



**HAL**  
open science

## **Iota2 ne fait pas que de la classification, elle produit aussi des indicateurs environnementaux !**

Mathieu Fauvel, Ludovic Arnaud, Eric Ceschia, Benjamin Tardy, Alice Lorillou, Mickaël Savinaud

### ► **To cite this version:**

Mathieu Fauvel, Ludovic Arnaud, Eric Ceschia, Benjamin Tardy, Alice Lorillou, et al.. Iota2 ne fait pas que de la classification, elle produit aussi des indicateurs environnementaux!. 2023. hal-04595790

**HAL Id: hal-04595790**

**<https://hal.inrae.fr/hal-04595790>**

Submitted on 31 May 2024

**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.



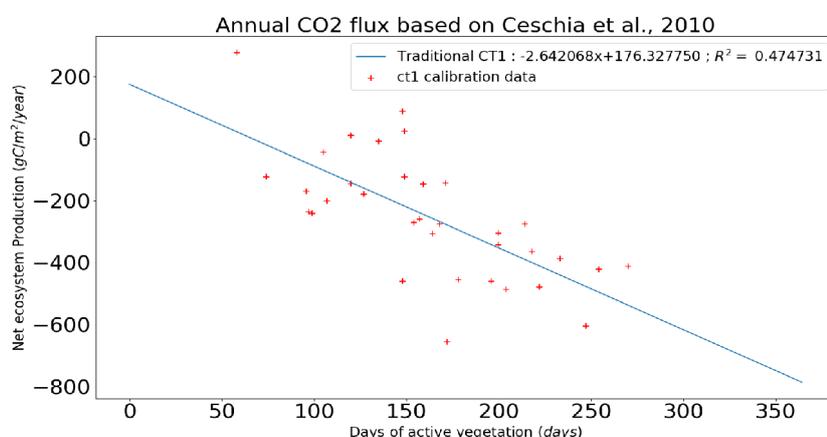
Distributed under a Creative Commons Attribution - NonCommercial - ShareAlike 4.0 International License



- en tant qu'indicateur de l'effet des pratiques sur les bilans de C pour la Politique Agricole Commune (PAC)
- pour quantifier l'évolution des stocks de C pour les marchés du C en agriculture.

Cet indicateur ayant été au préalable testé sur de petites zones en Europe, un des objectifs du projet SOCCROP était de tester la possibilité de l'appliquer sur de vastes territoires et dans des contextes pédoclimatiques plus contrastés (i.e. sur plusieurs continents).

Cet indicateur est calculé en utilisant la chaîne de traitement iota2 à partir des images Sentinel-2 de niveaux 2A afin d'estimer la fixation nette du CO<sub>2</sub> trimestrielle sur un an. L'indicateur du projet SOCCROP est basé sur une méthodologie très similaire à celle de l'indicateur "Carbon Tier 1" (CT1) développé dans le cadre du projet H2020 [NIVA](#). Le CT1 NIVA avait été développé en se basant sur la relation empirique décrite par Ceschia et al. (2010) reliant le flux net annuel de CO<sub>2</sub> (NEP, pour Net Ecosystem Production) aux nombres de jours où la végétation est photosynthétiquement active (NDAV pour "Number of days of active végétation") (Figure 1).



Cette relation s'appuie sur une quarantaine d'années de mesures de flux net de CO<sub>2</sub> obtenues par la méthode d'Eddy covariance des fluctuations turbulentes, cumulées sur une quinzaine de sites en Europe (couvrant [une large gamme de pédoclimats et types de culture](#)). Cette relation est applicable aux principales grandes cultures en Europe. L'indicateur NIVA CT1 utilise donc cette relation relativement simple pour estimer le flux net de CO<sub>2</sub> à partir de l'observation du nombre de jours où la végétation est active (NDAV) :

$$NEP = a \cdot NDAV + b \quad (1)$$

où les paramètres a et b ont été calibrés sur des mesures de flux du réseau [ICOS](#) réalisées en Europe selon  $a = -2.64 \text{ gC} \cdot \text{m}^{-2} \cdot \text{d}^{-1}$  et  $b = 176 \text{ gC} \cdot \text{m}^{-2}$  avec des erreurs respectives de  $\sigma_a = 0.49 \text{ gC} \cdot \text{m}^{-2} \cdot \text{d}^{-1}$  et  $\sigma_b = 84 \text{ gC} \cdot \text{m}^{-2}$ . Une expression analytique de l'incertitude peut être dérivée pour cet indicateur. Elle combine l'incertitude intrinsèque du modèle de régression (contenue dans l'incertitude des paramètres) avec l'incertitude sur la mesure du NDAV. Elle s'écrit :

$$\sigma_{NEP} = \sqrt{\sigma_a^2 N_{GD}^2 + \sigma_b^2 + \sigma_{N_{cv}}^2 a^2}$$

A ce stade, l'idée naturelle est d'utiliser la télédétection pour estimer le NDAV. Cela repose toutefois sur l'hypothèse forte que la végétation verte observée par satellite est toujours active d'un point de vue photosynthèse. A partir d'images optiques, on produit un indice de végétation comme le NDVI. Dans le cadre du CT1 NIVA, nous avons fait l'hypothèse que le NDAV est bien

approximé en comptant le nombre de jours où le NDVI est supérieur ou égal à 0.3 qui est une valeur typique caractérisant un sol nu.

Cependant, dans le cadre du projet SOCRROP, plusieurs améliorations ont été suggérées pour améliorer la précision de l'indicateur. En particulier la prise en compte de variables climatiques comme la température de l'air et le rayonnement global. En effet, en fonction des conditions climatiques, la végétation et le sol n'ont pas le même niveau d'activité.

Ainsi, une analyse précise des données expérimentales du réseau de stations de flux Européennes labellisées par [ICOS](#) montre que le modèle initial (1) peut être modifié sous la forme suivante :

$$NEP = c N_{GD} + d N_{AS} \quad (2)$$

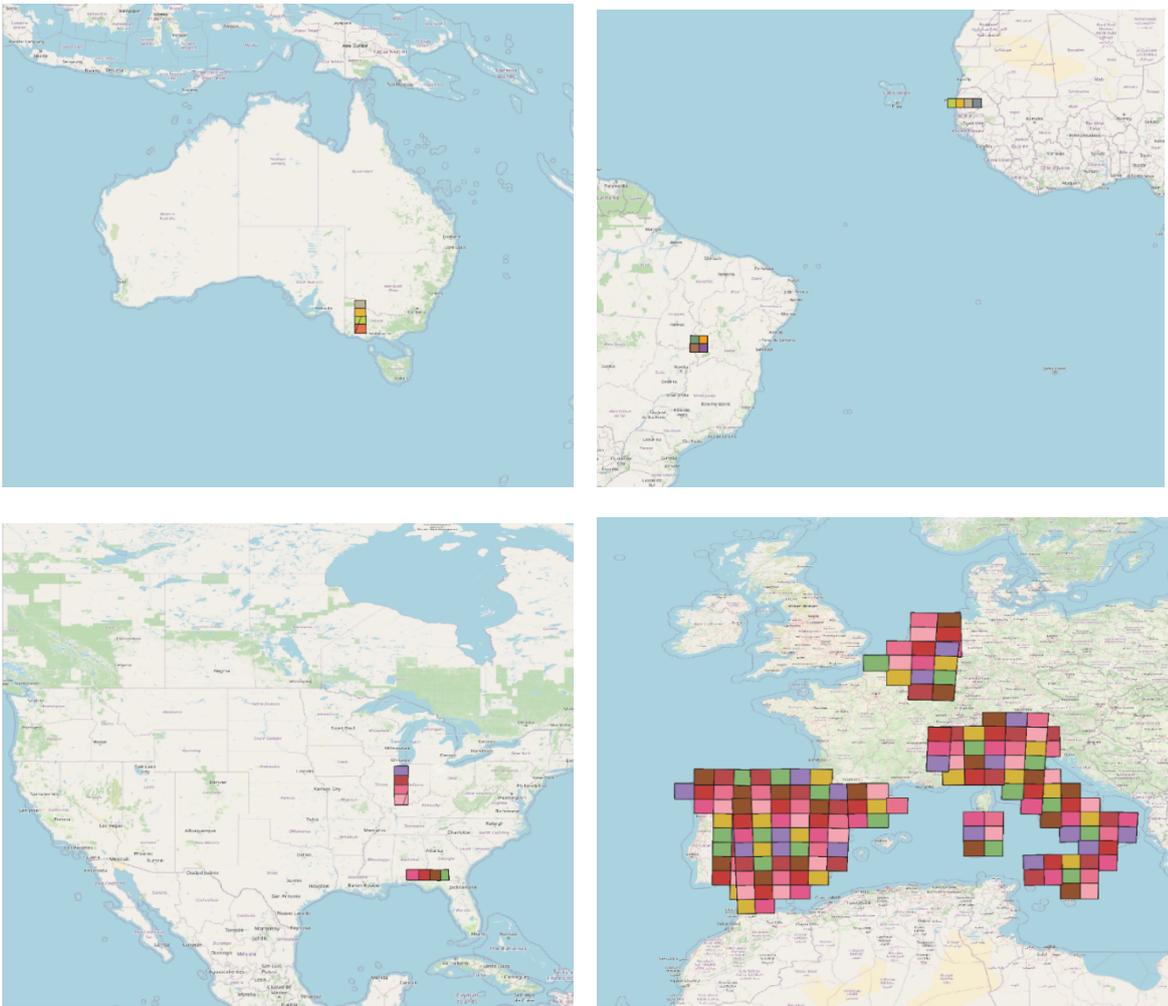
où  $N_{GD}$  est le nombre de jours vert,  $N_{AS}$  est le nombre de jours durant lesquels la respiration du sol est potentiellement plus active car il fait chaud, et c et d sont des paramètres de régression. Dans ce contexte,  $N_{GD}$  est défini comme le nombre de jours où  $NDVI > 0.3$ .  $N_{AS}$  est définie comme le nombre de jours tel que  $NDVI < 0.3$  et tel que le rayonnement global est supérieur à un certain seuil  $R_g^{th}$ . La calibration de ce modèle sur les données flux du réseau ICOS permet d'obtenir que  $c = -0.85 \pm 0.24 \text{ gC} \cdot \text{m}^{-2} \cdot \text{j}^{-1}$ ,  $d = 0.14 \pm 0.50 \text{ gC} \cdot \text{m}^{-2} \cdot \text{j}^{-1}$  et  $R_g^{th} = 65 \text{ W} \cdot \text{m}^{-2}$ . Comme pour l'indicateur originel, l'expression analytique de l'incertitude peut être dérivée comme :

$$\sigma_{NEP} = \sqrt{\sigma_c^2 N_{GD}^2 + \sigma_{N_{GD}}^2 c^2 + \sigma_d^2 N_{AS}^2 + \sigma_{N_{AS}}^2 d^2} \quad (3)$$

Dans le cadre du projet SOCCROP, c'est l'indicateur (2) avec son incertitude (3) qui sont considérés. Cependant, pour des raisons d'analyse, l'information du nombre de jours de vert peut se révéler intéressante. Ainsi, il a aussi été décidé de conserver le nombre de jours de vert par trimestre.

## Enjeux du projet

L'enjeu essentiel de ce projet pour l'équipe [CS GROUP](#) était de démontrer le passage à l'échelle du code développé par l'[INRAe](#) sur un environnement cloud en optimisant les temps de calcul et les coûts associés. Pour démontrer ce passage à l'échelle, 180 tuiles Sentinel-2 ont été produites avec cette chaîne. Ces tuiles couvrent, sur une année agricole entière, 4 pays européens (Belgique, Espagne, Italie et Pays-Bas) ainsi que qu'un ensemble de zones éco-climatiques variées en Australie, au Brésil, aux États-Unis (plus précisément en Géorgie au cœur de la Corn Belt) et au Sénégal.

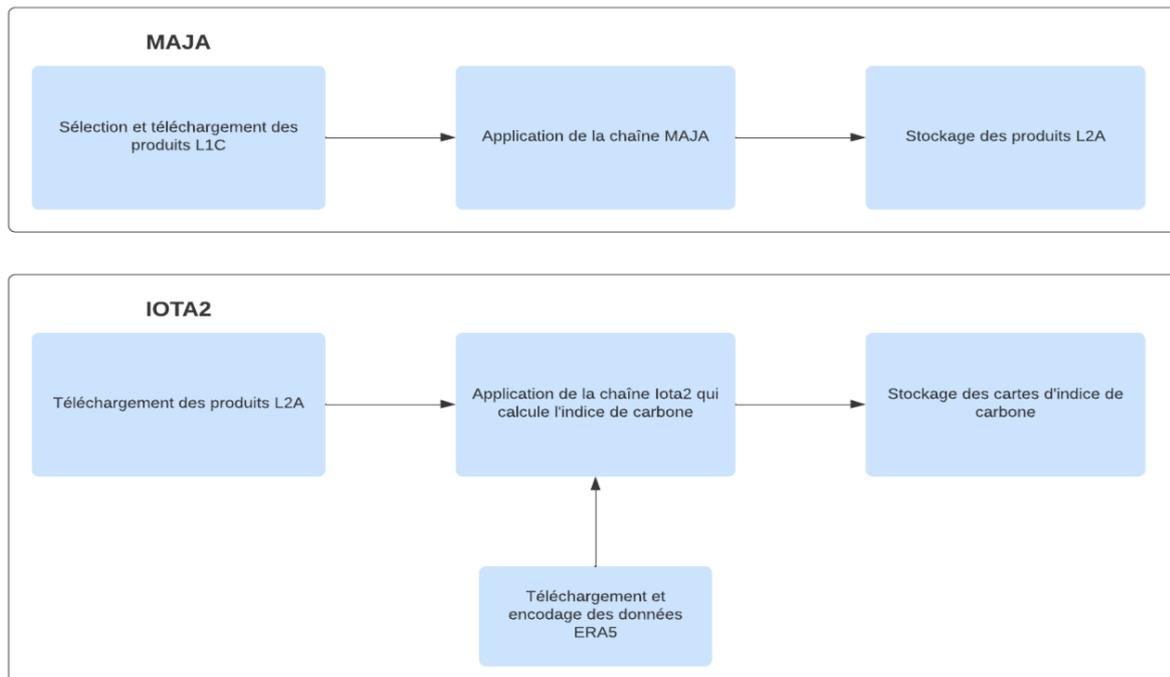


**Figure 2 : Répartition des tuiles**

Ce démonstrateur a été déployé sur l'infrastructure d'AWS, sur leur centre de calcul de Frankfurt pour bénéficier d'un accès optimisé aux données Sentinel-2 et de leur service d'orchestration de traitement serverless et managé Fargate. Pour ce faire, les chaînes de traitements [MAJA](#) et [iota2](#) ont été instanciées et configurées pour produire rapidement les cartes d'indices voulues en optimisant les coûts. L'objectif final de cette démonstration étant d'anticiper et d'estimer le coût d'une production annuelle de l'indice SOCCROP sur l'ensemble des surfaces continentales.

## Aspects techniques de la chaîne de traitement SOCCROP

La chaîne de traitement SOCCROP, mise en œuvre par CS Group, est constituée de 2 blocks de traitements, illustrée dans la figure 2, permettant la récupération des données Sentinel-2 et la production de l'indice.



Pour les données Sentinel-2 L2A non disponibles sur le site THEIA, la chaîne MAJA est utilisée pour produire des réflectances de surface (niveau 2A) à partir des images Sentinel-2 de niveau 1C. La chaîne MAJA a été choisie car elle permet la détection précise des nuages et de leurs ombres ainsi que la correction des effets atmosphériques sur des séries temporelles d'images. La précision des masques de nuages obtenus par MAJA offre une meilleure précision à l'indicateur SOCCROP qui est très sensible aux altérations du signal (nuage, saturation, ombre). Cette précision est nécessaire car elle évite les propagations d'erreurs lors d'étapes comme le sur échantillonnage temporel.

De plus, MAJA, dans ses dernières versions, offre la possibilité de ne produire qu'une partie des sorties (réflectances de surface ou masques) ce qui permet d'optimiser les temps de production et de réduire les volumes de données à stocker sur une année. En effet, seules les bandes rouges et proche-infrarouges et 3 masques (les masques de nuage, de saturation et de bord) sont nécessaires pour le calcul de l'indice SOCCROP par « iota2 ».

Le second pipeline encapsule la chaîne iota2 qui calcule l'indice de carbone à partir des bandes rouges et proche-infrarouges des produits L2A. La boîte à outils iota2 dispose de nombreuses fonctions permettant entre autres de calculer des cartes d'indices spectraux. La chaîne gère de manière automatique les données d'entrées, la configuration de la méthode d'interpolation temporelle et de correction des nuages. Elle offre également la possibilité d'exploiter des données tierces, comme les informations d'ensoleillement et de température disponible dans les données ERA5. De même, pour alléger les coûts de stockage et de calcul, iota2 a été modifié pour ne prendre en compte que les bandes et les masques nécessaires. Grâce à ces améliorations, l'espace stocké des images a pu être réduit de moitié et le temps de production des cartes diminué significativement. Le traitement par tuile natif de iota2 et les traitements hautement parallélisés offerts par l'OrfeoToolBox (<https://www.orfeo-toolbox.org/>), bibliothèque de traitement d'images sur laquelle les deux chaînes de traitement sont basées, ainsi que les capacités de distribution de traitements du service Fargate d'AWS ont permis de réaliser la production annuelle sur les 180 tuiles en un peu moins d'une semaine (hors temps de récupération des produits L2A depuis Theia).



RPG : MIS pour maïs grain, BTH pour blé tendre d'hiver, CZH pour colza d'hiver, SOG pour sorgho, PPH pour prairie permanente, LUZ pour luzerne, SOJ pour soja, MLO pour mélange d'oléagineux, TRN pour tournesol...voir

<https://files.geo.data.gouv.fr/link-proxy/www.geonormandie.fr/2020-01-24/5e2a6b991a98421bf73d869c/dc-dl-rpg-2-0.pdf>

Le tableau 1 présente les temps moyens de traitement pour chaque étape du pipeline. Le temps requis pour le traitement par MAJA est relativement long dans la mesure où une année entière de données Sentinel-2 est traitée d'un trait. Un découpage de la période temporelle pourrait être réalisé pour distribuer ces calculs, mais cela peut avoir un impact sur la qualité des produits. Les cartes de carbone sont produites en environ 7h par iota2. Ici la majorité du temps est dépensée dans l'interpolation temporelle de la série annuelle.

### **Temps de traitements des données L1C (peps)    Temps de traitement des données L2A (théia)**

Traitement par MAJA : 72h

Téléchargement L2A : 7h

Traitement iota2 : 7h

Traitement iota2 : 7h

Temps de traitement d'une tuile L1C : 79h

Temps de traitement d'une tuile L2A : 14h

**Tableau 1 : Coût d'exploitation d'une tuile pour une année de données Sentinel-2, dans le cadre SOCCROP**

## **Conclusion et perspectives**

La collaboration entre l'équipe CS GROUP et l'INRAE a permis de démontrer la possibilité d'un passage à l'échelle du code développé par l'INRAE sur un environnement cloud en optimisant les temps de calcul et les coûts associés. Cet outil nous a permis d'estimer les flux nets annuel de CO<sub>2</sub> des principales grandes cultures sur une bonne partie de l'Europe de l'Ouest et sur plusieurs autres zones test dans le monde grâce à l'utilisation des données Sentinel-2 et de la chaîne iota2.

L'indicateur a montré une bonne cohérence avec les pratiques connues sur le terrain comme la mise en œuvre de cultures intermédiaires (ex. en Bretagne ou en Géorgie).

C'est donc une approche assez opérationnelle permettant d'estimer un indicateur lié au bilan C des grandes cultures dans une optique de versement de primes environnementales (éco-schémas) pour la PAC ou dans un contexte d'inventaires nationaux d'émissions de GES. Cependant cette approche ne permet pas de prendre en compte l'effet de l'ensemble des pratiques sur les bilans C des parcelles. Pour ce faire, il faudrait intégrer au calcul de l'indicateur des données relatives aux récoltes et aux amendements organiques (ce qui correspond à la méthode TIER 2 du projet NIVA). C'est cette méthode qui serait à privilégier dans le cadre d'un financement des agriculteurs en fonction de la quantité de carbone qu'ils stockent. Pour y parvenir, les agriculteurs et les autorités devraient s'accorder pour que les données de pratiques soient accessibles à l'échelle de la parcelle.

L'approche a toutefois montré ses limites dans les zones à fort enneigement comme en Belgique ou au Brésil. Pour une production opérationnelle à l'échelle globale, il serait donc nécessaire d'utiliser des données satellitaires radar (ex. Sentinel-1) en complément des données optiques Sentinel-2. L'utilisation des données Sentinel-1 pour interpoler de manière opérationnelle les trous dans les

séries temporelles de NDVI issues de Sentinel-2 est explorée par plusieurs unités de recherches. Ce n'est donc probablement qu'une question de temps avant que l'approche mise en œuvre dans le cadre de SOCCROP puisse être appliquées de manière opérationnelle à l'échelle globale en s'appuyant sur l'utilisation combinée des données Sentinel 1 et 2.

## **Contributeurs**

Pour le CESBIO-INRAe: Ludovic Arnaud, Mathieu Fauvel et Eric Ceschia

Pour CS-Group: Alice Lorillou, Mickael Savinaud et Benjamin Tardy