

École Nationale Supérieure d'Informatique et Mathématiques  
Appliquées (ENSIMAG)

## Projet de Mathématiques Appliquées :

# Segmentation d'images satellitaires pour la détection des lacs en montagne

### Équipe :

Paul Bottner

Mohamed Khalil Becharai

Mostafa El Jardali

Obaida El Jouaid

Encadrant : Guillaume James

Année académique : 2025–2026

# Table des matières

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Données utilisées</b>	<b>2</b>
<b>3</b>	<b>Méthodes explorées</b>	<b>2</b>
3.1	Seuil fixe avec filtrage gaussien . . . . .	2
3.2	Chan–Vese . . . . .	3
3.3	Random Forest . . . . .	4
3.3.1	Principe général de l’algorithme . . . . .	4
3.3.2	Justification du choix . . . . .	4
3.3.3	Stratégie d’apprentissage . . . . .	4
<b>4</b>	<b>Métriques d’évaluation</b>	<b>5</b>
4.1	Distance de Hamming . . . . .	5
4.2	Différence d’aire . . . . .	5
4.3	Taux de fausse détection . . . . .	5
4.4	Taux de vraie détection . . . . .	6
4.5	Corrélation . . . . .	6
4.6	Similarité structurelle . . . . .	6
<b>5</b>	<b>Résultats</b>	<b>6</b>
5.1	Zones faciles . . . . .	6
5.2	Zones plus complexes . . . . .	8
<b>6</b>	<b>Pistes d’amélioration</b>	<b>10</b>
<b>7</b>	<b>Conclusion</b>	<b>10</b>
	<b>Bibliographie</b>	<b>11</b>

# 1 Introduction

Les lacs d’eau jouent un rôle essentiel dans l’équilibre écologique et hydrologique. Leur suivi est donc indispensable pour de nombreuses applications, notamment la gestion des ressources en eau et la prévention des risques naturels tels que les débordements dans les zones habitées. La délimitation automatique des contours des lacs à partir d’images satellitaires constitue ainsi un enjeu important.

Cependant, cette tâche reste complexe en raison des limites propres aux différents types d’images. Les images optiques, comme celles fournies par Sentinel-2, sont sensibles à la présence de nuages, d’ombres et aux variations d’illumination, ce qui complique l’identification fiable des surfaces d’eau. Les images radar SAR issues de Sentinel-1, quant à elles, peuvent présenter des distorsions liées à la géométrie radar, en particulier en terrain montagneux, rendant la segmentation des lacs difficile.

Dans le cadre de ce projet, plusieurs méthodes de segmentation d’images ont d’abord été explorées afin de mieux comprendre les caractéristiques des données et les difficultés liées à la segmentation des surfaces d’eau. Ces méthodes incluent le seuillage fixe associé à un filtrage gaussien ainsi que la méthode de Chan–Vese. Cette phase exploratoire a permis d’identifier les limites des approches purement basées sur l’intensité.

L’essentiel du travail se concentre ensuite sur une approche d’apprentissage supervisé basée sur la méthode des forêts aléatoires (Random Forest). Plusieurs modèles ont été entraînés, mais avec des ensembles de caractéristiques (features) différents. L’objectif est d’analyser l’influence du choix des caractéristiques, ainsi que de l’intégration du contexte spatial, sur la qualité de la segmentation des surfaces d’eau. Les performances des différents modèles sont évaluées à l’aide de métriques quantitatives.

## 2 Données utilisées

Les données utilisées pour la segmentation sont des moyennes mensuelles des images SAR du satellite Sentinel-1 (résolution de 20m), dont les caractéristiques suivent l’indice OASIS sur la période 2021–2024. Nous avons également utilisé une superposition entre les images optiques NDWI et les images SAR pour l’année 2024. Enfin, en l’absence de véritables images terrain, des images segmentées avec une méthode plus avancée, testée par les professeurs, nous ont été fournies afin de calculer les différentes métriques et d’entraîner les modèles.

Les expérimentations sont réalisées sur une zone d’étude située dans le secteur de la Bérarde, qui a été marquée en 2024 par la vidange d’un lac supraglacière du glacier de Bonne Pierre. Ce cas d’étude réel et complexe constitue un contexte pertinent pour évaluer la robustesse de la méthode proposée.

## 3 Méthodes explorées

Comme dans la majorité des méthodes de segmentation, une étape de filtrage est nécessaire afin de réduire le bruit présent dans les images.

Nous avons choisi d’appliquer un filtre gaussien simple, dont le principe est de remplacer la valeur d’un pixel par une moyenne pondérée des pixels voisins, les poids dépendant de la distance par rapport au pixel central. Le paramètre  $\sigma$  du filtre gaussien a été fixé à 2. Ce choix permet de réduire efficacement le bruit tout en préservant les contours des lacs, lesquels représentent généralement une portion relativement réduite de l’image. Un  $\sigma$  trop élevé risquerait en effet de lisser excessivement les contours et de dégrader la précision de la segmentation. Le filtrage gaussien a été implémenté à l’aide de la fonction dédiée du module `scipy.ndimage`.

### 3.1 Seuil fixe avec filtrage gaussien

La méthode du seuil fixe est l’une des méthodes de segmentation les plus simples. Elle consiste à définir un seuil à partir duquel un pixel est classé soit comme appartenant à un lac, soit comme

non-lac. Dans la littérature, une zone est généralement considérée comme de l’eau lorsque les valeurs de rétrodiffusion SAR sont comprises entre  $-20$  et  $-10$  dB. Toutefois, dans notre cas, les images utilisées suivent l’indice OASIS. Par conséquent, le seuil a été fixé empiriquement à la valeur 0.2, après plusieurs tests exploratoires visant à obtenir une séparation satisfaisante entre les zones d’eau et les autres surfaces.

Après des tests réalisés sur plusieurs zones d’étude, la méthode du seuil fixe a montré des résultats satisfaisants dans les zones dites faciles (zones 5 et 7). En revanche, cette approche échoue largement dans les zones plus complexes, notamment celles contenant des glaciers. Cette dégradation des performances s’explique par le fait que le seuillage conduit à une sur-segmentation des zones de neige humide, qui sont alors confondues avec les surfaces lacustres. De plus, cette méthode ne prend pas en compte le contexte spatial, chaque pixel étant traité indépendamment de ses voisins, ce qui la rend particulièrement sensible au bruit et aux variations locales d’intensité.

### 3.2 Chan–Vese

La méthode de *Chan–Vese* est une approche de segmentation par évolution de contour visant à séparer une image en deux régions aussi homogènes que possible. Elle repose sur la minimisation d’une fonctionnelle d’énergie qui dépend à la fois de la régularité du contour et de l’homogénéité des régions qu’il délimite. Cette fonctionnelle s’écrit :

$$F = \mu \cdot \text{Length}(C) + \nu \cdot \text{Area}(\text{inside}(C)) + \lambda_1 \int_{\text{inside}(C)} |u_0(x, y) - c_1|^2 dx dy \\ + \lambda_2 \int_{\text{outside}(C)} |u_0(x, y) - c_2|^2 dx dy.$$

où  $C$  désigne le contour recherché,  $u_0(x, y)$  l’intensité du pixel de coordonnées  $(x, y)$ , et  $c_1$  et  $c_2$  les moyennes des intensités respectivement à l’intérieur et à l’extérieur du contour. Les paramètres  $\lambda_1, \lambda_2, \mu > 0$  contrôlent l’importance relative des différents termes de l’énergie, tandis que  $\nu \in \mathbb{R}$  agit comme un terme de pondération de l’aire. La minimisation de cette fonctionnelle est réalisée de manière itérative, généralement à l’aide d’une descente de gradient dans le cadre des méthodes de *level-set* Guiot et al. (2023).

Le paramètre  $\mu$  a été choisi égal à 0,07. Ce choix s’explique par la nature des lacs, qui occupent généralement une faible portion de l’image. Une valeur faible de  $\mu$  permet de limiter la régularisation du contour et de préserver les détails fins des frontières des lacs. Les paramètres  $\lambda_1$  et  $\lambda_2$  ont été fixés à 1, afin de donner un poids équivalent aux régions situées à l’intérieur et à l’extérieur du contour, sans privilégier a priori l’une des deux classes.

Malgré ces ajustements, la méthode de Chan–Vese n’a pas permis d’obtenir des résultats satisfaisants. Elle tend à sur-segmenter les lacs et présente un phénomène d’inversion entre les régions « lac » et « non-lac ». Cette limitation s’explique principalement par le fait que la méthode ne repose que sur des critères d’homogénéité d’intensité et ne prend pas en compte la sémantique de la segmentation.

Plusieurs tentatives d’amélioration ont été explorées. Le problème d’inversion entre les régions a été corrigé en s’appuyant sur la moyenne d’intensité des régions segmentées, l’eau présentant généralement une intensité plus élevée. Cette stratégie a permis d’identifier correctement la région correspondant au lac après la segmentation. Par ailleurs, la fonction d’initialisation fournie par la bibliothèque `scikit-image`, initialement basée sur un damier, a été remplacée par une initialisation issue d’un seuillage fixe avec un seuil de 0,2. Malgré ces ajustements, les performances globales de la méthode sont restées limitées et n’ont pas permis d’obtenir une segmentation satisfaisante.

### 3.3 Random Forest

#### 3.3.1 Principe général de l'algorithme

Random Forest est un algorithme d'apprentissage automatique supervisé. Il existe deux variantes : l'une utilisée pour la classification et l'autre pour la régression. Dans notre cas d'application, qui est la segmentation d'images, nous nous intéressons uniquement à la classification. L'algorithme repose sur la combinaison de plusieurs arbres de décision, d'où son nom.

On peut donner l'intuition de l'algorithme de la manière suivante : pour chaque arbre de décision, on sélectionne aléatoirement  $k$  caractéristiques parmi les  $m$  disponibles. À partir de ces caractéristiques, chaque nœud de l'arbre est construit en choisissant le meilleur critère de division, puis est subdivisé en nœuds fils de manière récursive jusqu'à atteindre la profondeur maximale fixée.

Pour effectuer une prédiction sur de nouvelles données, chaque arbre de la forêt fournit sa propre prédiction. La classe finale est déterminée par un vote majoritaire parmi les prédictions de tous les arbres, ce qui rend le modèle plus robuste et précis que l'utilisation d'un seul arbre de décision. Banerjee (2020)

#### 3.3.2 Justification du choix

Notre choix d'appliquer l'algorithme des forêts aléatoires (*Random Forest*) est motivé par plusieurs raisons. Tout d'abord, cette méthode est reconnue pour sa robustesse, liée à la combinaison d'un grand nombre d'arbres de décision, ce qui permet de réduire significativement les phénomènes de surapprentissage (*overfitting*).

De plus, l'algorithme offre la possibilité d'estimer l'importance de chaque caractéristique, facilitant ainsi l'analyse et la sélection des caractéristiques les plus pertinentes. Cet aspect est particulièrement intéressant dans notre cas, où différentes informations peuvent être intégrées, notamment celles liées à la texture ou à la période d'acquisition des images.

Cette capacité d'enrichissement par de nouvelles caractéristiques constitue un avantage majeur par rapport à d'autres méthodes explorées, telles que le seuillage fixe, qui repose uniquement sur l'intensité des pixels, ou la méthode de Chan–Vese, principalement basée sur l'évolution des contours.

Enfin, le choix de cette approche est également appuyé par les performances rapportées dans l'article Guiot et al. (2023), où les forêts aléatoires obtiennent des résultats encourageants pour une problématique similaire en télédétection.

#### 3.3.3 Stratégie d'apprentissage

Afin d'améliorer la qualité de la segmentation et de la rendre plus flexible, aussi bien pour le suivi des lacs dans des zones déjà connues (sur lesquelles le modèle a été entraîné) que pour l'application à de nouvelles zones, plusieurs modèles basés sur la méthode des forêts aléatoires ont été entraînés. Dans chaque cas, le modèle repose sur la même méthode d'apprentissage, mais utilise des ensembles de caractéristiques différents.

Les paramètres des arbres ont été conservés identiques pour l'ensemble des modèles, avec un nombre de 20 arbres et une profondeur maximale fixée à 10. Ce choix représente un compromis entre les performances obtenues et le temps d'entraînement. Les modèles ont été entraînés à partir d'images de moyenne mensuelle couvrant les années 2023 et 2024, tandis que les phases de test ont été réalisées sur les années 2021 et 2022. L'implémentation a été effectuée à l'aide de la bibliothèque `scikit-learn`.

Le premier modèle, appelé **RF\_SAR**, repose uniquement sur les images SAR. Les caractéristiques utilisées sont l'intensité du pixel issue de l'image normalisée, un seuillage fixe à 0,2, ainsi qu'une mesure de texture obtenue à partir de la variance locale calculée sur une fenêtre de  $20 \times 20$  pixels. Ce modèle a fourni des résultats satisfaisants, en particulier dans les zones dites faciles. Toutefois, ses performances restent proches de celles obtenues par la méthode de seuillage fixe.

Le deuxième modèle **RF\_Z** intègre, en plus des caractéristiques issues des images SAR, une image colocalisée correspondant à une superposition entre une image optique et une image SAR de la même

région. Ce modèle peut être appliqué à des zones sur lesquelles il n’a pas été entraîné, dès lors que l’image colocalisée est disponible, ce qui lui confère une meilleure capacité de généralisation.

Enfin, le dernier modèle, noté **RF\_ZM**, enrichit le second modèle en intégrant des informations contextuelles supplémentaires, à savoir le mois d’acquisition et le numéro de la zone. De plus, pour ce modèle, l’image colocalisée est sélectionnée de manière à minimiser l’écart temporel entre les acquisitions, en retenant systématiquement le mois le plus proche. Cette meilleure proximité temporelle contribue à améliorer la cohérence des données d’entrée et explique en partie les performances supérieures observées. Toutefois, ce modèle est principalement destiné au suivi des lacs dans des zones sur lesquelles il a été entraîné.

Les résultats détaillés obtenus par ces trois modèles sont présentés et analysés dans le chapitre consacré aux résultats.

## 4 Métriques d’évaluation

Afin d’évaluer les performances des méthodes de segmentation des lacs, les résultats obtenus ont été comparés aux segmentations de référence (vérités terrain) à l’aide de plusieurs métriques quantitatives. Ces indicateurs permettent de mesurer les divergences entre images binaires et également de prendre en compte la distribution des pixels. Les métriques utilisées sont détaillées ci-dessous.

Considérons la comparaison entre une image binaire issue de la segmentation automatique et une image de vérité terrain. On définit :

- $n$  : le nombre total de pixels considérés ;
- $a$  : le nombre de pixels correctement classés comme lac (vrais positifs) ;
- $b$  : le nombre de pixels classés comme lac alors qu’ils ne le sont pas (faux positifs) ;
- $c$  : le nombre de pixels appartenant au lac mais non détectés (faux négatifs).

### 4.1 Distance de Hamming

La distance de Hamming est une mesure mathématique correspondant à la proportion de pixels non concordants entre deux images binaires. Elle est comprise entre 0 et 1, où 0 indique une correspondance parfaite entre les deux images.

Elle est définie par :

$$h = \frac{b + c}{n}.$$

### 4.2 Différence d’aire

La différence d’aire quantifie l’écart entre les surfaces détectées dans les deux images binaires. Elle est calculée comme la différence absolue entre le nombre de pixels à 1 dans l’image segmentée et dans l’image de référence, normalisée par le nombre maximal de pixels à 1 observé entre les deux images. Cette métrique est comprise entre 0 et 1.

### 4.3 Taux de fausse détection

Le taux de fausse détection (*False Alarm Rate*, FAR) mesure la proportion de pixels incorrectement détectés comme lac parmi l’ensemble des pixels détectés comme lac. Il est défini par :

$$FAR = \frac{b}{a + b}.$$

Cette métrique est comprise entre 0 et 1, une valeur faible indiquant peu de fausses détections.

#### 4.4 Taux de vraie détection

Le taux de vraie détection (*Hit Rate*, HR) mesure la proportion de pixels du lac correctement détectés par rapport au nombre total de pixels du lac observés dans la vérité terrain. Il est défini par :

$$HR = \frac{a}{a + c}.$$

Cette mesure est également comprise entre 0 et 1, une valeur élevée traduisant une bonne capacité de détection.

#### 4.5 Corrélation

La corrélation permet d'évaluer le lien statistique entre deux images. Elle a été calculée à l'aide du module Python NumPy, qui applique la relation suivante :

$$R_{ij} = \frac{C_{ij}}{\sqrt{C_{ii}C_{jj}}}.$$

où  $C_{ij}$  désigne l'élément d'indice  $(i, j)$  de la matrice de covariance  $C$ .

Le coefficient de corrélation est compris entre  $-1$  et  $1$ , où  $1$  indique une corrélation positive parfaite,  $-1$  une corrélation négative parfaite, et  $0$  l'absence de corrélation.

#### 4.6 Similarité structurelle

La similarité structurelle (*Structural Similarity Index Measure*, SSIM) permet d'évaluer la similarité entre deux images en tenant compte de la luminance, du contraste et de la structure. Dans ce travail, la SSIM a été calculée à l'aide de l'implémentation fournie par la bibliothèque `scikit-image`.

### 5 Résultats

#### 5.1 Zones faciles

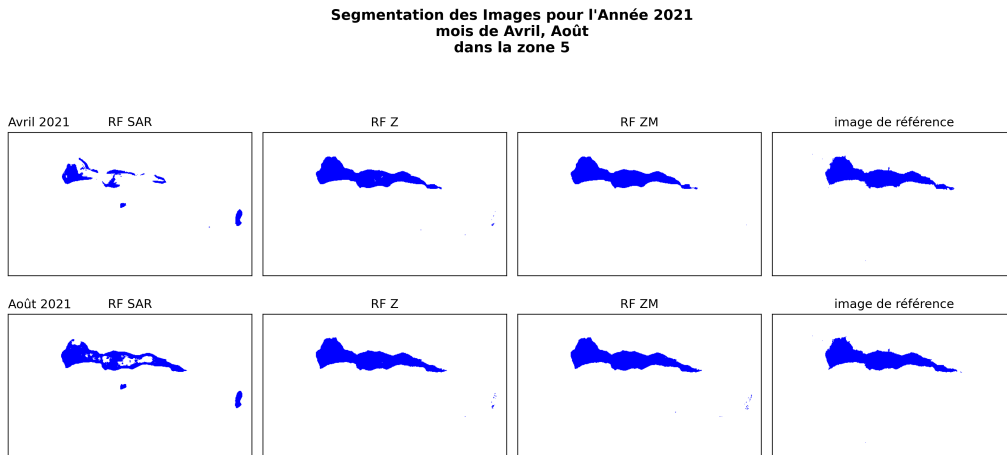


FIGURE 1 – Exemple de segmentation d'images SAR avec image colocalisée (Zone 5, lac du Chambon — zone facile), réalisée à l'aide des trois modèles de Random Forest, pour une période estivale (août 2021) et une période non estivale (avril 2021), avec comparaison aux images de référence.

Dans la Figure 1, on observe clairement la différence de performance entre le modèle **RF\_SAR** et les modèles **RF\_Z** et **RF\_ZM**, qui prennent en considération à la fois les images SAR et les images

colocalisées. Ces deux modèles produisent une segmentation quasi parfaite, aussi bien pour les périodes estivales que non estivales. On peut toutefois noter un léger sur-lissage des contours, dû à l'utilisation du filtrage gaussien. Néanmoins, la suppression de ce filtre entraîne une dégradation des performances globales ; il a donc été conservé dans la chaîne de traitement.

À l'inverse, le modèle **RF\_SAR** présente davantage de difficultés : certaines zones du lac ne sont pas correctement segmentées et les performances se dégradent nettement durant les périodes non estivales. Cette baisse de performance est principalement liée à la présence de neige pendant cette période, contrairement à la période estivale où la segmentation s'avère plus fiable.

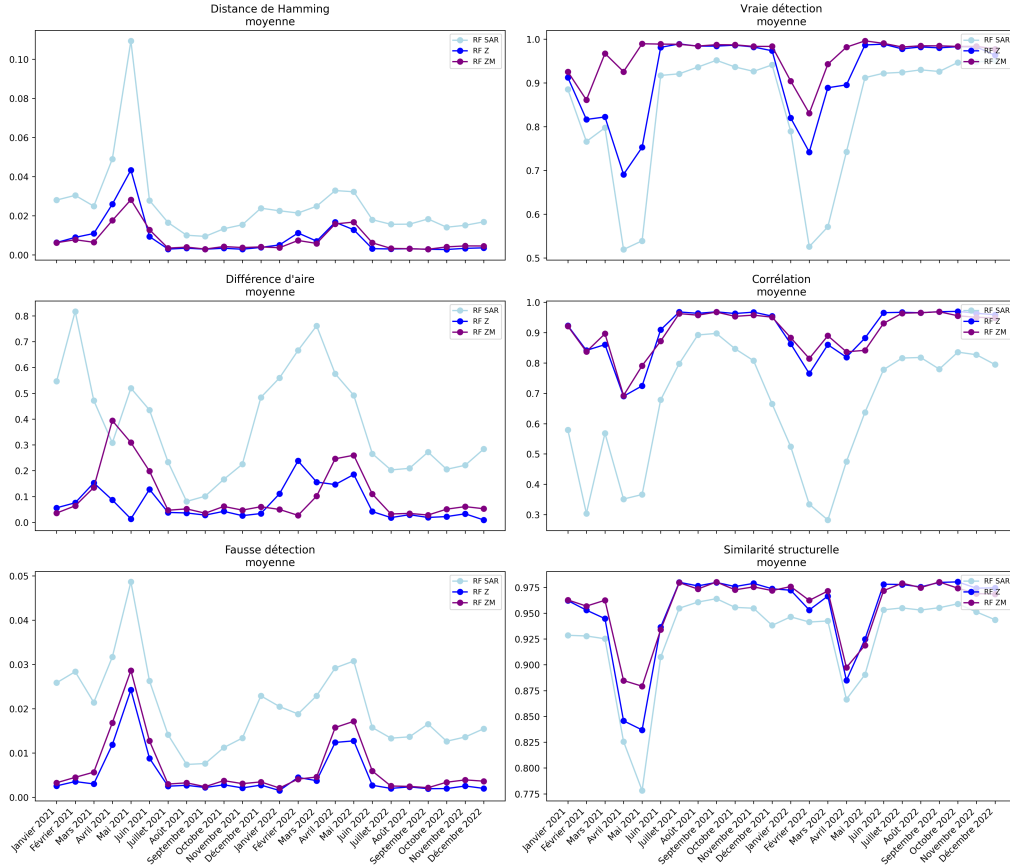


FIGURE 2 – moyenne des scores obtenus sur les deux zones (Zone 5 et Zone 7) pour les métriques de corrélation, distance de Hamming, vraie détection, fausse détection, différence d'aires et similarité structurelle (SSIM), sur les années 2021 et 2022, pour les trois modèles.

Comme illustré dans la Figure 2, on remarque que les trois modèles obtiennent des scores relativement élevés pour la vraie détection, la corrélation et la similarité structurelle. Cependant, un creux significatif est observé entre les mois de mars et juin pour l'ensemble des modèles.

Par ailleurs, les métriques de différence d'aire, de fausse détection et de distance de Hamming présentent des valeurs proches de zéro sur la majorité des périodes, avec néanmoins des pics correspondant aux mêmes intervalles temporels où les autres scores diminuent.

Cette dégradation des performances peut s'expliquer par le début de la fonte de la neige durant cette période. La présence de neige humide complique alors la distinction entre les surfaces d'eau et leur environnement, ce qui rend la segmentation plus délicate pour l'ensemble des modèles.

On observe également que les modèles **RF\_Z** et **RF\_ZM** affichent des performances très proches, tandis que le modèle **RF\_SAR** s'en écarte nettement, confirmant ses performances inférieures par rapport aux deux autres approches.



## 5.2 Zones plus complexes

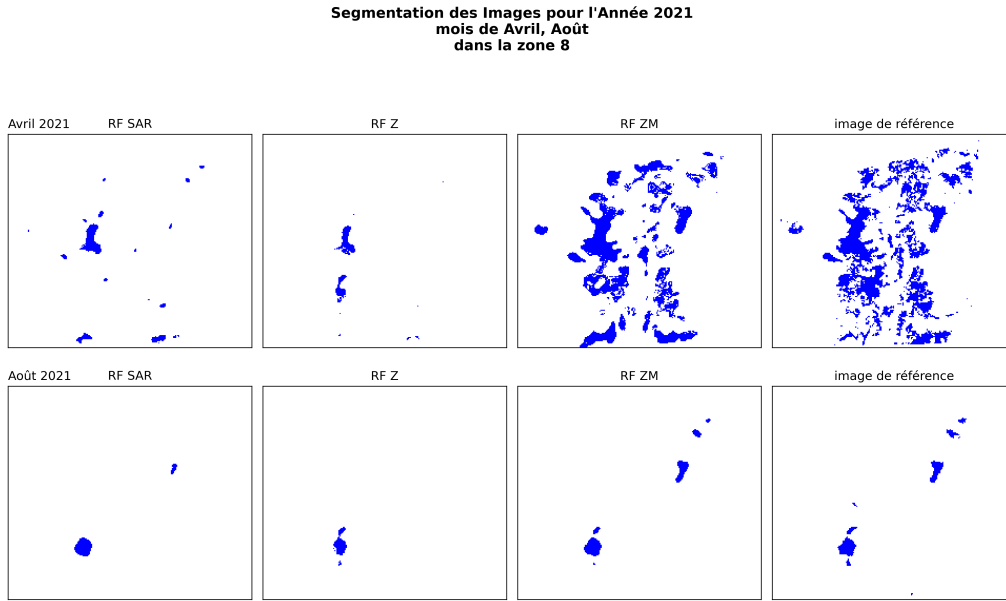


FIGURE 3 – Exemple de segmentation d’images SAR avec image colocalisée (Zone 8 — zone difficile, présence de neige humide), réalisée à l’aide des trois modèles de Random Forest, pour une période estivale (août 2021) et une période non estivale (avril 2021), avec comparaison aux images de référence.

L’analyse de la Figure 3 met en évidence, pour le mois d’avril, une différence marquée entre les segmentations produites par les deux premiers modèles et celle obtenue avec le modèle **RF\_ZM**, cette dernière apparaissant nettement plus pertinente. En revanche, les modèles **RF\_SAR** et **RF\_Z** fournissent des segmentations plus fragmentées, un comportement également observé pour le mois d’août.

Par ailleurs, pour ces deux mois, le modèle **RF\_SAR** estime une surface segmentée correcte légèrement plus étendue que celle obtenue avec le modèle **RF\_Z**, contrairement à ce qui avait été observé pour les zones dites faciles. Cette différence peut être attribuée à une image colocalisée plus éloignée temporellement pour cette zone, ce qui affecte la qualité des informations fournies au modèle **RF\_Z**.

Enfin, un lissage excessif des contours est observé pour l’ensemble des modèles, de manière similaire à ce qui avait déjà été constaté sur la Figure 1.

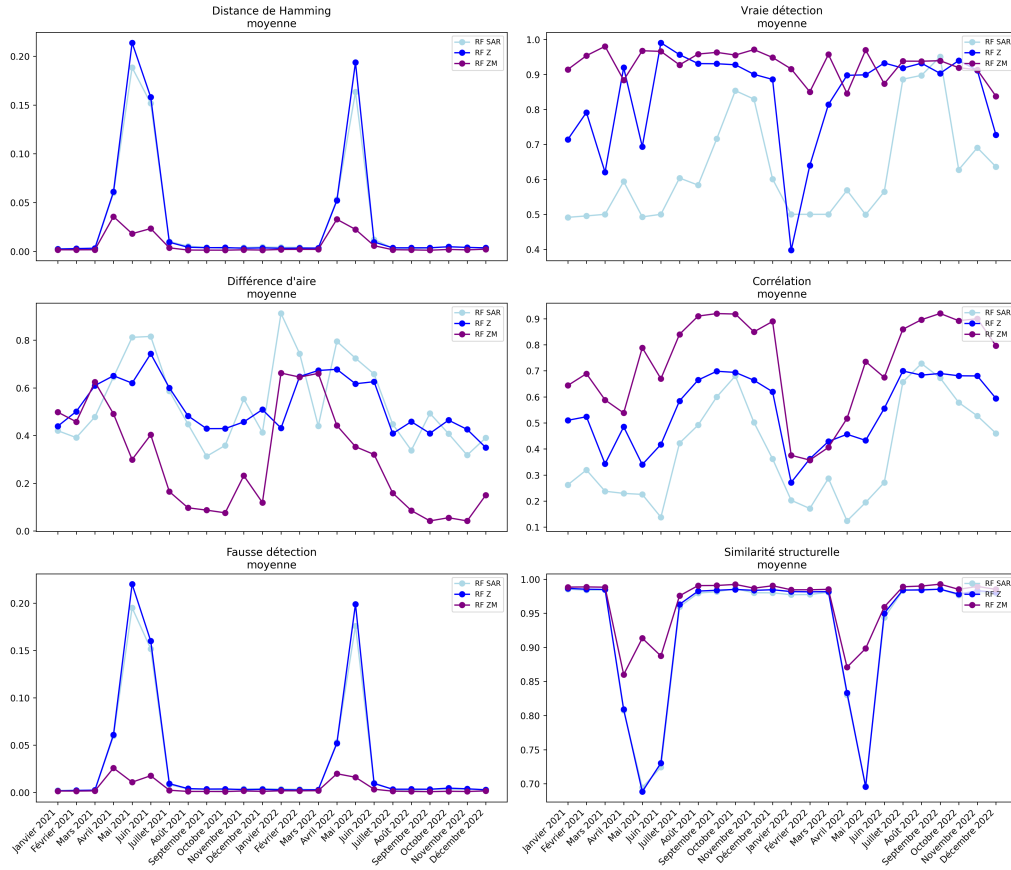


FIGURE 4 – moyenne des scores obtenus sur les deux zones (Zone 3 et Zone 8) pour les métriques de corrélation, distance de Hamming, vraie détection, fausse détection, différence d’aires et similarité structurelle (SSIM), sur les années 2021 et 2022, pour les trois modèles.

Dans la Figure 4, nous observons des pics et des creux similaires à ceux de la Figure 2, mais avec une amplitude beaucoup plus marquée, notamment pour les modèles **RF\_SAR** et **RF\_Z**. Cela pourrait être dû à une évolution rapide des surfaces d’eau dans ces régions. En ce qui concerne les courbes représentant la différence d’aire, le taux de vraies détections et la corrélation moyenne, on remarque des fluctuations importantes des scores, avec des variations considérables à la fois au fil du temps et entre les modèles.

TABLE 1 – Moyennes des scores de segmentation des moyennes mensuelles de 2021-2022 sur les zones 1 à 8

Scores	RF SAR	RF Z	RF ZM
Distance de Hamming moyenne	0.02	0.014	0.007
Différence d’aire moyenne	0.47	0.411	0.347
Fausse détection moyenne	0.016	0.013	0.006
Vraie détection moyenne	0.65	0.851	0.94
Corrélation moyenne	0.469	0.631	0.68
Similarité structurelle moyenne	0.94	0.951	0.963

Dans la Table 1, nous présentons une comparaison des trois modèles **RF\_SAR**, **RF\_Z** et **RF\_ZM**, basée sur les moyennes des différents indicateurs de performance introduits précédemment. Ces scores ont été calculés sur la période 2021–2022 et pour l’ensemble des huit zones d’étude. Les résultats montrent que le modèle **RF\_ZM** surpasse globalement le modèle **RF\_Z**, ce qui peut s’expliquer par l’utilisation d’un nombre plus important de caractéristiques ainsi que par une colocalisation des

images présentant un écart temporel réduit. Par ailleurs, une amélioration notable des performances est observée entre les modèles **RF\_SAR** et **RF\_Z**, suggérant que l'intégration d'images colocalisées apporte une information supplémentaire significative sur la surface étudiée.

Les valeurs moyennes obtenues pour la distance de Hamming, la similarité structurelle et le taux de fausses détections apparaissent globalement satisfaisantes. En revanche, les indicateurs liés à la différence d'aire, au taux de vraies détections et aux corrélations présentent des résultats plus contrastés. Ces derniers suggèrent que, bien que les modèles localisent correctement les zones contenant des lacs, celles-ci ne sont pas toujours entièrement détectées.

## 6 Pistes d'amélioration

Une première piste d'amélioration consisterait à étendre la classification binaire actuelle en introduisant des classes supplémentaires, telles que *neige* et *glacier*, au lieu de se limiter à la distinction lac / non-lac. Une telle approche permettrait de mieux discriminer les surfaces d'eau des zones de neige humide, qui présentent des signatures radiométriques proches dans les images SAR. La réduction de cette ambiguïté devrait conduire à une segmentation plus précise des lacs, en particulier dans les zones complexes caractérisées par la présence de neige et de glaciers.

Par ailleurs, l'utilisation de méthodes de filtrage plus adaptées aux images SAR, telles que le filtre de Lee, pourrait contribuer à une meilleure atténuation du bruit de speckle tout en préservant les structures pertinentes pour la segmentation.

Comme pour tout modèle d'apprentissage, disposer de davantage de données permettrait également d'augmenter la robustesse et la précision. On aurait souhaité avoir plus d'images colocalisées pour chaque mois. Enfin, l'intégration d'informations supplémentaires, telles que l'altitude ou la pente des zones étudiées, aurait probablement permis d'améliorer encore les performances.

## 7 Conclusion

La segmentation des lacs de montagne à partir des données des satellites Sentinel-1 et Sentinel-2 constitue un outil pertinent pour l'analyse de leur évolution temporelle. Dans ce travail, plusieurs méthodes de segmentation ont été étudiées à partir de données acquises dans des zones montagneuses des Alpes françaises sur la période 2021–2024, en s'appuyant sur des segmentations de référence fournies à l'aide d'une méthode plus élaborée. Parmi les approches évaluées, la méthode de segmentation Random Forest s'est révélée globalement efficace et présente une bonne capacité d'adaptation aux variations du terrain. Toutefois, les modèles proposés montrent encore certaines limites, notamment pour assurer une détection complète et fiable des lacs dans l'ensemble des conditions climatiques observées, ce qui ouvre plusieurs perspectives d'amélioration. Par ailleurs, les résultats obtenus avec la méthode de Chan–Vese suggèrent qu'elle dispose d'un potentiel important, sous réserve d'optimisations supplémentaires. Enfin, ce travail met en évidence que la combinaison d'images SAR et optiques, à travers l'utilisation de données colocalisées, a un impact significatif sur les performances de segmentation, comme l'indiquent nos méthodes de mesure.

## Références

Banerjee, P. (2020), ‘Random forest classifier tutorial’. Tutoriel en ligne.

**URL:** *<https://www.kaggle.com/code/prashant111/random-forest-classifier-tutorial>*

Guiot, A., Karbou, F., James, G. & Durand, P. (2023), ‘Insights into segmentation methods applied to remote sensing sar images for wet snow detection’, *Geosciences* **13**(7).