențierea relațiilor dintre factorii analizați și susceptibilitate. Integrarea acestei metode permite interpretarea rezultatelor în termeni fizico-geografici, facilitând înțelegerea mecanismelor care determină predispoziția la acumulări de apă. Rezultatele evidențiază rolul dominant al altitudinii, al pantei, al claselor de utilizare a terenului și al distanței față de rețeaua hidrografică, confirmate prin metodele de evaluare a importanței variabilelor (SHAP, Gini și importanța prin permutare). De asemenea, analiza SHAP permite identificarea unor praguri critice ale variabilelor, relevante pentru delimitarea zonelor cu susceptibilitate ridicată.

Validarea modelului se realizează prin metode spațiale independente, incluzând analiza pe zone concentrice de distanță față de rețeaua hidrografică și analiza comparativă Nord–Sud, care demonstrează coerența geografică a rezultatelor și capacitatea modelului de a reflecta distribuții spațiale realiste ale susceptibilității. Aceste metode depășesc evaluarea statistică clasică, oferind o perspectivă asupra comportamentului modelului în raport cu structura spațială a fenomenului analizat. Integrarea rezultatelor într-o analiză aplicativă privind mediul construit evidențiază distribuția spațială a construcțiilor în raport cu clasele de susceptibilitate, permițând identificarea sectoarelor cu expunere ridicată la acumulări de apă.

Contribuțiile principale ale tezei constau în dezvoltarea unui cadru metodologic integrat pentru evaluarea susceptibilității la inundații la scară locală, fundamentat pe utilizarea combinată a datelor de teledetecție de înaltă rezoluție și a metodelor de învățare automată. Un prim element de noutate îl reprezintă utilizarea sistematică a datelor LiDAR pentru caracterizarea microreliefului la rezoluție metrică, care permite identificarea detaliată a

depresiunilor topografice și a variațiilor subtile de pantă relevante pentru acumularea apei. În același timp, integrarea informațiilor privind utilizarea și acoperirea terenului la rezoluție submetrică conduce la o reprezentare mult mai fidelă a condițiilor de suprafață, comparativ cu abordările bazate pe produse generalizate.

Un aport important constă în evaluarea comparativă a algoritmilor de învățare automată într-un cadru unitar, care permite evaluarea performanței acestora în condițiile unui set de date complex și dezechilibrat. Analiza comparativă evidențiază avantajele modelului Random Forest în captarea relațiilor non-liniare dintre variabilele de mediu și susceptibilitate. În completare, utilizarea tehnicilor de interpretabilitate, în special metoda SHAP, reprezintă o contribuție semnificativă, deoarece permite explicarea rezultatelor modelului și corelarea acestora cu procesele fizice reale.

De asemenea, teza aduce contribuții prin integrarea unei validări spațiale independente, care permite verificarea coerenței geografice a rezultatelor. Evaluarea influenței modului de configurare a seturilor de date evidențiază stabilitatea modelului și impactul acestei configurări asupra rezultatelor obținute. Integrarea finală a rezultatelor într-un cadru aplicativ, prin raportarea la mediul construit, completează demersul metodologic și evidențiază utilitatea practică a modelului dezvoltat.

Prin aceste elemente, teza propune un model coerent, reproductibil și adaptabil, aplicabil în analiza susceptibilității la inundații în contexte similare caracterizate prin variabilitate spațială ridicată și disponibilitatea datelor de înaltă rezoluție.

Direcțiile viitoare de dezvoltare ale cercetării vizează extinderea cadrului metodologic prin testarea unor algoritmi suplimentari de învățare automată, precum algoritmi de tip Support Vector Machine (SVM) sau modele bazate pe rețele neuronale convoluționale de tip CNN, în vederea evaluării capacității acestora de a gestiona seturi de date caracterizate prin dezechilibru pronunțat al claselor. De asemenea, cercetările ulterioare pot include utilizarea unor indicatori suplimentari de evaluare, precum aria de sub curba Precision–Recall (PR-AUC), complementară indicatorului AUC–ROC, precum și identificarea unui prag optim de clasificare pentru îmbunătățirea echilibrului dintre rata de detecție și precizie.

În ansamblu, cercetarea contribuie la dezvoltarea metodelor moderne de evaluare a susceptibilității la inundații prin integrarea datelor geospațiale de înaltă rezoluție cu tehnici avansate de modelare și interpretare. Cadrul metodologic propus are aplicabilitate directă în sprijinirea proceselor de planificare teritorială și în identificarea zonelor cu predispoziție ridicată la inundare, în special în contexte caracterizate prin variabilitate spațială ridicată.