

# Inventaire des dolines supposées dans le Doubs par apprentissage automatique des dépressions détectées par traitement du LiDAR HD de l'IGN

## Inventory of Suspected Dolines in the Doubs Using Machine Learning on Depressions Detected from IGN High-Resolution LiDAR

Margaux FLIPO<sup>1\*</sup>, Florian MASSON<sup>2</sup>, Daniel BLANCHARD<sup>2</sup>, Gildas NOURY<sup>2</sup>, Eglantine HUSSON<sup>2</sup>, Aurélien VALLET<sup>1</sup>

<sup>1</sup>*Bureau de Recherches Géologiques et Minières (BRGM), Dijon, France*

<sup>2</sup>*Bureau de Recherches Géologiques et Minières (BRGM), Orléans, France*

### Résumé

Les paysages karstiques sont souvent marqués par d'innombrables dolines qu'il est impossible de cartographier par une inspection de terrain. Le traitement du Modèle Numérique de Terrain (MNT) issu des levés LiDAR HD permet désormais de s'affranchir de cette difficulté. Les levés de l'IGN couvrant le Doubs (25) ont été utilisés par le BRGM dans le cadre d'une étude visant à y évaluer, à la demande de la Direction Départementale des Territoires, les risques de mouvements de terrain.

Les méthodes préexistantes, qui repèrent les dépressions fermées, ont été améliorées à trois niveaux : 1) les dépressions « ouvertes » sur une pente ont été repérées en ajoutant une opération lissant le MNT ; 2) les dépressions coalescentes ont été séparées les unes des autres via un test sur les altitudes des polygones extraits ; 3) enfin un apprentissage automatique (machine-learning) par la technique de la forêt d'arbres décisionnels a été mis en œuvre pour écarter les dépressions d'origine non naturelle (fossés, descentes de garage, etc.). Cette technique consiste à entraîner un algorithme à catégoriser une entité (ici doline supposée ou « pas doline » supposée) en fonction d'un certain nombre de paramètres a priori prédictifs (ici : les paramètres de forme : dimensions, circularité, élongation, etc. et les caractéristiques environnementales (distance aux habitations, aux routes, etc...))

2 234 dépressions réparties sur 43 km<sup>2</sup> et sur 3 zones géologiquement différentes et aux occupations de sol variables (forêt, plaine, ville) constituent la population d'entraînement, avec 70% de dépressions non naturelles et 30 % naturelles. L'algorithme arrive in fine à un taux de prédiction de 96%, ce qui permet de l'utiliser à plus large échelle et de compléter les inventaires existants pour le département : environ 220 000 dépressions sur les 1,5 millions détectées sur le département sont actuellement considérées comme des dolines supposées.

### Abstract

Karst landscapes are often characterized by countless dolines, making exhaustive field mapping impractical. The processing of high-resolution LiDAR-derived Digital Elevation Models (DEMs) now makes it possible to overcome this limitation. LiDAR datasets acquired by the French National Mapping Agency (IGN) over the Doubs department (25) were used by the BRGM in a study commissioned by the Departmental Directorate of Territories to assess ground-movement hazards.

Pre-existing methods for detecting closed depressions were improved at three levels: (1) depressions "open" on a slope were identified by incorporating a DEM smoothing operation; (2) coalescent depressions were separated using an altitude-based test applied to the extracted polygons; and (3) a machine-learning approach based on a decision-tree ensemble was implemented to filter out depressions of non-natural origin (ditches, driveway slopes, etc.). This approach consists in training an algorithm to

classify an entity (here, presumed doline or presumed non-doline) based on a set of a priori predictive parameters, including shape descriptors (dimensions, circularity, elongation, etc.) and environmental characteristics (distance to buildings, roads, etc.).

A total of 2,234 depressions distributed over 43 km<sup>2</sup>, across three geologically distinct areas with contrasting land-use types (forest, plain, urban), were used as the training dataset, comprising 70% non-natural and 30% natural depressions. The algorithm ultimately achieves a prediction accuracy of 96%, allowing its application at a broader scale and enabling the completion of existing inventories for the department. Approximately 220,000 of the 1.5 million depressions detected across the Doubs are currently classified as presumed dolines.

**Mots-clés:** Dolines ; machine-learning ; Lidar HD ; karst.

## 1. INTRODUCTION

Les dolines sont des dépressions fermées d'origine naturelle, liées à la karstification des paysages (dissolution de certains types de roches, notamment carbonatées). Leur diamètre varie de quelques mètres à plusieurs centaines de mètres et leur profondeur peut atteindre plusieurs mètres. Elles constituent des points privilégiés d'infiltration des eaux (Figure 1). Elles sont des indicateurs clés du fonctionnement des systèmes karstiques : les recenser et les caractériser est essentiel pour comprendre l'histoire et le fonctionnement du karst, notamment en termes hydrogéologiques et pour la gestion des risques naturels. Les dolines signalent en effet la présence de vides ou de zones fragiles en profondeur. Des effondrements peuvent survenir soudainement et menacer les infrastructures, habitations, routes ou exploitations agricoles. Une cartographie fine et exhaustive des dolines permet ainsi d'identifier les zones sensibles qui devront faire l'objet d'attentions particulières, comme des investigations géotechniques préalables à la construction.



Figure 1 : Photographie d'une doline à Montrond-le-Château (Doubs, France)

La cartographie des dolines peut être réalisée selon différentes méthodes offrant des niveaux variables de précision et d'automatisation. Dans des régions telles que le Jura où elles sont très nombreuses, le repérage

par visite systématique est rapidement limité. Le pointage manuel de cartes topographiques, de Modèles Numériques de Terrain (MNT) ombrés ou d'orthophotographies a été longtemps utilisé, mais cette approche est chronophage et difficile à appliquer à grande échelle. Plus récemment, des méthodes semi-automatisées permettent de repérer automatiquement toutes les dépressions présentes en surface avec un tri entre dépressions naturelles et dépressions anthropiques (mares aménagées, fossés, descentes de garage, etc.) souvent réalisé à dire d'expert, aboutissant à des résultats moyennement satisfaisants en terme d'exhaustivité sur de grands territoires. C'est pourquoi, l'intelligence artificielle constitue une avancée majeure : cette technique permet de traiter simultanément un grand nombre de variables morphométriques et de s'adapter à la diversité des formes karstiques. Les modèles supervisés, tels que les forêts aléatoires constituent un sous-ensemble de l'intelligence artificielle. Ils permettent d'ajuster automatiquement les combinaisons de critères les plus discriminants. Cette approche réduit les biais liés à la subjectivité, améliore la robustesse de la détection et permet ainsi de recenser de manière fiable les dolines en écartant les dépressions qui ne sont pas naturelles.

Le Doubs ( $\approx 5\,200$  km<sup>2</sup>), avec ses calcaires jurassiques, sa tectonique structurée en plis et failles, son héritage quaternaire et sa topographie de plateaux et massifs, présente l'une des plus fortes concentrations de dolines en France (Camus, 2003). Un Atlas d'indices karstiques a déjà été réalisé en combinant des données ponctuelles (Cerema, 2013). Depuis, le Doubs est couvert par un MNT haute résolution produit par l'IGN, ce qui en fait le candidat idéal pour développer une méthode de détection automatique de dolines à grande échelle.

L'outil développé a pour objectif de produire un inventaire automatisé et quasi exhaustif des formes karstiques de surface à partir d'un MNT à haute

résolution. Il se concentre sur la détection des dolines, définies par une morphologie pseudo-circulaire à elliptique (diamètre  $\geq 1$  m, profondeur  $\geq 0,5$  m). Les métadonnées (analyses morphométriques, spatiales et environnementales) sont valorisées dans un modèle d'apprentissage automatique basé sur la méthode d'arbres décisionnels, afin de classer les objets et d'identifier les dolines, tout en permettant un traitement efficace de grandes surfaces comme le département du Doubs.

## 2. ZONE D'ETUDE

Le département du Doubs (5 232 km<sup>2</sup>) présente une géologie dominée par les calcaires jurassiques plissés et faillés (Figure 2c) et un relief variant des plaines alluviales aux plateaux et aux massifs dont l'altitude approche 1 000 m (Figure 2b). Les calcaires peuvent être nus ou recouverts de dépôts quaternaires (générant un karst sous couverture). Les aquifères jurassiques incluent des résurgences, des pertes et sont caractérisés par des circulations rapides, ce qui rend leur alimentation et leur sensibilité à la pollution très variables, nécessitant une connaissance précise des points d'infiltration (dolines, pertes).

Six zones représentatives des différents contextes géologiques, géomorphologiques et d'occupation du sol du département ont été sélectionnées pour calibrer et tester la détection (Figure 2b). La zone test n°1, située sur le plateau de Montrond-le-Château, s'étend

sur les calcaires massifs du Jurassique supérieur, avec de nombreuses dolines profondes et bien structurées. La zone test n°2, aux alentours de Valdahon (Figure 2a), se caractérise par un karst de plateau recouvert localement de dépôts quaternaires discontinus, où les dolines présentent des morphologies plus diffuses. La zone n°3 située dans les Avant-Monts se caractérise par un relief de transition, intermédiaire entre la plaine et les hauts plateaux jurassiens. La zone test n°4, concerne le faisceau de Quingey, ensemble de plis serrés du Jura externe se caractérisant par un relief contrasté associant crêtes anticlinales, vallées encaissées et plateaux structuraux discontinus, ainsi qu'un karst bien développé fortement guidé par la fracturation. La zone test n°5, à mi-chemin entre le plateau de Levier et la Haute-Chaine (environ 900 – 1 100 m d'altitude), présente un karst influencé par les processus glaciaires et périglaciaires, avec des dolines larges et peu profondes souvent associées à des surfaces nivéo-karstiques (forme de relief karstique liée à la dissolution des roches calcaires qui se développe sous l'influence conjointe de la neige et du karst). Enfin, la zone test n°6, implantée en périphérie nord de Besançon, est constituée d'un karst sous couverture en milieu périurbain, avec une occupation du sol hétérogène (zones agricoles, boisées ou urbanisées). Ces six zones représentent la diversité des paysages karstiques du Doubs et permettent de tester la robustesse des algorithmes dans des environnements contrastés, tant du point de vue géologique que topographique et anthropique.

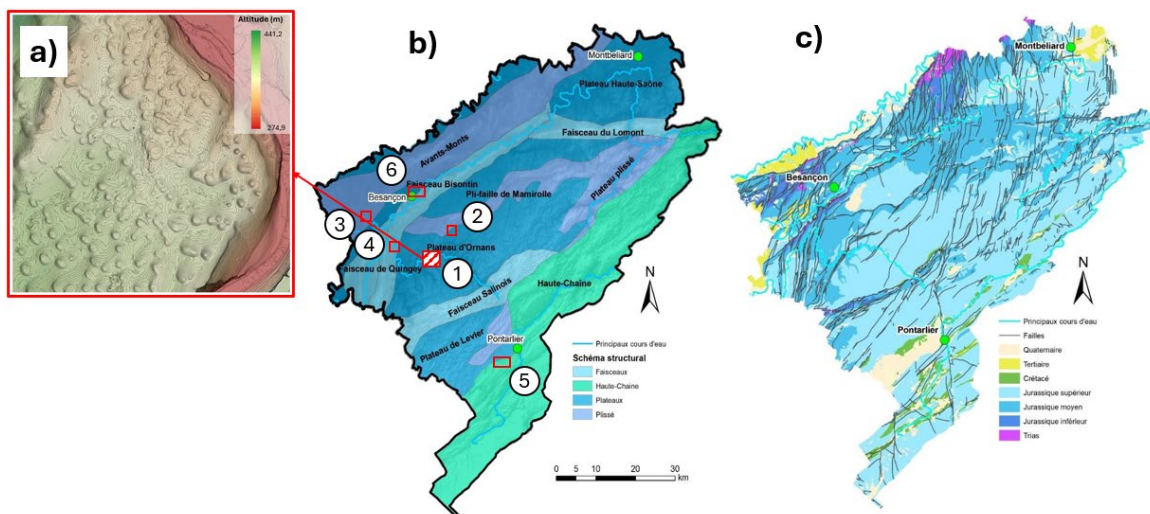


Figure 2 : a) Exemple de zone couverte par des dolines visibles sur le MNT, b) et c) représentations de la structure et de la géologie du Doubs tirées du rapport BRGM/RP-70682-FR, et zones sélectionnées pour la calibration de l'outil (encadrés en rouge).

### 3. DONNEES ET METHODE

La détection de dolines repose principalement sur l'utilisation du LiDAR HD de l'IGN (échantillonnage

moyen de 20 à 30 pts/m<sup>2</sup>), traité pour obtenir un MNT raster de résolution 50 cm. Des couches auxiliaires (BD TOPO : bâtiments, carrières ; BD CARTO : hydrographie) ont également utilisées.

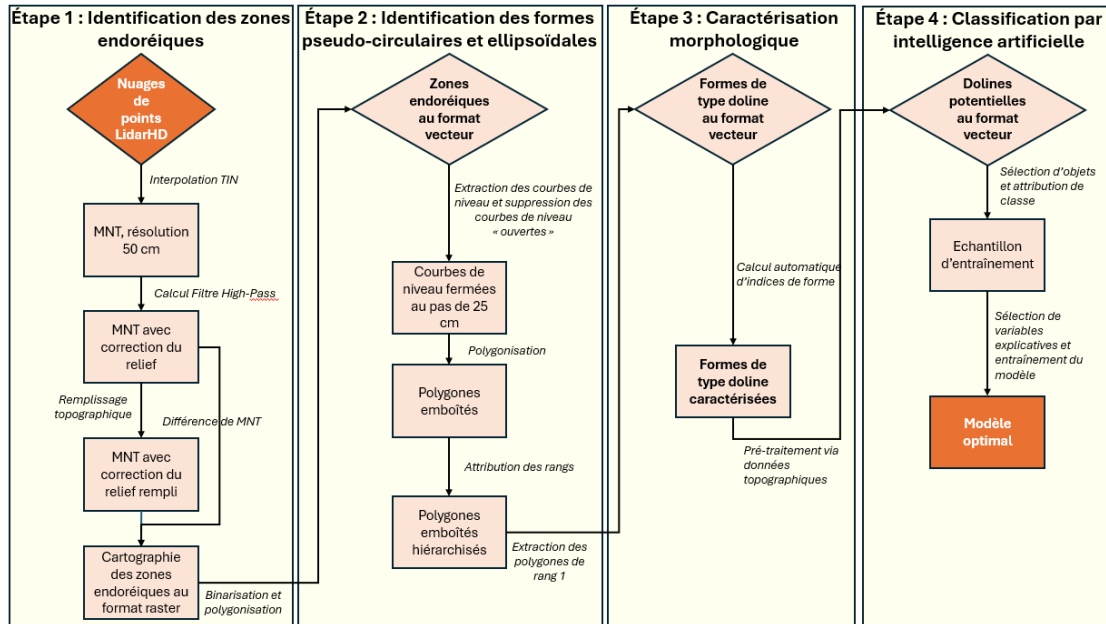


Figure 3 : Diagramme général de la méthode.

La détection des dolines sur le département s'articule en quatre étapes (Figure 3) :

- 1) Identification des dépressions topographiques à partir d'un MNT issu du LiDAR. La méthode repose sur une analyse altimétrique différentielle entre le MNT d'origine et sa version corrigée topographiquement remplie. Un prétraitement est réalisé afin de mettre en évidence les microformes et les dépressions topographiques ;
- 2) Extraction des formes pseudo-circulaires correspondant aux courbes de niveau fermées. Ces courbes sont organisées hiérarchiquement dans l'espace pour isoler celles qui correspondent potentiellement au pourtour d'une doline ;
- 3) Caractérisation morphométrique de ces polygones (algorithme BRGM) ;
- 4) Classification par intelligence artificielle pour distinguer les véritables dolines potentielles des dépressions a priori non naturelles, à partir d'un ensemble d'entraînement et de variables explicatives sélectionnées. Une validation croisée sur un ensemble de test est menée afin d'identifier le modèle optimal.

Le modèle présentant les meilleurs scores sur l'ensemble de test est retenu puis appliqué à l'ensemble du département.

#### 3.1 Détection des zones de dépression topographique.

La première étape consiste à identifier toutes les zones de dépression topographique dans le MNT. La méthode se base sur une analyse altimétrique différentielle, entre le MNT d'origine et sa version corrigée où toutes les altitudes ont été ajustées afin de permettre un écoulement continu de l'eau à la surface (remplissage topographique).

Le MNT d'origine fait l'objet d'un prétraitement de rééchantillonnage permettant de séparer les basses fréquences (relief régional) des hautes fréquences (variations locales). Le MNT haute fréquence ainsi obtenu met en évidence les microformes en éliminant les tendances générales du relief ce qui permet d'identifier notamment les zones de dépression topographique situées sur des versants en pente (pas identifiables sans ce traitement. Les zones de

dépression topographique identifiées sont ensuite polygonisées. Seules les zones présentant une profondeur supérieure à 50 cm sont retenues afin d'éliminer de potentiels artefacts (irrégularités du MNT, végétation résiduelle ou erreurs de mesure).

### 3.2 Détection des dolines potentielles

Cette étape consiste à repérer, parmi toutes les zones de dépression topographique détectées, celles qui ont réellement la forme de dolines (pseudo-circulaire, ellipsoïdale). Pour cela, une analyse basée

sur les courbes de niveau au pas de 25 cm du MNT lissé est réalisée : seules les courbes fermées sont conservées, puis transformées en polygones. Leur organisation est hiérarchisée selon le principe de l'arbre de contours développé par Kweon et Kanade (1994). Leur hiérarchisation, fondée sur les relations de contenance entre polygones, permet de reconstituer l'organisation interne des zones de dépression topographique : les polygones de rang 1 correspondent aux minima locaux, tandis que les rangs supérieurs délimitent des surfaces plus élevées. Les dolines potentielles sont définies comme les polygones de rang 1 situés à la plus haute altitude (Figure 4).

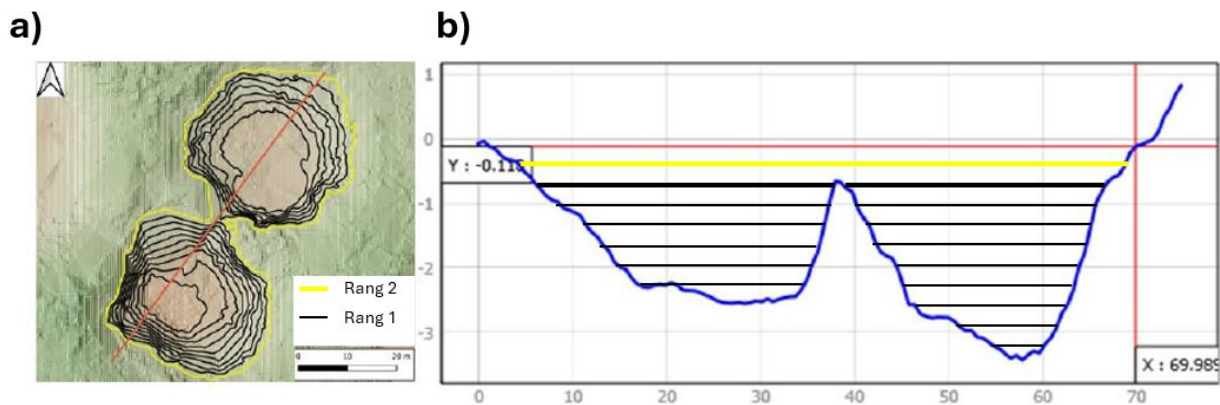


Figure 4 : Exemple d'attribution du numéro de rang aux courbes de niveau. a) vue en plan, b) vue en coupe. La courbe de niveau retenue comme « doline » est représentée en noir gras sur la coupe.

### 3.3 Caractérisation morphométrique

L'ensemble de ces polygones hiérarchisés sont caractérisés suivant leur géométrie. Pour chaque dépression, 22 variables sont calculées (Tableau 1) —

planimétriques (aire, longueur, largeur, circularité, rectangularité, dimension fractale), altimétriques (profondeur, altitude des pixels voisins, coefficients d'asymétrie et d'aplatissement, volume), topographiques (orientation interne, pente radiale, distance centroïde-fond) et géographiques (distance aux routes, distance aux bâtiments).

Tableau 1 : Caractérisation géométrique des dépressions repérées sur le MNT : exemples de neuf variables calculées

Type	Variable calculée	Description
Planimétrie	Longueur	Longueur du grand axe du rectangle englobant minimal de la dépression (axe majeur), exprimée en mètres.
	Indice de circularité	Indicateur de compacité basé sur le rapport entre l'aire et le périmètre ; plus il est proche de 1, plus la forme est régulière et circulaire.
	Aire	Surface totale de la dépression, exprimée en mètres-carrés.
	Indice de rectangularité	Rapport entre l'aire de la dépression et celle de son rectangle englobant minimal ; mesure le degré d'occupation de cet espace théorique.
Orientation et pente	Azimut	Orientation du grand axe du rectangle englobant minimal, exprimé en degrés (0-180°).
	Pente moyenne radiale (°)	Pente moyenne des flancs de la dépression, calculée depuis le point le plus bas vers la bordure de la dépression.
Altimétrie	Volume (m3)	Volume topographique correspondant à la quantité de matière nécessaire pour combler la dépression jusqu'au niveau moyen du terrain environnant.
	Coefficient d'asymétrie	Indicateur statistique (skewness) décrivant l'asymétrie de la distribution altimétrique : négatif pour un fond piqué, positif pour un fond étalé.

### 3.4 Classification des dolines potentielles

Il s'agit maintenant de trier les dolines de l'ensemble des 1 511 779 dépressions détectées précédemment, et d'écarter les dépressions d'origine non naturelle, telles que les fossés routiers, les descentes de garage ou autre aménagement urbain, les carrières, etc.

La superposition des dépressions avec les couches BD TOPO et BD CARTO permet d'éliminer 259 266 artefacts (dépressions intersectant bâtiments, carrières, plans d'eau et cours d'eau principaux, etc.).

Le modèle XGBoost (Extreme Gradient Boosting) proposé par Chen et Guestrin en 2016, a ensuite été retenu pour classer les formes détectées en « doline supposée » ou « non doline supposée » sur la base des variables morphométriques calculées précédemment. Cet algorithme d'apprentissage supervisé est fondé sur le *gradient boosting*, méthode qui combine plusieurs arbres de décision pour construire un modèle global performant. À chaque itération, il améliore les prédictions en corrigeant les erreurs des arbres de décision précédents, ce qui lui permet de modéliser efficacement des relations complexes en régression comme en classification. Sa robustesse face aux données bruitées est assurée par des mécanismes de régularisation et de contrôle de la complexité, limitant le sur-apprentissage (Friedman, 2001).

Un échantillon de 2 170 entités réparties sur les six zones test précédemment décrites a été constitué puis fait l'objet d'une classification experte à partir de l'analyse du MNT et de l'orthophotographie. A cet échantillon s'ajoutent 7 événements récents issus de la DDT25 et 57 observations de terrain. Au final, 760 objets sont classés en « doline supposée » et 1 474 objets en « non doline supposée ». Les dépressions couvrent les trois principales occupations des sols présentes sur le territoire : zones forestières (43 %), agricoles (36 %) et urbanisées (21 %).

Une analyse de corrélation (coefficient de Spearman) entre la classification experte et les variables morphométriques calculées précédemment permettent d'identifier les variables explicatives principales. Les variables fortement corrélées entre elles ( $> 0,9$ ) sont écartées pour réduire la redondance et optimiser le temps de calcul.

L'échantillon est ensuite divisé en un ensemble d'entraînement (80 % des données) servant à entraîner le modèle et un ensemble de test (20 % des données) servant à évaluer la performance des algorithmes sur des données non vues. Au total, 1 000 entraînements et évaluations indépendants ont été réalisés afin d'explorer de manière robuste l'influence des paramètres intrinsèques au modèle (hyperparamètres) et du découpage des données. Cette procédure combine la variabilité liée au réglage des hyperparamètres et celle induite par les ensembles d'apprentissage et de validation. L'objectif est de sélectionner le meilleur classificateur capable de distinguer efficacement dolines et objets non karstiques, tout en évitant le surapprentissage.

Cinquante combinaisons distinctes d'hyperparamètres du modèle XGBoost sont donc générées aléatoirement. Pour chacune de ces combinaisons, une validation croisée répétée en 20 essais est appliquée : à chaque essai, l'ensemble d'entraînement est redécoupé aléatoirement en sous-ensembles d'apprentissage et de validation, sur lesquels le modèle est respectivement ajusté puis évalué. Les performances obtenues au cours des 20 essais sont ensuite moyennées pour chaque combinaison d'hyperparamètres. L'ensemble du processus permet d'identifier les configurations les plus performantes et les plus robustes vis-à-vis des effets du hasard.

Chaque modèle est évalué quantitativement selon plusieurs indicateurs : TP (vrais positifs), FP (faux positifs), FN (faux négatifs) et TN (vrais négatifs). (Figure 5). D'autres valeurs, plus synthétiques, sont ensuite calculées pour comparer les performances des modèles et pour la validation finale : la précision (TP/TP+FP), le rappel TP/TP+FN), le score F1 (moyenne harmonique de la précision et du rappel) et l'aire sous la courbe précision-rappel (AUC-PR) qui mesure la qualité du compromis précision/rappel sur tous les seuils possibles. Le modèle retenu in fine correspond à celui qui présente le meilleur score d'AUC-PR.

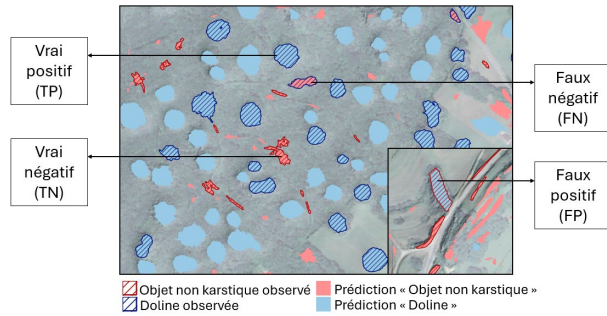


Figure 5 : Exemple de classification des prédictions d'un modèle XGBoost par rapport à la classification attribuée par l'expert sur l'échantillon d'entraînement.

#### 4. RESULTATS

Les variables explicatives retenues pour l'entraînement du modèle sur les 6 sites pilotes précédemment décrits sont : largeur, élongation, circularité et rectangularité.

Les indicateurs moyens obtenus sur la validation croisée du meilleur modèle sont : Précision  $\approx 0,88$ , Rappel  $\approx 0,85$ , Score F1  $\approx 0,865$  et AUC-PR = 0,93 (Figure 6). La variation des scores selon le seuil de probabilité montre un compromis favorable entre taux de faux positifs et faux négatifs.

Le modèle le plus performant sur l'ensemble des tests est ensuite appliqué à l'échelle du département du Doubs. Il attribue à chaque objet une probabilité d'être une doline. Pour décider si un objet est peu, moyennement ou très probablement une doline, les seuils suivants sont utilisés : en dessous d'une probabilité de 0,5 (standard statistique), l'objet est considéré comme ayant peu de chances d'être une doline (moins de 50%) ; au-dessus, il est plutôt considéré comme une doline. Un second seuil, 0,68 (68%), correspond au niveau où le modèle est statistiquement le plus performant (meilleur score F1).

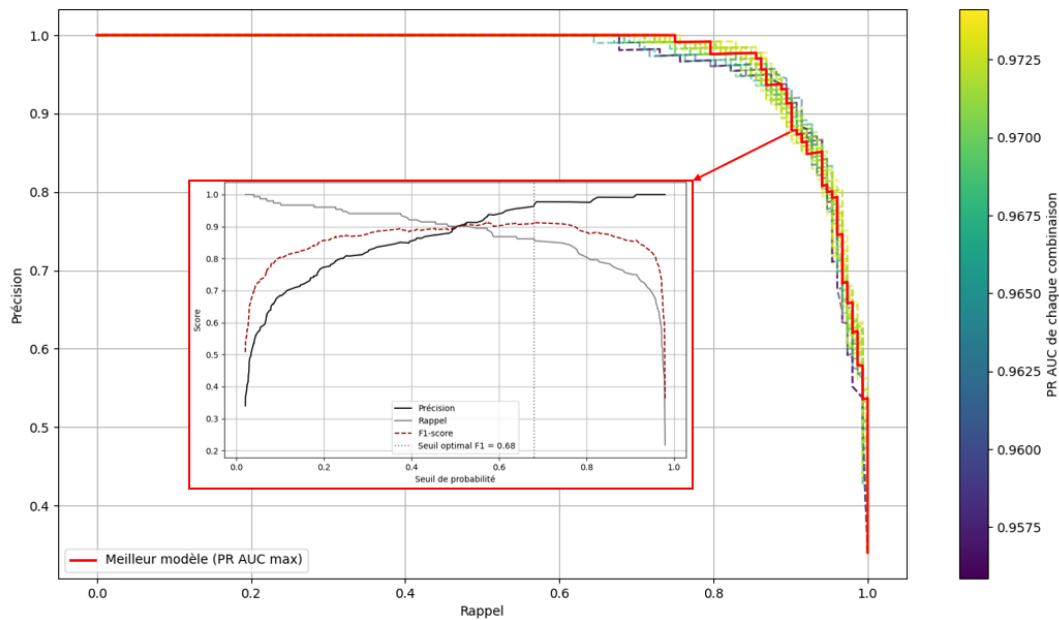


Figure 6 : Courbes précision-rappel pour les tests réalisés. Le modèle présentant l'AUC la plus importante correspondant au meilleur classificateur est indiqué en rouge. Le détail des variations de ses métriques en fonction du seuil de probabilité est présenté dans l'encadré en rouge.

Appliqué à l'ensemble du département, le modèle a permis d'analyser 1 511 779 dépressions détectées et d'attribuer à chacune une probabilité d'appartenance à la classe 'doline'. In fine, 1 219 054 dolines potentielles sont classées en « objet non karstique » (68,8%), 134 115 dolines potentielles sont classées en

« doline probable » (7,5%) et 158 610 sont classées en « doline très probable » (8,9%).

Un contrôle à dire d'expert sur le département est réalisé pour valider la pertinence du modèle retenu. Ce contrôle a été réalisé en comparant les résultats obtenus avec l'algorithme avec les données de l'Atlas Mouvements de Terrain (Cerema, 2013). Au total, 18

737 événements de l'Atlas sur les 24 416 coïncident avec les entités classées en karst/doline dans l'Atlas, soit un taux de reconnaissance satisfaisant de 77 %. Les secteurs déjà reconnus comme à forte densité de dolines (Valdahon, Montrond-le-Château, secteur nord de Besançon) sont bien retrouvés par la méthode. De nouvelles zones d'intérêt sont également mises en évidence, telles que les secteurs de La-Chapelle-des-Bois et Arcey.

## 5. DISCUSSION

Plusieurs pistes d'amélioration peuvent être envisagées afin d'accroître la performance, la robustesse et la pertinence du modèle de détection des indices karstiques.

- L'intégration des caractéristiques géologiques et géographiques comme variables explicatives permettrait d'affiner la détection en tenant compte des formations favorables au développement du karst et d'améliorer la contextualisation spatiale des résultats (distance aux routes, bâtiments). Utiliser ces variables pourrait toutefois conduire à exclure des dolines avérées situées en milieu urbanisé.

- Le présent travail s'est concentré sur la cartographie d'indices karstiques de type doline, caractérisés par une morphologie circulaire. Préciser la méthodologie suivant les formes karstiques (effondrement, soutirage, dissolution) et à d'autres formes karstiques (lapiaz, poljés, vallées sèches) compléterait la vision du fonctionnement karstique de la zone d'étude.

- La quantification des performances se limite actuellement au taux de reconnaissance du modèle. Y adjoindre une estimation de l'incertitude des géométries détectées (surface, profondeur ou contours réels) permettrait de mieux juger de la fiabilité du modèle et d'en améliorer l'interprétation. En partie liées à la qualité du LiDAR, les incertitudes pourraient être précisées via un indicateur de la couverture LiDAR. Les zones bénéficiant d'une résolution plus fine seraient ainsi distinguées de celles où la couverture est lacunaire, permettant une pondération adaptée des détections.

## 6. CONCLUSION

Cette étude démontre la pertinence d'une approche automatique pour la détection et la classification des

dolines à l'échelle départementale, avec des performances élevées. Le modèle permet de reconnaître 77 % des dolines recensées dans les bases de données existantes et conduit à l'identification de 292 725 dolines potentielles sur l'ensemble du territoire étudié. Ces résultats confirment l'intérêt des méthodes fondées sur l'analyse de MNT LiDAR et l'apprentissage automatique pour dépasser les limites des inventaires traditionnels, souvent incomplets et hétérogènes. L'inventaire obtenu constitue une base de données spatialisée et surfacique inédite, offrant un cadre de référence homogène pour l'analyse des formes karstiques de surface. Il permettra d'améliorer les analyses entre la distribution des dolines, la nature des formations géologiques et les caractéristiques morphologiques du relief. Ces données ouvrent également des perspectives pour une meilleure quantification des processus hydrologiques karstiques, notamment l'estimation des flux de percolation et des zones préférentielles d'infiltration. Enfin, cette approche fournit un socle méthodologique robuste pour le développement de produits opérationnels, en particulier la production de cartographies d'aléa liées aux mouvements de terrain et aux effondrements karstiques.

## REMERCIEMENTS

Cette étude, menée par le BRGM, s'inscrit dans le cadre d'un projet visant à caractériser l'aléa d'effondrement karstique, en partenariat avec la Direction Départementale des Territoires du Doubs et le Cerema, à qui les auteurs adressent leurs remerciements pour leur collaboration.

## REFERENCES

- Camus H. (2003). Hydrogéologie des systèmes karstiques du Jura. Presses Universitaires.
- Chen T., & Guestrin C. (2016). XGBoost: A scalable tree boosting system. In KDD 2016.
- J. Friedman. Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, 29(5):1189–1232, 2001.
- Cerema (2013). Atlas des mouvements de terrain du Doubs.
- Kweon I.S., & Kanade T. (1994). Contour tree approaches for terrain analysis. CVGIP.
- IGN (2021–2024). Programme LiDAR HD