



**HAL**  
open science

## Le Pôle Thématique National des Surfaces Continentales Theia : produits et services pour l'agriculture.

Nicolas Baghdadi, Jordi Inglada, Marie Weiss, Hassan Bazzi, Valérie Demarez,  
Philippe Lagacherie, Isabelle Biagiotti, Arnaud Selle

### ► To cite this version:

Nicolas Baghdadi, Jordi Inglada, Marie Weiss, Hassan Bazzi, Valérie Demarez, et al.. Le Pôle Thématique National des Surfaces Continentales Theia : produits et services pour l'agriculture.. Innovations Agronomiques, 2021, 83, pp.11-27. 10.15454/0xy8-yc88 . hal-03542910

**HAL Id: hal-03542910**

**<https://hal.inrae.fr/hal-03542910>**

Submitted on 25 Jan 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0  
International License

## **Le Pôle Thématique National des Surfaces Continentales Theia: produits et services pour l'agriculture**

**Baghdadi N.<sup>1</sup>, Inglada J.<sup>2,5</sup>, Weiss M.<sup>3</sup>, Bazzi H.<sup>1</sup>, Demarez V.<sup>2</sup>, Lagacherie P.<sup>4</sup>,  
Biagiotti I.<sup>1</sup>, Selle A.<sup>5</sup>**

<sup>1</sup> INRAE, UMR TETIS, 500, rue François Breton, F-34093 Montpellier Cedex 5

<sup>2</sup> CESBIO, Université de Toulouse, CNES/CNRS/INRAE/IRD/UPS, Toulouse, France

<sup>3</sup> INRAE, UMR EMMAH, 228, route de l'aérodrome, CS 40509, F-84914 Avignon Cedex

<sup>4</sup> INRAE, UMR LISAH, 2, place Viala, F-34060 Montpellier Cedex 2

<sup>5</sup> CNES, Centre de Toulouse, France

**Correspondance** : nicolas.baghdadi@inrae.fr

### **Résumé**

Le Pôle thématique national des surfaces continentales Theia a pour objectif d'accroître l'utilisation par la communauté scientifique et les acteurs publics de la donnée spatiale en complémentarité d'autres types de données, notamment les données *in situ* et aéroportées. Depuis quelques années, Theia, met à disposition des acteurs publics nationaux, des scientifiques (nationaux et internationaux) et des acteurs privés, des données et produits à valeur ajoutée issus de la télédétection par satellite, complémentaires à l'offre européenne Copernicus. Theia vise également à structurer la communauté scientifique nationale, à mutualiser les données image, les traitements et l'expertise scientifique, ainsi qu'à rendre visibles les réalisations nationales à l'échelle internationale. Theia constitue un écosystème d'innovation au service de la recherche, de l'action publique et du développement économique dans les domaines de l'environnement, des hydro- et agro-systèmes et de l'aménagement des territoires, notamment par la mise à disposition de données, aussi bien en France qu'en Europe et dans les pays du Sud. Il est une des composantes de l'Infrastructure de Recherche (IR) Data Terra comme les trois autres pôles de données – Aeris, Odatis, et ForM@Ter.

**Mots-clés** : Pôle de données, Theia, Surfaces continentales, Télédétection, Agriculture

### **Abstract: Theia, the French Thematic Consortium for Land Surfaces: Products and Services for Agriculture**

The French Land Data and Services consortium Theia aims to increase the use by the scientific community and public actors of spatial data in complementarity with other types of data, in particular *in situ* and airborne data. For several years, Theia has been providing national public actors, scientists (national and international) and private actors with data and value-added products from satellite remote sensing that complement the European Copernicus offer. Theia also aims to structure the national scientific community, to pool image data, processing and scientific expertise, and to make national achievements visible on an international scale. Theia constitutes an ecosystem of innovation at the service of research, public action and economic development in the fields of the environment, hydro and agro-systems and regional planning, in particular by providing data in France as well as in Europe and in countries in the South. Theia is a component of the Data Terra Research Infrastructure (IR) like the other three data consortia – Aeris (Atmosphere), Odatis (Ocean) and ForM@Ter (Solid Earth).

**Keywords**: Data centre, Theia, Land, Remote-sensing, Agriculture

## 1. Contexte et objectifs de Theia

Le pôle de données et de services Theia ([www.theia-land.fr](http://www.theia-land.fr)) est un consortium de 10 institutions publiques françaises impliquées dans l'observation de la Terre et les sciences de l'environnement (CEA, CEREMA, CIRAD, CNES, IGN, INRAE, CNRS, IRD, Météo France, AgroParisTech et ONERA). Il a été créé en 2012 dans le but d'augmenter l'utilisation des données spatiales par la communauté scientifique et les acteurs publics. Les premières années ont permis de structurer les communautés scientifiques et d'utilisateurs nationales, de mutualiser les ressources, de faciliter l'accès aux données et aux capacités de traitement, de fédérer diverses initiatives jusque-là indépendantes, et de diffuser les acquis français à l'échelle nationale et internationale. Des actions de diffusion et de formation ciblant des utilisateurs dans d'autres pays ont depuis été développées. Le pôle Theia fait partie de l'Infrastructure de Recherche Data Terra avec ODATIS (Données et Service pour l'Océan), ForM@Ter (Données et Service pour la Terre Solide) et AERIS (Données et Service pour l'Atmosphère) (Huynh et al., 2019). Theia structure la communauté scientifique à travers : 1) les centres d'expertise scientifique (CES), organisés autour de la conception et du développement de produits à valeur ajoutée de pointe ; 2) des réseaux d'animation régionaux (ART) fédérant les utilisateurs (scientifiques et acteurs publics / privés) ; 3) la mise en place d'une infrastructure de services et de données (IDS) mutualisée, répartie entre plusieurs centres et permettant l'accès à une variété de produits (Figures 1 et 2, Tableau 1). Actuellement, l'infrastructure Theia s'appuie sur le centre de données CNES (Toulouse), le centre de données Geosud (Montpellier, sélectionné dans le cadre de l'appel à projets « Équipements d'Excellence » du Programme national d'Investissements d'Avenir en 2011) avec une connexion au centre de données Geoportail (Paris). Chaque centre de données est indépendant et responsable du prétraitement, du stockage, de l'archivage et de la distribution des données aux utilisateurs. Récemment l'IDS A2S de l'Université de Strasbourg a intégré les IDS de Theia.

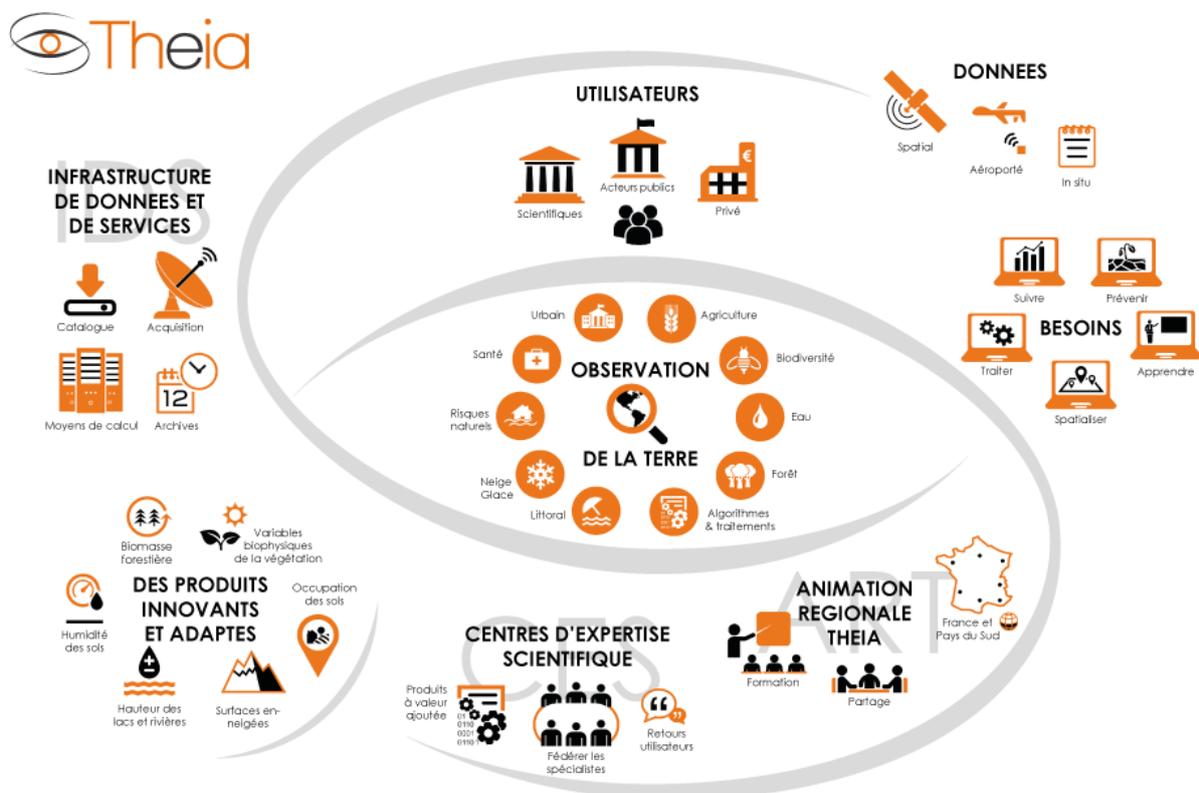
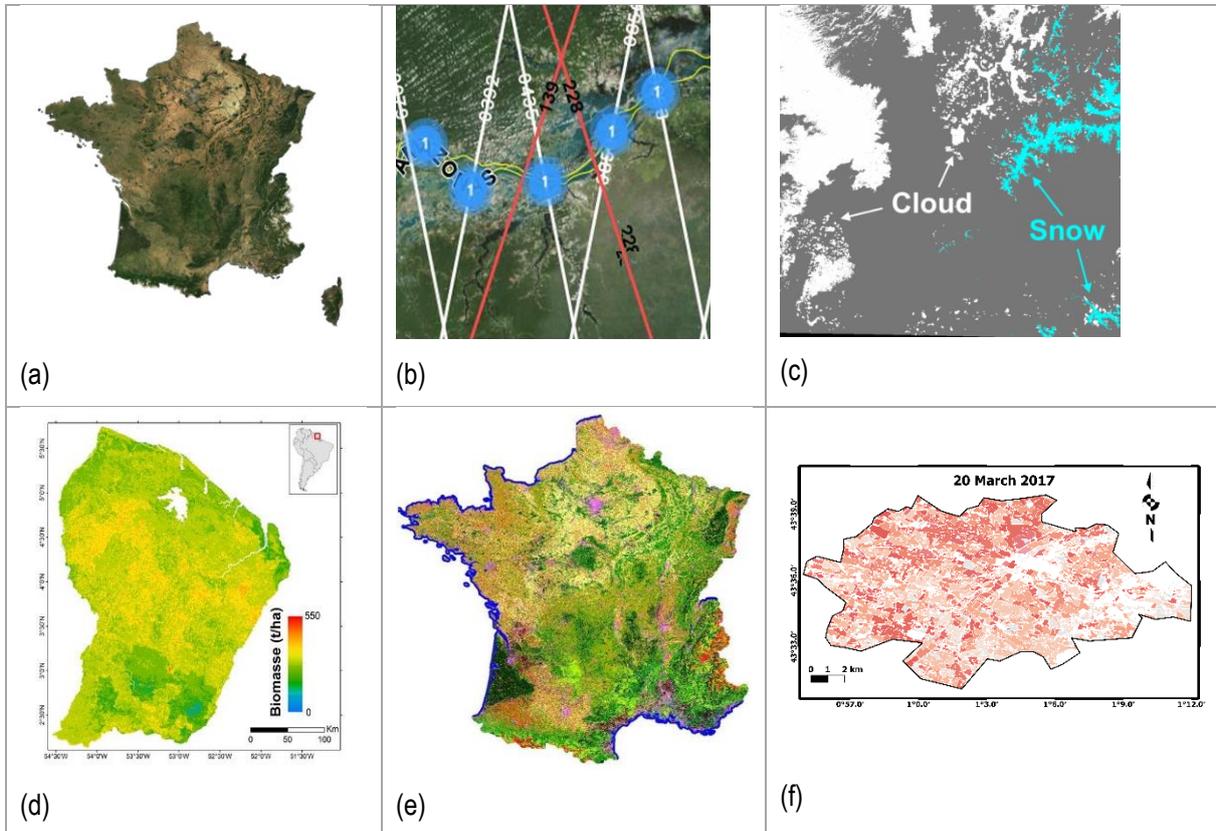


Figure 1 : Structure du pôle Theia.

La communauté scientifique et les acteurs publics sont le principal public cible de l'action, mais le secteur privé peut également bénéficier des synergies créées par le pôle Theia. En effet, la plupart des données sont diffusées sous licence ouverte et les algorithmes en Open Source. Le volet formation, à consolider, contribuera à renforcer à plus long terme la capacité de tous ces utilisateurs.



**Figure 2** : Exemple de produits thématiques Theia. (a) Réflectance de surface Sentinel-2, (b) Hauteur des lacs et des rivières, (c) Surfaces enneigées, (d) Biomasse forestière, (e) Occupation du sol, (f) Humidité du sol.

## 2. Organisation du pôle Theia

Theia propose une animation scientifique à l'échelle régionale et nationale (sections 2.1 & 2.2) et fournit un accès aux produits satellitaires prétraités et à valeur ajoutée (déjà accessibles ou en développement-section 2.3).

### 2.1 Réseau de centres d'expertise scientifique (CES)

Les centres d'expertise scientifique (CES) de Theia sont des groupes de laboratoires de recherche qui développent des technologies innovantes pour utiliser les données spatiales pour la surveillance de la surface de la Terre. Ils se concentrent sur le développement de produits à valeur ajoutée, éventuellement avec des services associés à ces produits. Les CES sont mono ou multi-laboratoires. Leurs objectifs sont de participer au développement et à la validation des produits, de proposer de nouvelles applications et de fédérer la communauté scientifique aux niveaux national, autour de domaines thématiques (agriculture, foresterie, eau, sol, urbain, côtier, neige / glace ; Figure 1, Tableau 1).

Les CES ont été initialement créés pour regrouper des chercheurs autour du développement de chacun des produits à valeur ajoutée. On distingue les CES ayant des produits déjà opérationnels (réflectance de surface, occupation des sols, variables biophysiques de la végétation, humidité du sol, surfaces enneigées, hauteurs des lacs et des rivières) de ceux qui sont en développement ou en validation (évapotranspiration, surfaces irriguées, cartographie numérique des sols, biomasse forestière et

changements du couvert forestier, couleurs des eaux continentales, zones urbaines et artificielles, risques associés aux maladies infectieuses). Les produits sont d'abord développés sur le territoire national pour démonstration. Ces produits sont accessibles pour la communauté scientifique et l'ensemble des utilisateurs via un nouveau catalogue offrant d'importantes fonctionnalités de recherche : [catalogue.theia-land.fr](http://catalogue.theia-land.fr)

**Tableau 1** : Extrait du portefeuille de produits Theia: <https://www.theia-land.fr/produits-thematiques/>

Classe	Produits	Zone	Période	Accès	Disponibilité
<b>Produits à valeur ajoutée</b>	Hydroweb	Hauteur des grands lacs et rivières du monde	1992 > présent	Tout utilisateur	<a href="http://catalogue.theia-land.fr">catalogue.theia-land.fr</a>
	Occupation des sols (OSO)	France	2016, 2017, 2018, 2019	Tout utilisateur	<a href="http://catalogue.theia-land.fr">catalogue.theia-land.fr</a>
	Surface enneigée	Pyrénées, Haut-Atlas, Alpes françaises	Juillet 2016 > présent	Tout utilisateur	<a href="http://catalogue.theia-land.fr">catalogue.theia-land.fr</a>
	Humidité des sols 25 km	Global	2002 > présent	Tout utilisateur	<a href="http://ftp.ifremer.fr">ftp.ifremer.fr</a>
	Humidité des sols THRS	Régionale (17 sites)	Septembre 2016 > août 2020	Tout utilisateur	<a href="http://catalogue.theia-land.fr">catalogue.theia-land.fr</a>
	Altitude de ligne d'équilibre glaciaire annuelle	240 glaciers des Alpes européennes	2016 > présent	Tout utilisateur	<a href="http://theia-land.fr">theia-land.fr</a>
	Biomasse, hauteur de la canopée	Guyane française, Madagascar, Gabon, A. subsaharienne	2013 2010 2010 2010	Tout utilisateur	<a href="http://theia-land.fr">theia-land.fr</a>
	Cultures irriguées	Sud-Ouest Adour, Tarn Catalogne	2015 2017 2018	Tout utilisateur	<a href="http://peps-vizo.cnes.fr">peps-vizo.cnes.fr</a>
<b>Réflectance de surface</b>	Sentinel-2 niveaux 2 & 3	Europe occidentale et autres régions du monde	2016 > présent	Tout utilisateur	<a href="http://catalogue.theia-land.fr">catalogue.theia-land.fr</a>
	Landsat	France et Rom-Com	2005 > 2011   2013 > 2017 2018 > présent	Tout utilisateur	<a href="http://catalogue.theia-land.fr">catalogue.theia-land.fr</a>
	Venµs niveaux 2 & 3	110 sites	2017 > présent	Tout utilisateur	<a href="http://catalogue.theia-land.fr">catalogue.theia-land.fr</a>

## 2.2 Réseaux régionaux d'animation Theia (ART)

Les objectifs des ART (Animation Régionale Theia) sont de coordonner les utilisateurs (scientifiques et acteurs publics) à l'échelle régionale et de participer aux efforts de formation, notamment en ce qui concerne les produits à valeur ajoutée développés par les CES. Huit ART sont répartis sur le territoire national français (dont un en Nouvelle-Calédonie). Un ART « Pays du Sud » (GeoDEV) assure le développement de la communauté des utilisateurs dans les pays en développement. Une relation solide entre les CES et les ART est en cours de développement pour diffuser les résultats des CES aux communautés d'utilisateurs.

### 2.3 Portefeuille de produits

Les produits fournis par Theia (Tableau 1) sont soumis à un contrôle qualité, couvrent de vastes territoires et de longues périodes. Il s'agit de séries chronologiques de réflectance de surface à très hautes résolutions spatiales et temporelles, des cartes d'occupation du sol, et de séries chronologiques de variables bio-géophysiques (humidité du sol, niveaux d'eau, biomasse, etc.).

## 3. Les CES Theia au service de l'agriculture

Quatre CES Theia offrent aujourd'hui des produits et des algorithmes exploitables pour l'agriculture.

### 3.1 CES Occupation du sol (OSO)

Le Centre d'Expertise Scientifique (CES) Occupation des sols (OSO) regroupe des équipes du Cesbio, Dynafor, UMR Tetis, IGN-Matis qui ont comme objectif la définition et de développement d'algorithmes automatiques pour la production de cartes d'occupation des sols à partir d'imagerie satellitaire. Le CES est en production depuis 2017. Les caractéristiques principales du produit OSO France sont les suivantes: (1) le produit couvre la totalité de la France métropolitaine avec une nomenclature à 23 classes (17 classes pour les 2 premiers millésimes 2016 et 2017) ; (2) il est livré sous 2 formats différents : une version raster avec des pixels de 10 m x 10 m et une version vecteur avec une unité minimale de collecte de 0.1 ha ; et (3) le produit est généré tous les ans et le millésime correspond à une année civile. Le millésime de l'année N est livré avant la fin du premier trimestre de l'année N+1.

Il est important de noter que ces caractéristiques sont des spécifications du produit qui ont été établies après consultation auprès d'utilisateurs potentiels. Cette consultation a eu lieu entre 2014 et 2016 et des produits prototypes ont été fournis aux utilisateurs pour évaluer le degré d'adéquation des caractéristiques du produit aux besoins. Le CES étant passé à la phase de production, les éventuelles évolutions des spécifications (nomenclature, extension géographique, fréquence de production, résolution spatiale) doivent se faire de façon à ce que le produit reste compatible avec les spécifications précédentes. Par exemple, l'extension de 17 à 23 classes réalisée à partir de 2018 s'est faite par le découpage de certaines classes de la nomenclature initiale, permettant ainsi aux utilisateurs d'avoir une correspondance non ambiguë avec les millésimes précédents.

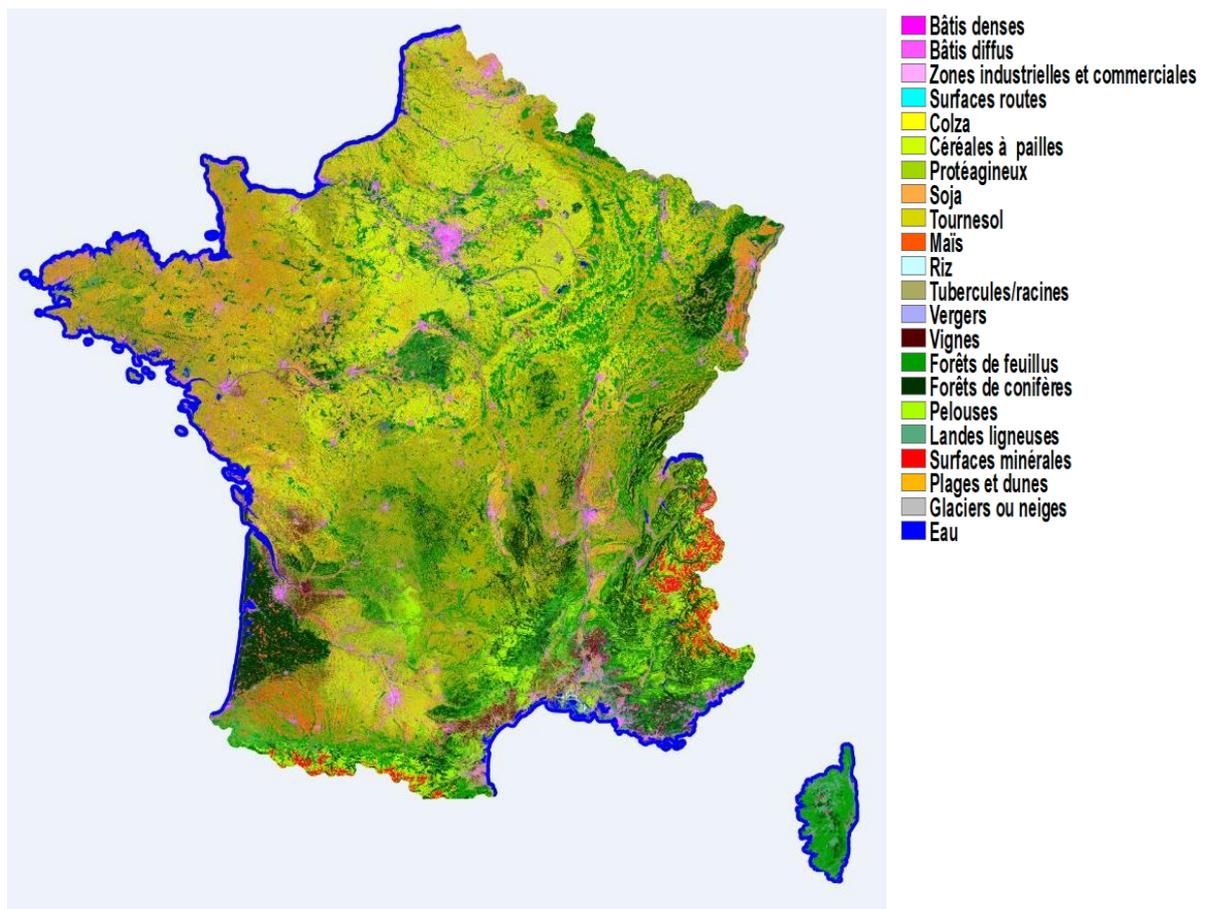
La carte OSO France est produite par classification automatique d'images satellite. La procédure est détaillée dans Inglada et al. (2017). À l'heure actuelle, la procédure utilise exclusivement des séries temporelles d'imagerie optique à haute résolution spatiale issues du système Sentinel-2. Des études ont été réalisées pour évaluer l'apport des séries temporelles de capteurs radar (Sentinel-1) et de l'imagerie à très haute résolution spatiale (SPOT-6/7 et Pléiades-HR), mais les résultats n'ont pas été suffisamment probants pour que ces approches soient adoptées pour la production opérationnelle. Les laboratoires partenaires du CES continuent le travail sur le développement de méthodes pour l'amélioration de la qualité du produit OSO.

Les algorithmes de classification utilisés sont de type supervisé, ce qui nécessite des données de référence pour réaliser la calibration (apprentissage) des modèles. Les données utilisées pour l'apprentissage sont issues de bases de données publiques (Corine Land Cover, Urban Atlas, BD Topo, etc.), mais aussi de partenariats spécifiques. Ainsi, par exemple, depuis 2018, le Cesbio a une convention avec l'Agence de services et de paiement (ASP) qui lui donne accès à des données du RPG (Registre parcellaire graphique) avant consolidation, ce qui permet de classer les cultures annuelles et donc être capables d'enrichir les postes agricoles de la nomenclature.

À ce jour, le produit OSO n'a pas d'équivalent aux mêmes échelles spatiales ni temporelles (fréquence de mise à jour). Par exemple, Corine Land Cover a une fréquence de mise à jour beaucoup plus faible et

un détail spatial moins fin, et d'autres produits Copernicus ont des nomenclatures simplifiées (High Resolution Layers) ou des étendues géographiques limitées (Urban Atlas).

La qualité thématique du produit est évaluée à la fin de chaque production. D'un côté, une validation par comparaison à un échantillon de la donnée de référence est réalisée. Cet échantillon contient des zones différentes de celles utilisées pour l'apprentissage et donc permet de mesurer la capacité de l'algorithme à généraliser à des régions géographiques différentes. On peut ainsi générer une matrice de confusion avec, en lignes, les classes attribuées par l'algorithme et, en colonnes, les classes dans la donnée de référence, pour un grand nombre de pixels. La carte OSO millésime 2019 est montrée en Figure 3 avec la matrice de confusion en Figure 4. La limite de cette validation réside dans le fait que les données de référence (Corine Land Cover, BD Topo, etc.) peuvent être obsolètes. Un deuxième type de validation est réalisée en utilisant des données correspondant à la même période que celle du millésime. Ces données sont difficiles à obtenir et ne sont pas disponibles en quantité suffisante pour les utiliser pour l'apprentissage. Aussi, ces données sont souvent concentrées sur des petites portions du territoire et ont des nomenclatures différentes de celles du produit OSO. Elles peuvent cependant être utilisées pour réaliser une validation indépendante du produit. Plus d'informations sur cette approche de validation peuvent être consultées ici : <https://labo.obs-mip.fr/multitemp/premieres-validations-de-la-carte-doccupation-du-sol-oso/>.



**Figure 3 :** Carte OSO Millésime 2019 avec ses 23 classes.

KAPPA : 0.888 OA : 0.898

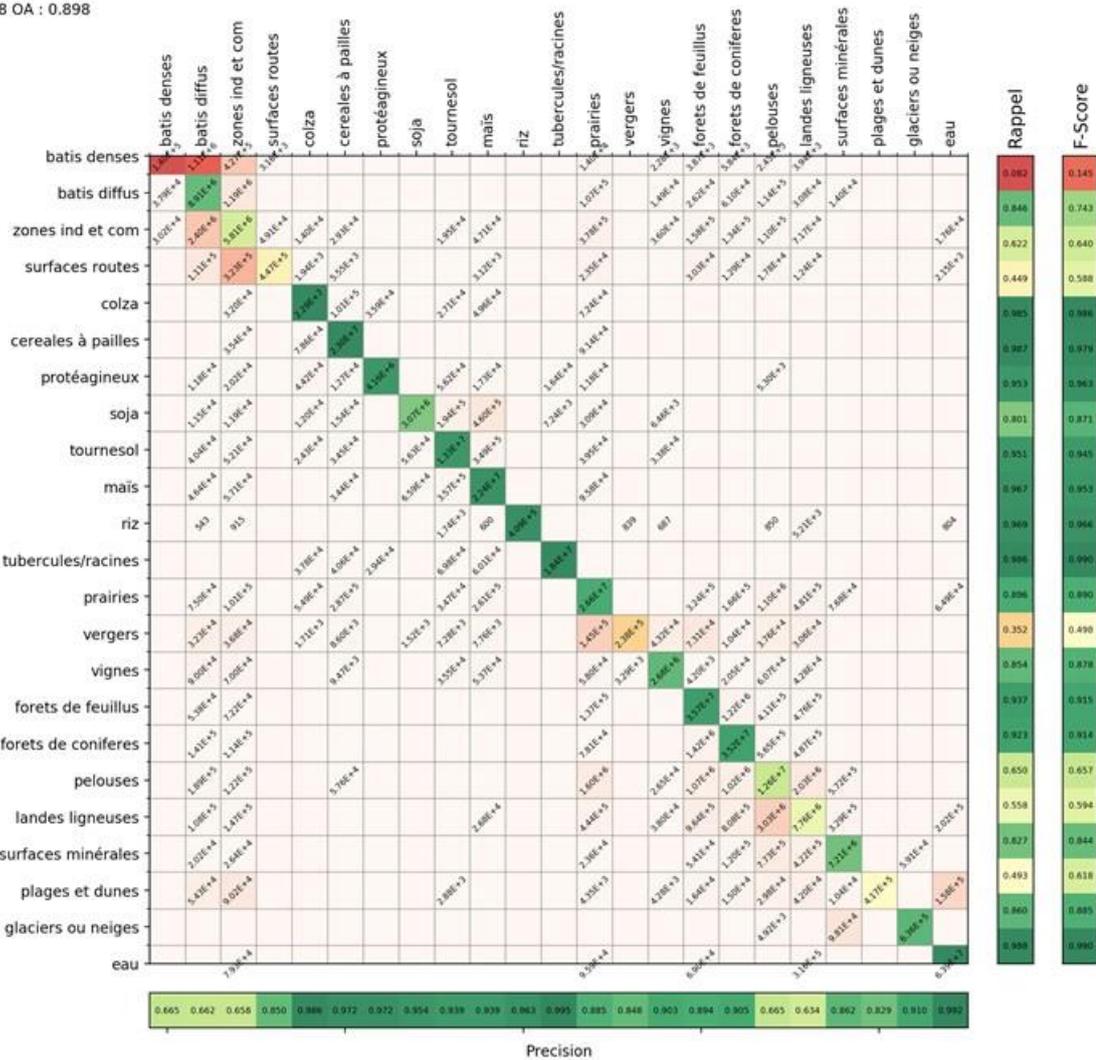


Figure 4 : Matrice de confusion pour le millésime OSO 2019.

Enfin, une validation par des photo-interprètes a aussi été réalisée pour le millésime 2016. Le même protocole utilisé pour les produits Copernicus a été appliqué, et le produit OSO a été considéré comme étant de qualité opérationnelle (plus de détails ici : <https://labo.obs-mip.fr/multitemp/another-validation-of-cesbios-2016-france-land-cover-map/>).

Ces 2 validations indépendantes sont cohérentes avec les matrices de confusion fournies avec les produits.

Pour compléter les informations de fiabilité du produit, 2 cartes supplémentaires sont distribuées avec chaque millésime. La première, appelée carte de « validité », donne, pour chaque pixel de la carte OSO, le nombre d'acquisitions non nuageuses utilisées pour la détermination de la classe thématique. Un pixel avec peu d'observations claires aura une information moins fiable. La deuxième carte est appelée « confiance » et donne la probabilité du pixel d'appartenir à la classe renseignée dans le produit OSO. En effet, la méthode de classification choisit la classe la plus probable mais, dans certains cas, l'écart de probabilité entre la première et la deuxième classe peut être faible. L'intérêt de ces 2 cartes supplémentaires est de donner une information locale qui complète l'information globale fournie par la matrice de confusion.

Dans le cas du produit vecteur, chaque polygone contient les validités et confiances moyennes des pixels le composant.

Pour les applications d'agriculture de précision ou de statistique agricole en cours de saison, le produit OSO annuel avec 11 classes agricoles n'est pas très utile. Dans le cas de la statistique agricole, d'autres sources de données comme le RPG ou Agreste sont préférables à OSO. En revanche, les applications qui s'intéressent aux cultures hors RPG avec une spatialisation plus fine que la Statistique Agricole Annuelle peuvent faire appel à OSO. D'autres usages constatés du produit OSO sont en rapport avec la connaissance de l'environnement proche des parcelles agricoles pour des besoins de modélisation ou de suivi de la déprise agricole, etc.

### 3.2 CES Humidité du sol à THRS (très haute résolution spatiale)

Le suivi spatio-temporel de l'humidité du sol dans les zones agricoles est d'une grande importance pour de nombreuses applications, notamment celles liées au cycle continental de l'eau. L'utilisation de capteurs *in situ* assure cette surveillance mais cette technique est très coûteuse et elle ne peut être réalisée que sur une très petite zone agricole, d'où l'importance de la télédétection spatiale qui permet désormais une cartographie opérationnelle à grande échelle de l'humidité du sol à forte résolution spatio-temporelle.

Les images radar sont utilisées depuis longtemps pour estimer et cartographier l'humidité de surface des sols nus. Pour les surfaces avec une couverture végétale, le couplage de données radar et optique est souvent nécessaire pour estimer l'humidité du sol. Les données optiques sont complémentaires des données radar, et leur intérêt réside dans leur potentiel à estimer les paramètres biophysiques de la végétation, par exemple le *Leaf Area Index* (LAI) et le *Normalized Difference Vegetation Index* (NDVI). Ces paramètres permettent d'évaluer la contribution de la végétation dans le signal radar rétrodiffusé, et par conséquent d'extraire la contribution du sol pour ensuite l'inverser et estimer l'humidité du sol.

Pour cartographier l'humidité du sol dans le cas d'un sol couvert de végétation, la plupart des études utilisent le modèle semi-empirique Water Cloud développé par Attema et Ulaby en 1978. Généralement, dans ce modèle, le signal radar total rétrodiffusé est modélisé comme la somme du signal rétrodiffusé du sol multiplié par l'atténuation bidirectionnelle, et du signal direct réfléchi par la végétation. Dans la plupart des études, la contribution de la végétation est exprimée par un paramètre biophysique de la végétation (biomasse, LAI, teneur en eau de la végétation, ou NDVI). La contribution du sol est généralement modélisée en fonction de l'humidité et de la rugosité du sol pour des paramètres instrumentaux donnés: angle d'incidence, longueur d'onde et polarisation. Il peut être simulé à l'aide d'un modèle physique de rétrodiffusion radar (*Integral Equation Model "IEM"* développé par Fung en 1994), ou d'un modèle de rétrodiffusion semi-empirique (par exemple le modèle de Baghdadi, Baghdadi et al., 2016).

Récemment, l'arrivée du satellite radar à synthèse d'ouverture (SAR) Sentinel-1 (S1) a fourni aux utilisateurs des données SAR en libre accès à une résolution spatiale élevée (10 m x 10 m) et avec un temps de revisite élevé (six jours en Europe). La mission S1 de l'Agence spatiale européenne (ESA) est une constellation de deux satellites SAR en orbite polaire (Sentinel-1A et Sentinel-1B) fonctionnant dans la bande C (~ 5,4 GHz). Les données SAR de la mission S1 à des résolutions spatiales et temporelles élevées ont encouragé la cartographie de l'humidité du sol en mode opérationnel.

El Hajj et al. (2017) ont développé une méthode opérationnelle pour cartographier l'humidité de surface du sol (SSM) à l'échelle de la parcelle sur les zones agricoles basée sur le couplage des données S1-SAR et des données optiques Sentinel-2 (S2) en utilisant la technique du réseau neuronal (S<sup>2</sup>MP). Les cartes S<sup>2</sup>MP sont produites pour les cultures été-hiver dans les zones agricoles et les prairies (elles ne sont pas appliquées aux vignobles et aux vergers). Le pôle de données Theia (<https://www.theia-land.fr/en/>) utilise l'algorithme développé par El Hajj et al. (2017) pour la fourniture de cartes d'humidité du sol à l'échelle de la parcelle pour plusieurs sites dans le monde (France, Italie, Espagne, Maroc, Liban, etc.). Les cartes S<sup>2</sup>MP pourraient être dérivées pour les images S1 acquises dans les modes d'acquisition

ascendant et descendant (heure d'acquisition ~ 18: 00 et 06:00 TU). Près de 20 images S1 couvrent totalement ou partiellement la plupart des bassins versants en France par mois. Ainsi, une estimation de l'humidité du sol pourra être disponible au maximum tous les 3 ou 4 jours.

Le Centre d'Expertise Scientifique (CES) humidité du sol à très haute résolutions spatiale (THRS) regroupe des équipes de l'UMR Tetis et du Cesbio qui ont comme objectif le développement d'algorithmes d'estimation de l'humidité à l'échelle parcellaire, sub-parcellaire mais aussi à moyenne résolution spatiale sur des mailles de 500 m x 500 m ou 1 km x 1 km. Le CES est en production depuis 2018 (Figure 5).

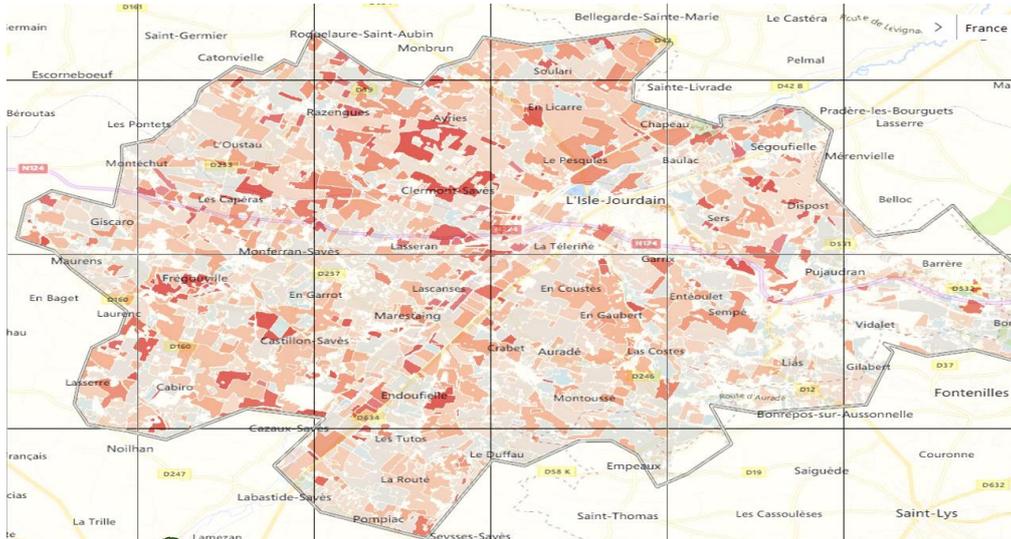


Figure 5 : Un aperçu du produit d'humidité du sol à l'échelle de la parcelle (S2MP).

### 3.3 CES Cartographie numérique des sols

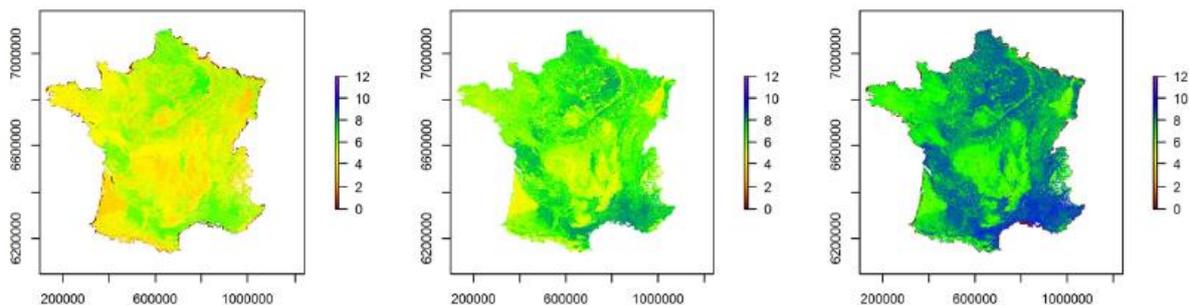
Les sols sont connus pour rendre un certain nombre de services écosystémiques indispensables à la vie sur terre : production d'aliments, de fibre et de combustibles, purification des eaux, régulation des crues, siège de biodiversité, régulation du climat par le stockage du carbone, etc. L'aptitude des sols à assurer ces services écosystémiques est gouvernée par un certain nombre de propriétés-clé du sol comme la profondeur, la texture ou le pH. Ces propriétés sont éminemment variables dans l'espace sous l'action de facteurs naturels de la pédogénèse (climat, relief, roche mère, végétation) et de l'activité humaine. Il est donc important de disposer de données spatiales précises sur les propriétés des sols afin d'orienter les décisions, tant aux échelles régionales que locales.

Traditionnellement, l'information spatialisée sur les sols est fournie par les cartes pédologiques. La couverture et la résolution spatiales de ces cartes est cependant trop limitée pour satisfaire le besoin de connaissance locale sur les sols. La cartographie numérique des sols (CNS) constitue une évolution prometteuse de la cartographie pédologique traditionnelle (Lagacherie et al., 2013). Le principe général est de prédire des types de sol ou des propriétés de sol par des algorithmes d'apprentissage automatique (ex : Random Forest) utilisant les données spatiales numériques représentant des éléments du paysage en relation (de causalité ou non) avec les sols (ex MNT) ou considérées comme proxys de certaines propriétés de sol (images de télédétections VIS-NIR). Ces algorithmes sont calibrés à partir de données pédologiques « *in situ* » disponibles sur la zone à étudier. La CNS est désormais mise en œuvre à l'échelle globale selon des spécifications définies dans le projet GlobalSoilMap (Arrouays et al., 2014) : images à 90 m de résolution de propriétés de sol clé, estimées à différentes profondeurs avec un degré estimé d'incertitude.

Le CES Cartographie Numérique des Sols s'inscrit dans cette perspective. Son objectif général est d'accompagner le passage à l'opérationnalité de la cartographie numérique des sols à l'échelle de la France. Ceci se décline selon les objectifs spécifiques suivants :

- Produire sur le territoire national des premières images d'estimation de propriétés de sol selon les spécifications GlobalSoilMap.
- Fédérer les efforts des différentes équipes françaises compétentes en télédétection des sols et CNS pour relever les défis méthodologiques dont la résolution nous apparaît comme un préalable pour améliorer ces premières estimations à l'échelle nationale.
- Transférer et diffuser les compétences en matière de cartographie numérique et de télédétection des sols vers des acteurs intervenant aux échelons régional ou local et dans les pays du Sud.

Les premières images de propriétés des sols à 90 m de résolution couvrant le territoire français ont été produites pour les propriétés suivantes (GlobalSoilMap France, Mulder et al, 2016): profondeur du sol, pH, taux d'argile, de limon, de sable, de carbone organique, d'éléments grossiers et capacité d'échange cationique, et ce, sur les 6 intervalles de profondeurs spécifiés par GlobalSoilMap (voir exemple Figure 6). Ces images sont le résultat de prédiction d'un outil d'apprentissage automatique (Cubist) utilisant des données spatiales disponibles à l'échelle nationale et calibré sur une moyenne de 30 000 observations de sol (de 3 000 à 60 000 selon les propriétés et les profondeurs). Ces produits seront mis à disposition début 2021 par le Groupement d'Intérêt Scientifique Sol (GISSOL, <https://data.inra.fr/dataverse/gissol>). Des produits similaires ont été également réalisés pour les deux régions Languedoc-Roussillon (diffusé par OpenIG <https://www.openig.org>) et Bretagne (non encore diffusé à ce jour).



**Figure 6** : Un exemple de produit GlobalSoilMap France. Images de prédiction du pH du sol (5-15 cm) ; au centre valeur modale, à gauche et à droite bornes inférieures et supérieures de l'intervalle de confiance à 90% (Mulder et al, 2016)

Au travers de l'ensemble des images d'estimation de propriétés de sol, GlobalSoilMap France offre un aperçu inédit des variations de sol au sein du territoire national. Cependant, comme le montrent les fortes variations de prédiction entre les deux bornes extrêmes de l'intervalle de confiance (Figure 6), ces premières estimations ont encore des précisions trop modestes pour être effectivement mobilisées pour des décisions opérationnelles à des échelles locales. Des travaux sont menés par les équipes du CES Theia Cartographie numérique des sols pour améliorer cette situation. Trois voies sont principalement explorées :

- Intégration des produits de télédétection, notamment les images Sentinel et les données Gamma-radiométriques, dans le pool de données spatiales utilisées pour prédire les propriétés des sols ;
- Densification des échantillonnages d'observations de sol sur lesquels se calent les outils d'apprentissage automatique par a) numérisation massive de données pédologiques anciennes ou par b) des approches participatives collectant de nouveaux sites d'observation ;

- Estimations non biaisées des incertitudes de prédiction résiduelles permettant d'éclairer les décisions supportées par les produits de CNS.

Une réflexion est également en cours pour aborder la cartographie numérique des fonctions du sol et d'indice de qualité de sol, ceci afin de répondre à une demande croissante des utilisateurs dans ce domaine. Des essais ont été notamment entrepris dans ce sens dans les régions Occitanie et Bretagne.

### 3.4 CES Paramètres biophysiques de la végétation

Le CES Paramètres biophysiques de la végétation est dédié à l'estimation de diverses variables biophysiques présentant la double caractéristique de jouer un rôle-clé dans les processus liés à la croissance de la plante et d'être accessibles par télédétection optique (Weiss et al., 2020). Ce sont des variables indispensables pour le suivi des cultures, en particulier pour l'évaluation du rendement, pour le suivi des pratiques agricoles ou l'analyse de la réponse d'une culture à des stress environnementaux comme des déficiences en azote ou la sécheresse. Elles constituent en effet des variables importantes du bilan hydrique, servant directement d'entrée pour les travaux de nombreux autres CES, comme par exemple le CES Irrigation. En agriculture de précision, utilisées conjointement avec des modèles de cultures, ces variables contribuent à mieux prédire les rendements comme à optimiser la fertilisation azotée et l'irrigation pour prendre en compte les hétérogénéités intra-parcellaires au service d'une agriculture plus durable et viable économiquement. À des échelles régionales, voire globales, les applications principales concernent la prévention des risques (famines, prévision des rendements, baisse drastique de revenus et assurance pour les agriculteurs suite à un événement climatique...) ainsi que le suivi du changement climatique. La revisite temporelle et la résolution spatiale de Sentinel-2 sont inédites et bien adaptées pour les applications agricoles ou forestières, en particulier dans les zones où de fortes hétérogénéités sont observées. Les pays dont le paysage agricole est particulièrement morcelé comme en Europe ou en Afrique ou les zones d'agriculture biologique constituent de bons exemples d'utilisation.

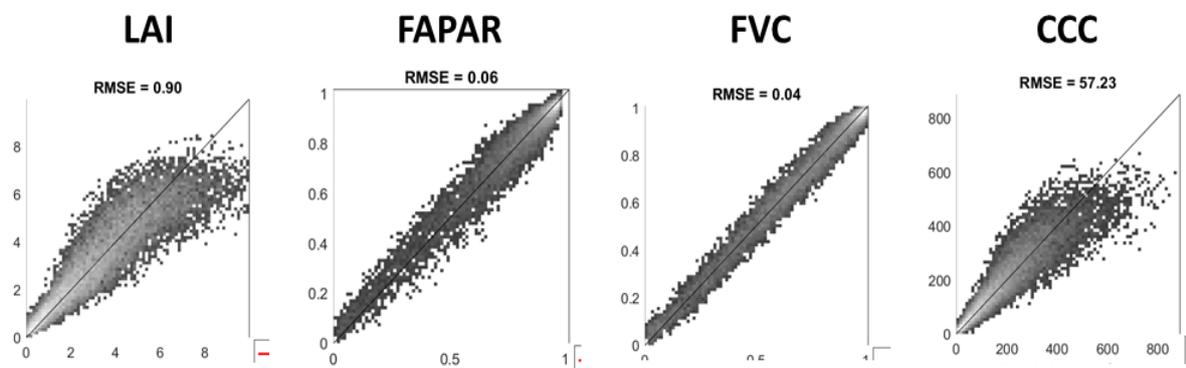
Le CES Paramètres biophysiques de la végétation propose quatre variables :

- L'indice foliaire vert (GAI : *Green Area Index*) : il représente la somme de la demi-surface des organes (feuilles, tiges, épis...) photosynthétiquement actif par unité horizontale de sol. Il est également souvent assimilé au terme LAI. Il correspond à la surface d'échange d'énergie et de masse entre la plante et l'atmosphère.
- La fraction de rayonnement photosynthétiquement actif absorbé par la végétation (fAPAR : *fraction of Absorbed Photosynthetically Active Radiation*). Le fAPAR intervient directement dans le processus de la photosynthèse et représente la quantité d'énergie absorbée par la plante pour produire de la biomasse.
- La fraction de couverture du sol par la végétation (FVC : *Fractional Vegetation Cover*) : représente la fraction de surface recouverte par la végétation verte lorsqu'on vise une surface à la verticale. Le FVC est indispensable pour séparer l'influence du sol de celle de la végétation dans les bilans d'énergie, en particulier pour l'estimation de l'évapotranspiration.
- Le contenu en chlorophylle du couvert végétal (CCC : *Canopy Chlorophyll Content*) est un bon indicateur de l'état azotée des plantes (Delloye et al., 2019). Une étude récente montre que ce type de variables est particulièrement intéressant dans l'utilisation de modèles de production primaire car le CCC détermine partiellement l'efficacité de la photosynthèse.

À l'échelle internationale, le LAI et le fAPAR sont par ailleurs reconnues par le GCOS (Global Climate Observing System) comme des variables essentielles du climat (ECV : Essential Climate Variables, <https://gcos.wmo.int/en/essential-climate-variables/table>).

L'algorithme d'estimation de ces variables, basé sur l'état de l'art (Verrelst et al., 2018), utilise des techniques d'apprentissage automatique (réseaux de neurones) entraînées sur une base de données

simulées. Il s'appuie sur l'expérience acquise pour le développement des produits issus des satellites moyenne résolution européens (MERIS, VEGETATION, PROBAV), dont l'efficacité a été démontrée et comparée avec les produits issus d'autres capteurs satellitaires (Baret et al., 2007 ; Weiss et al., 2014). Un modèle de transfert radiatif générique permet à partir d'un ensemble de données d'entrée (comprenant, entre autres, les différentes variables biophysiques) de simuler les réflectances telles que les mesurerait Sentinel-2. Un grand nombre de cas, environ 41 500, correspondant à des combinaisons variées de valeurs des paramètres d'entrée sont ainsi simulés. Les réseaux de neurones sont ensuite entraînés sur ces cas puis sont appliqués sur les images Sentinel-2 pour produire les cartes de variables biophysiques. L'implémentation de cet algorithme dans la chaîne Theia est en cours et devrait être opérationnelle fin 2020. La Figure 7 présente les performances théoriques d'estimation de ces algorithmes. Plusieurs exercices d'évaluation de ces produits à partir de mesures *in situ* ont également été publiés (Delloye et al., 2018 ; Estévez et al., 2020, Kganyago et al., 2020). La Figure 8 présente un extrait d'image Sentinel-2 dans la région de la Camargue et la carte de LAI obtenue. On observe une bonne cohérence entre les deux images, avec un faible LAI (couleur jaune) pour les parcelles de sol nu (en brun sur l'image RGB) et un fort LAI (couleur rouge) pour les parcelles de végétation dense (vert sombre sur l'image RGB).



**Figure 7 :** Performances de l'algorithme d'estimation des quatre variables biophysiques d'intérêt (données simulées).



**Figure 8 :** Exemple d'application des réseaux de neurones sur une image Sentinel-2 (région de la Camargue). À gauche, image en « vraie couleur » (RGB) ; à droite, carte de LAI estimée.

Cet algorithme a été développé pour fonctionner de façon générique, c'est-à-dire sans distinction entre les types de couverts végétaux (cultures, forêts, prairies). Il présente donc des capacités d'estimation variables selon le type de couvert considéré. Il est donc mieux adapté pour les surfaces agricoles et grandes cultures que pour la forêt qui présentent des architectures complexes. Des améliorations sont en cours, en particulier pour développer des réseaux de neurones mieux adaptés pour chaque type de culture ou pour des couverts forestiers. L'utilisation d'algorithme spécifique sera donc fortement dépendante de la qualité des cartes de couverture du sol développé par le CES OSO.

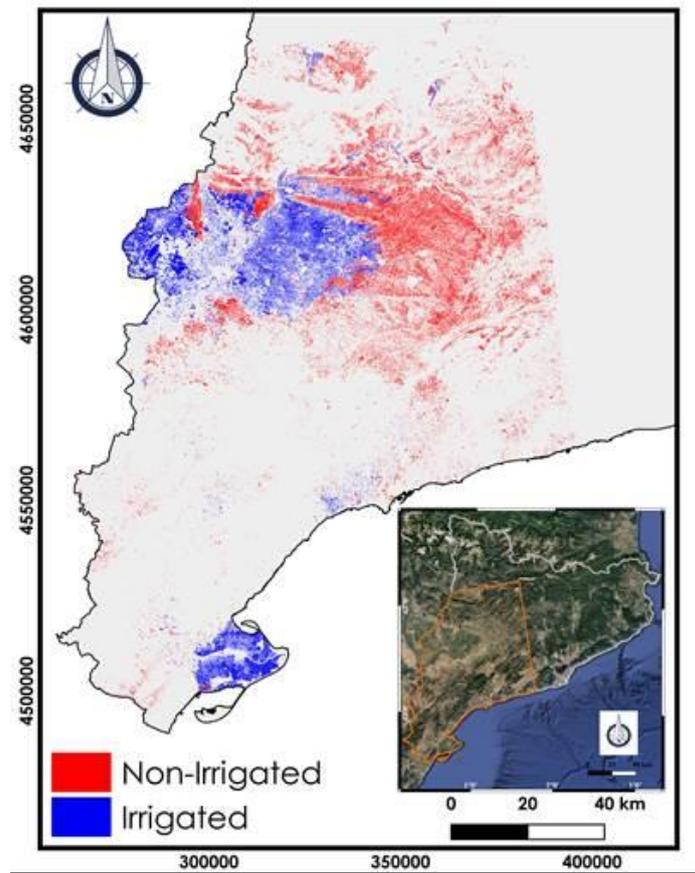
### 3.5 CES Irrigation

L'importance de l'irrigation dans la production agricole s'accroît d'autant que les conditions climatiques changent et qu'il est nécessaire de répondre à des besoins alimentaires mondiaux croissants. Une meilleure planification de l'irrigation est devenue indispensable pour répondre à la forte demande en nourriture due à l'augmentation de la population mondiale et pour faire face aux aléas climatiques à venir. La connaissance précise de l'étendue des zones irriguées est une étape nécessaire pour arriver à optimiser la gestion des ressources en eau.

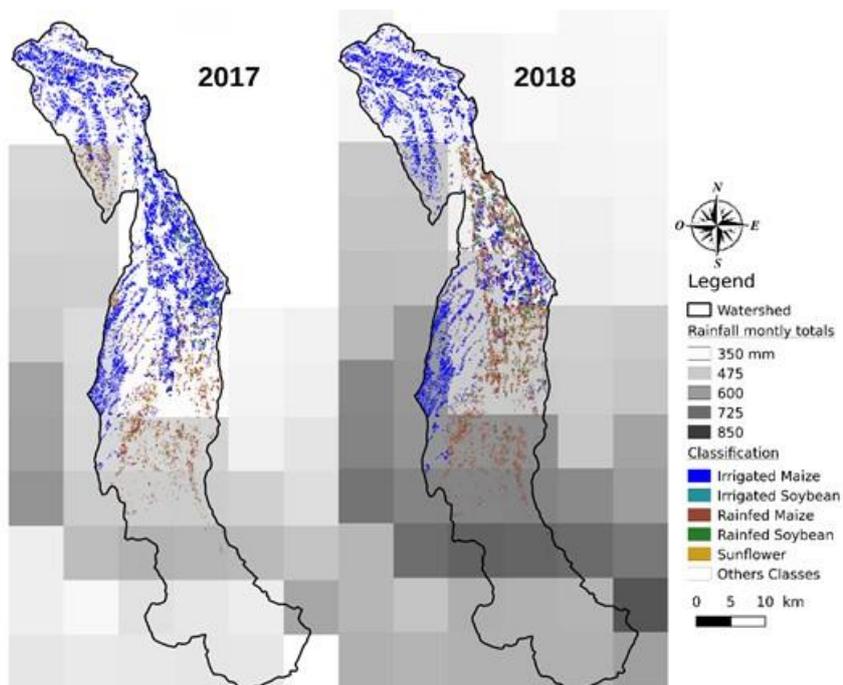
Malgré leur importance, l'étendue et la répartition des zones irriguées dans le monde restent encore incertaines. Les cartes existantes, en particulier celles qui couvrent de grandes superficies, ont été établies principalement à partir de statistiques nationales. La télédétection apparaît comme un outil efficace pour cartographier et surveiller les terres irriguées dans des conditions climatiques et des territoires variés. À l'échelle régionale peu d'efforts ont été entrepris jusqu'ici pour cartographier les zones irriguées en utilisant des données de télédétection, y compris des données optiques et radar (Boken et al., 2004 ; Gao et al., 2018 ; Thenkabail et al., 2005).

Le CES Irrigation du pôle Theia regroupe des équipes des deux UMRs Cesbio et TETIS qui développent des algorithmes pour la cartographie de surfaces irriguées à haute résolution spatiale. Ces algorithmes sont destinés aux chambres régionales d'agriculture, aux agences de l'eau ainsi qu'aux gestionnaires des bassins versants. Ils utilisent des données optiques issues principalement de Sentinel-2 et des données radar issues de Sentinel-1. La cartographie est réalisée à l'échelle parcellaire ou sub-parcellaire (périmètres irrigués, sous bassins versants).

Récemment, Bazzi et al. (2019) ont proposé un modèle de classification pour cartographier les zones irriguées à l'échelle de la parcelle en utilisant de données radar Sentinel-1 temporelles. L'approche proposée s'appuie sur des méthodes statistiques et mathématiques telles que l'analyse en composantes principales (ACP) et la transformation en ondelettes (TO). L'approche proposée a été testée avec des données de 2018 sur la région de Catalogne, dans le nord-est de l'Espagne. L'approche de classification est basée sur des forêts aléatoires (RF pour Random Forest) ; des réseaux neuronaux convolutifs (CNN) ont été utilisés pour construire des modèles de classification utilisant les paramètres de l'ACP ou de la TO. Les classificateurs RF construits à l'aide de l'ACP ou de la TO sur les séries temporelles S1 donnent de bons résultats pour la cartographie des zones irriguées avec une précision similaire de l'ordre de 90%. Quant à la classification avec les CNN, elle atteint une précision globale significative de 94,1 % (Figure 9). Plus récemment, Pageot et al. (2020) ont proposé une méthode de détection des parcelles irriguées et pluviales dans une zone tempérée (sud-ouest de la France) utilisant conjointement des séries temporelles optiques (Sentinel-2), radar (Sentinel-1) et météorologiques (SAFRAN), grâce à un algorithme de classification basé sur les Random Forest. Des indices cumulés mensuels calculés à partir de ces données satellitaires ont été utilisés dans le classificateur. Les données utilisées (deux années, 2017 et 2018) ont permis d'évaluer la robustesse de la méthode dans des conditions météorologiques différentes. L'utilisation combinée des données radar, optique et météo a permis d'obtenir des Fscore de 80 % pour les cultures irriguées et une précision globale de 70 % pour l'ensemble des cultures (irriguées et non-irriguées) (Figure 10).



**Figure 9** : Carte d'irrigation produite sur la Catalogne à l'aide des séries temporelles radar Sentinel-1, en utilisant une approche basée sur les réseaux de neurones (Bazzi et al., 2019).



**Figure 10** : Carte d'irrigation produite à l'aide des séries temporelles Sentinel-1/2 et de données météorologiques dans le sud-ouest de la France (Pageot et al., 2020).

Bien que d'énormes efforts aient été déployés pour quantifier l'étendue et la répartition spatiale des zones irriguées, les dates et la fréquence des irrigations n'ont pas encore fait l'objet d'une grande attention malgré leur grande importance dans la gestion des ressources en eau. Plus récemment, dans le cadre du CES Irrigation de Theia, Bazzi et al. (2020) ont proposé une approche opérationnelle et innovante basée sur l'arbre de décision pour détecter les événements d'irrigation à l'échelle de la parcelle en temps quasi réel et en utilisant des séries temporelles Sentinel-1. L'arbre de décision proposé s'appuie sur la détection de changements dans les coefficients de rétrodiffusion Sentinel-1 à l'échelle de la parcelle. Pour éliminer l'incertitude entre les précipitations et l'irrigation, l'étude a eu recours au signal Sentinel-1 et à des estimations de l'humidité du sol sur une grille de 10 km × 10 km issues du CES Humidité du sol à très haute résolution spatiale. Une méthode d'analyse a été ensuite construite pour détecter les événements d'irrigation à chaque date Sentinel-1, en introduisant des filtres supplémentaires afin de réduire les ambiguïtés dues au développement de la végétation liée au cycle de croissance des différents types de cultures ainsi qu'à la rugosité de la surface du sol. Un filtre utilisant le NDVI obtenu à partir des images optiques de Sentinel-2 a encore permis d'améliorer la détection des événements d'irrigation. L'apport de la méthode proposée réside ainsi dans la construction d'un outil non supervisé efficace pour la détection en temps quasi-réel des événements d'irrigation à l'échelle de la parcelle.

Les territoires sont soumis à des conditions climatiques et pédologiques différentes entraînant une grande diversité dans les pratiques culturales (grandes cultures d'été ou d'hiver, vignes, vergers, lavandin, etc.). De ce fait, un modèle élaboré sur un site et une année spécifique est difficilement applicable dans un autre contexte. La cartographie de l'irrigation nécessite donc de développer des approches robustes dans le temps et l'espace. C'est un des défis du CES Irrigation pour les années à venir.

Un autre objectif de ce CES est de proposer des cartes de volumes irrigués (dates et doses d'irrigation). Cette cartographie nécessite un couplage des données Sentinel-1/2 avec un modèle de bilan d'eau. Cette approche, déjà publiée (Battude et al., 2017), fait l'objet d'une thèse en cours (Y. Pageot, Cesbio) et sera poursuivie dans le cadre de projets à venir.

#### 4. Conclusions et perspectives

Depuis 2012, Theia permet une structuration des travaux scientifiques et rend visibles les résultats de ces travaux par une production avec passage à l'échelle et une diffusion des produits à valeur ajoutée qui en découlent. Du fait de leur production préalable dans Theia, plusieurs produits des CES Theia ont ainsi été sélectionnés pour contribuer aux réseaux européens Copernicus comme l'étendue du couvert neigeux et les hauteurs des lacs et rivières.

Theia permet également une structuration mutualisée des serveurs d'accès de ses différents partenaires en offrant depuis 2020 un catalogue centralisé d'accès ([catalogue.theia-land.fr](https://catalogue.theia-land.fr)) ainsi qu'un serveur cartographique ([maps.theia-land.fr](https://maps.theia-land.fr)) permettant un usage des données par des utilisateurs non spécialistes.

Le pôle Theia contribue enfin au développement de l'IR système Terre, Data Terra, qui ouvre des perspectives nouvelles en termes de partage des données et d'outils validés scientifiquement, pour observer, comprendre et prévoir de manière intégrée et transdisciplinaire le fonctionnement et l'évolution de l'environnement et du système Terre.

#### Références bibliographiques

Arrouays D., Grundy M.G., Hartemink A.E., Hempel J.W., Heuvelink G.B.M., Hong S.Y., Lagacherie P., Lelyk G., McBratney A.B., McKenzie N.J., Mendonca-Santos M.D., Minasny B., Montanarella L., Odeh I.O.A., Sanchez P.A., Thompson J.A., Zhang G.-L., 2014. GlobalSoilMap. Toward a Fine-Resolution Global Grid of Soil Properties, *Advances in Agronomy*.

- Baghdadi N., Choker M., Zribi M., El Hajj M., Paloscia S., Verhoest N., Lievens H., Baup F., Mattia F., 2016. A new empirical model for radar scattering from bare soil surfaces. *Remote Sensing*, vol. 8, Issue 11, pp. 1-14, doi: 10.3390/rs8110920.
- Baret F., Hagolle O., Geiger B., Bicheron P., Miras B., Huc M., Berthelot B., Nino F., Weiss M., Samain O., Roujean J.-L., Leroy M., 2007. LAI, fAPAR and fCover CYCLOPES global products derived from VEGETATION: Part 1: Principles of the algorithm. *Remote Sensing of Environment* 110, 275-286.
- Battude M., Al Bitar A., Brut A., Tallec T., Huc M., Cros J., Weber J.-J., Lhuissier L., Simonneaux V., Demarez V., 2017. Modeling water needs and total irrigation depths of maize crop in the south west of France using high spatial and temporal resolution satellite imagery, *Agricultural Water Management*, 189, 123-136, 2017.
- Bazzi H., Baghdadi N., Fayad I., Zribi M., Belhouchette H., Demarez V., 2020. Near Real-Time Irrigation Detection at Plot Scale Using Sentinel-1 Data. *Remote Sens.* 2020, 12, 1456; doi:10.3390/rs12091456.
- Bazzi H., Baghdadi N., Ienco D., El Hajj M., Zribi M., Belhouchette H., Escorihuela M.J., Demarez V., 2019. Mapping Irrigated Areas Using Sentinel-1 Time Series in Catalonia, Spain. *Remote Sensing*, *Remote Sens.* 2019, 11, 1836; doi:10.3390/rs11151836
- Boken, V.K., Hoogenboom, G., Kogan, F.N., Hook, J.E., Thomas, D.L., Harrison, K.A., 2004. Potential of using NOAA-AVHRR data for estimating irrigated area to help solve an inter-state water dispute. *International Journal of Remote Sensing* 25, 2277–2286
- Delloye C., Weiss M., Defourny P., 2018. Retrieval of the canopy chlorophyll content from Sentinel-2 spectral bands to estimate nitrogen uptake in intensive winter wheat cropping systems. *Remote Sensing of Environment* 216, 245-261.
- Estévez J., Vicent J., Rivera-Caicedo J.-P., Morcillo-Pallarés P., Vuolo F., Sabater N., Camps-Valls G., Moreno J., Verrelst J., 2020. Gaussian processes retrieval of LAI from Sentinel-2 top-of-atmosphere radiance data. *ISPRS Journal of Photogrammetry and Remote Sensing* 167, 289-304.
- Gao Q., Zribi M., Escorihuela M., Baghdadi N., Segui P., 2018. Irrigation Mapping Using Sentinel-1 Time Series at Field Scale. *Remote Sensing* 10, 1495. <https://doi.org/10.3390/rs10091495>.
- Hajj M., Baghdadi N., Zribi M., Bazzi H., 2017. Synergic Use of Sentinel-1 and Sentinel-2 Images for Operational Soil Moisture Mapping at High Spatial Resolution over Agricultural Areas, *Remote Sensing*, 9 (12), 1292.
- Huynh H., et al., 2019. L'infrastructure de recherche « Pôle de données et services pour le système Terre », à la pointe des techniques d'imagerie et de cartographie numérique. *Responsabilité & Environnement*, avr-2019.
- Inglada J., Vincent A., Arias M., Tardy B., Morin D., Rodes I., 2017. Operational high resolution land cover map production at the country scale using satellite image time series, *Remote Sensing*, 9(1), 95 (2017). <http://dx.doi.org/10.3390/rs9010095>.
- Kganyago M., Mhangara P., Alexandridis T., Laneve G., Ovakoglou G., Mashiyi N., 2020. Validation of sentinel-2 leaf area index (LAI) product derived from SNAP toolbox and its comparison with global LAI products in an African semi-arid agricultural landscape. *Remote Sensing Letters* 11, 883-892.
- Lagacherie, P., Walter, C., Arrouays, D., 2013. Cartographie numérique des sols : principe, mise en œuvre et potentialités. *Étude Gest. des Sols* 20, 83–98.
- Mulder, V.L., Lacoste, M., Richer-de-Forges, A.C., Arrouays, D., 2016. GlobalSoilMap France: High-resolution spatial modelling the soils of France up to two meter depth. *Sci. Total Environ.* 573, 1352–1369.
- Pageot Y., Baup F., Inglada J., Baghdadi N., Demarez V., 2020. Detection of irrigated and rainfed crops in temperate areas using Sentinel-1 and Sentinel-2 time series. *Remote Sens.* 2020, 12, 3044; doi:10.3390/rs12183044
- Thenkabail, P.S., Schull, M., Turrall, H., 2005. Ganges and Indus river basin land use/land cover (LULC) and irrigated area mapping using continuous streams of MODIS data. *Remote Sensing of Environment* 95, 317–341. <https://doi.org/10.1016/j.rse.2004.12.018>
- Verrelst J., Malenovsky, Z., Van der Tol C., Camps-Valls G., Gastellu-Etchegorry J.-P., Lewis P., North P., Moreno J., 2018. Quantifying Vegetation Biophysical Variables from Imaging Spectroscopy Data: A Review on Retrieval Methods. *Surveys in Geophysics*, 1-41.

Weiss M., Jacob F., Duveiller G., 2020. Remote sensing for agricultural applications: A meta-review. *Remote Sensing of Environment* 236, 111402.

Weiss M., Baret F., Block T., Koetz B., Burini A., Scholze B., Lecharpentier P., Brockmann C., Fernandes R., Plummer S., Myneni R., Gobron N., Nightingale J., Schaepman-Strub G., Camacho F., Sanchez-Azofeifa A., 2014. On Line Validation Exercise (OLIVE): A Web Based Service for the Validation of Medium Resolution Land Products. Application to FAPAR Products. *Remote Sensing* 6, 4190-4216.

Cet article est publié sous la licence Creative Commons (CC BY-NC-ND 3.0).



<https://creativecommons.org/licenses/by-nc-nd/3.0/fr/>

Pour la citation et la reproduction de cet article, mentionner obligatoirement le titre de l'article, le nom de tous les auteurs, la mention de sa publication dans la revue « *Innovations Agronomiques* », la date de sa publication, et son URL ou DOI).