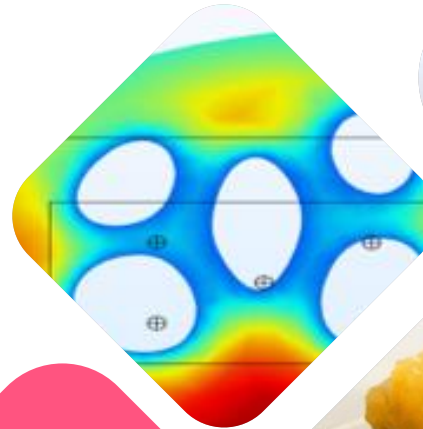


Publier un data paper

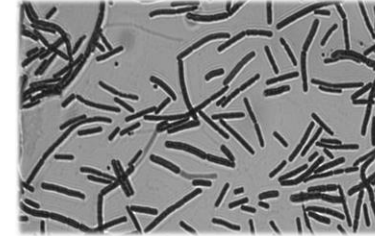
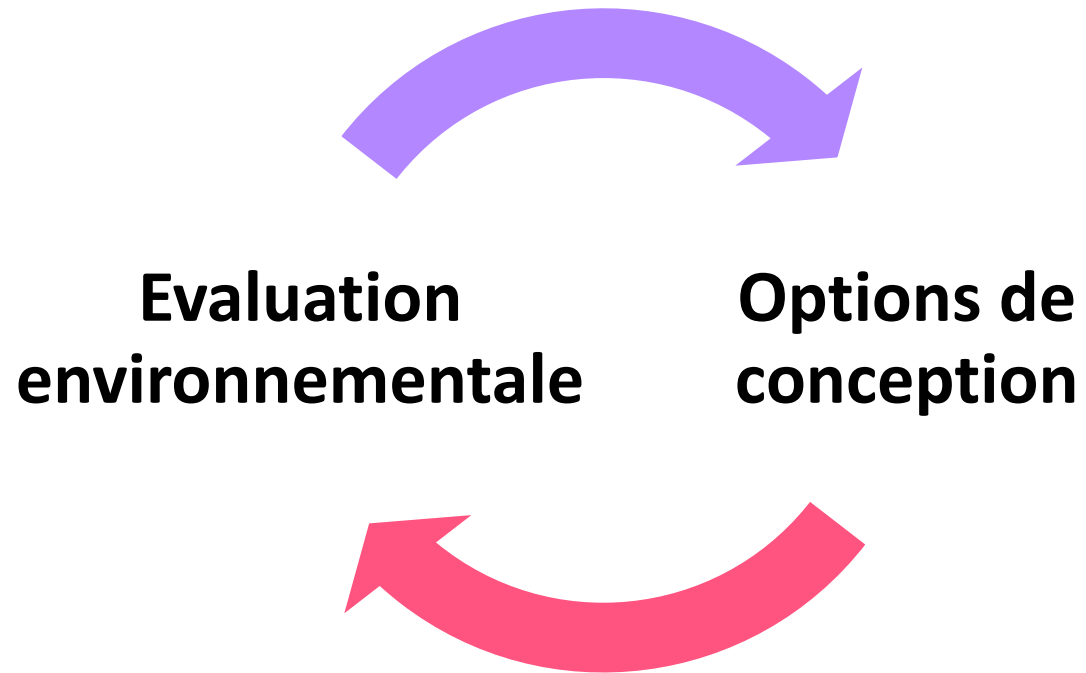
Journées science ouverte et données de la recherche AgroParisTech
5 avril 2023

Caroline Pénicaud



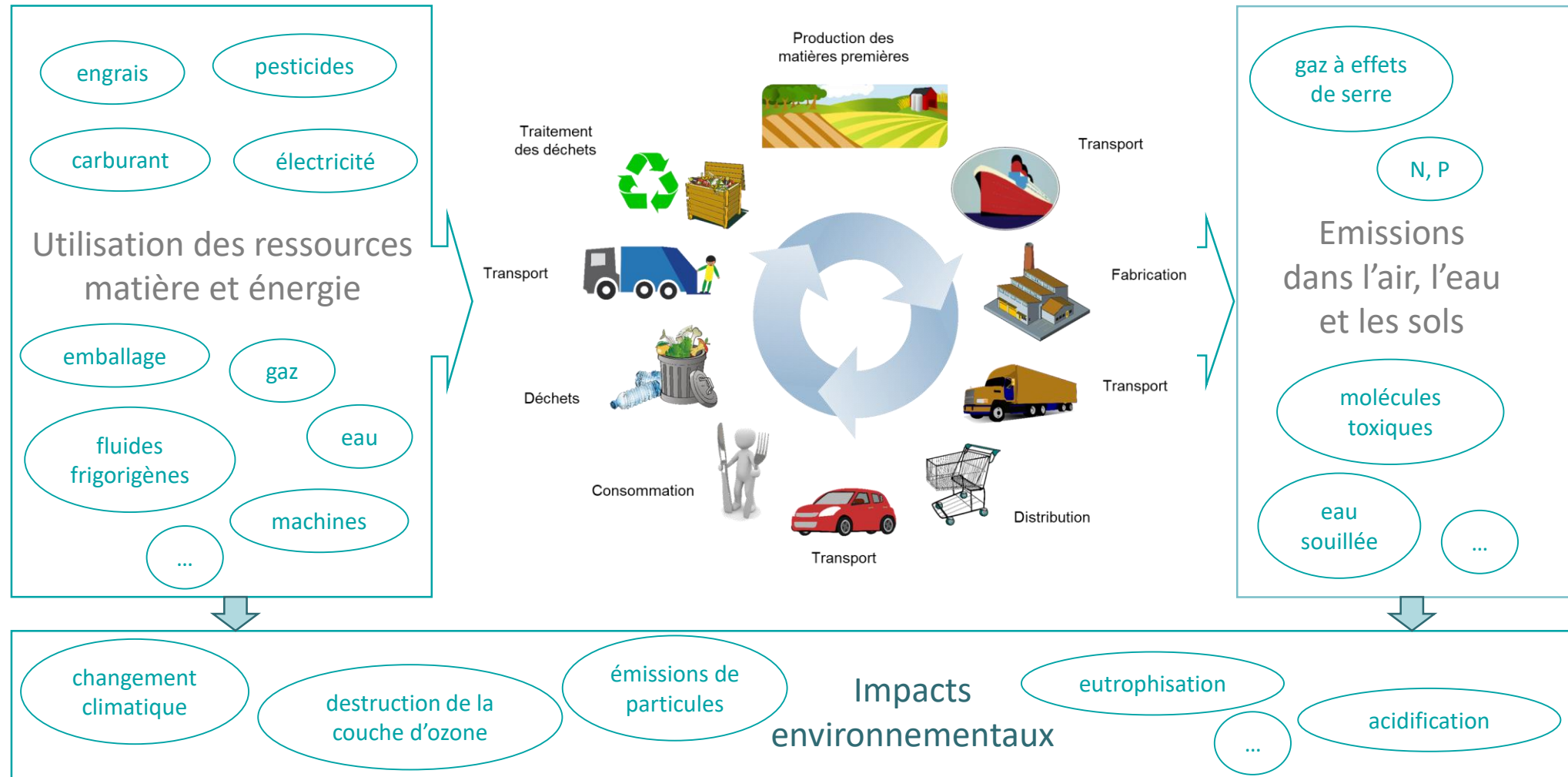
Mon domaine de recherche

- Ecoconception des aliments, des bioproduits et des (bio-)procédés alimentaires



Les données que je manipule

• Analyse du Cycle de Vie



Quels bénéfices à l'ouverture de ces données ?

- **Analyse du Cycle de Vie**
 - Méthode TRES coûteuse en données
 - Bilans matière et énergie
 - Tout au long du cycle de vie du produit
 - Bénéfice à avoir des retours pour éventuellement corriger/améliorer ce qu'on a fait
- **Bases de données ACV**
 - La plupart sont payantes
 - Pas toujours transparentes quant aux sources de données
- **Conviction de l'absolue nécessité de la Science Ouverte**
 - Restaurer la confiance dans la Science
 - « Tout seul on va plus vite, ensemble on va plus loin »

Dans quel contexte j'ai publié un data paper ?

- Pour accompagner un article « principal »

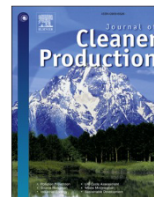
Journal of Cleaner Production 336 (2022) 130128



Contents lists available at ScienceDirect

Journal of Cleaner Production

journal homepage: www.elsevier.com/locate/jclepro



Does environmental impact vary widely within the same food category? A case study on industrial pizzas from the French retail market

Adeline Cortesi^a, Caroline Pénicaud^{a,*}, Anne Saint-Eve^a, Louis-Georges Soler^b, Isabelle Souchon^{a,c}

^a Université Paris-Saclay, INRAE, AgroParisTech, UMR SayFood, F-78850, Thiverval-Grignon, France

^b Université Paris-Saclay, INRAE, UR ALISS, 94205, Ivry-sur-Seine, France

^c Université D'Avignon, INRAE, UMR SQPOV, 84911, Avignon Cedex 9, France



Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib



Data Article

Life cycle inventory and assessment data for quantifying the environmental impacts of a wide range of food products belonging to the same food category: A case study of 80 pizzas representatives of the French retail market



Adeline Cortesi^a, Caroline Pénicaud^{a,*}, Anne Saint-Eve^a, Louis-Georges Soler^b, Isabelle Souchon^{a,c}

^a Université Paris-Saclay, INRAE, AgroParisTech, UMR SayFood, F-78850, Thiverval-Grignon, France

^b Université Paris-Saclay, INRAE, UR ALISS, 94205, Ivry-sur-Seine, France

^c Université d'Avignon, INRAE, UMR SQPOV, 84911 Avignon Cedex 9, France

- Apporter de la transparence sur les données utilisées
- Permettre la réutilisation des données
- Alléger le M&M de l'article

Dans quel contexte j'ai publié un data paper ?

- A la demande d'un journal suite la publication d'un dataset sur un entrepôt de données

Data in Brief 43 (2022) 108478



Data Article

Dataset on the Life Cycle Assessment of fructo- and galacto-oligosaccharides (FOS and GOS) produced by synthesis or hydrolysis



Esteban Gerbino^a, Camille Quentier^b, Caroline Pénicaud^{b,*}

^a Center for Research and Development in Food Cryotechnology (CIDCA, CCT-CONICET La Plata), La Plata, Argentina

^b Université Paris-Saclay, INRAE, AgroParisTech, UMR SayFood, 91120, Palaiseau, France

- Dataset publié en lien avec une communication orale dans un congrès
- Data In Brief nous a invité·es à soumettre un data paper pour présenter ces données qu'ils jugeaient d'intérêt

Dans quel contexte j'ai publié un data paper ?

- Pour valoriser des données qui ne vont pas forcément donner lieu à un article mais peuvent être utiles à d'autres

Data in Brief

Life cycle inventory and life cycle impact assessment datasets of PDO Feta production in Stymfalia region, Greece
--Manuscript Draft--

Manuscript Number:	DIB-D-23-00482
Article Type:	Data Article
Keywords:	Environmental assessment; LCA; cheese production; data inventory
Corresponding Author:	Caroline Pénicaud French National Institute for Agricultural Research INRAE Versailles, Île-de-France FRANCE
First Author:	Samuel Le Féon, Dr
Order of Authors:	Samuel Le Féon, Dr Andreas Papadakis Gwenola Yannou - Le Bris Julie Auberger Dimitrios Chatzitheodorou Joël Aubin Caroline Pénicaud

- Permettre la réutilisation des données

En pratique

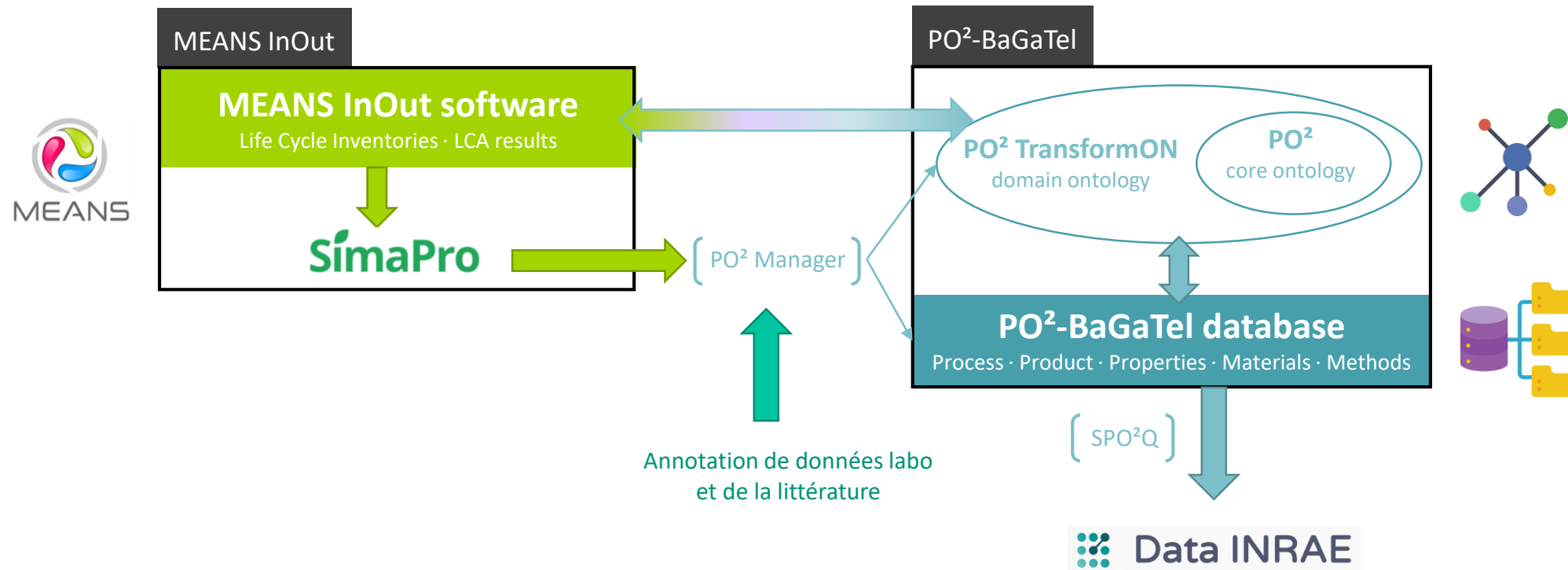
- **Données Excel : fichier doit pouvoir être « ingéré » en .tab**
 - 1 seul onglet par fichier
 - 1 seul tableau par fichier
 - Pas de couleurs / styles à conserver

[illegible]

- **Autres formats : pas tellement de contraintes**

Structuration des données

- Je n'étais pourtant pas complètement novice...



Publier les données sur un entrepôt

• Recherche Data Gouv

BaGaTeL Dataserve

(www.inrae.fr)

Recherche Data Gouv > Data INRAE > Experimental - Observation - Simulation Dataserve > BaGaTeL Dataserve >

Life Cycle Assessment of new fermented food products mixing cow milk and pea protein sources

Version 1.0



Pénicaud, Caroline; Hugué, Juliette; Chassard, Christophe; Lavigne, René; Irlinger, Françoise; Souchon, Isabelle; Marete, Stephan; Saint-Eve, Anne, 2023, "Life Cycle Assessment of new fermented food products mixing cow milk and pea protein sources", <https://doi.org/10.57745/X4QWKZ>, Recherche Data Gouv, V1, UNF:6:KRKRuU+eVU36WfHhb92aqQ== [fileUNF]

Citer le jeu de données

Pour en apprendre davantage sur le sujet, consulter le document Data Citation Standards [en].

Modalités d'accès au jeu de données

Modifier le jeu de données

Lier le jeu de données

Contact

Partager

Statistiques d'utilisation sur les jeux de données

24 consultations

0 téléchargement

0 citation

Description

Life Cycle Assessments (LCA) were performed to assess the environmental performance of 4 new fermented food products that mix animal (milk) and plant (pea) protein sources in different ratios (100% Pea, 75% Pea-25% Milk, 50% Pea-50% Milk, 25% Pea-75% Milk). The system perimeter goes from the agricultural production of the ingredients to the ready-to-eat products. Environmental impact results were obtained for 1 kg of ready-to-eat product, for all the environmental indicators calculated by the EF 3.0 Method on the SimaPro software. Life cycle inventories included the different flows in the LCA (raw materials, energy, water, cleaning products, packaging, transport, wastes). Foreground data have been acquired on the manufacturing site, background data were taken from the Ecoinvent 3.6 database. The dataset contains details on products, processes, equipment, infrastructures, mass and energy flows, Life Cycle Inventories (LCI) and Life Cycle Impact Assessment (LCIA) results. English

Sujet

Earth and Environmental Sciences; Engineering; Agricultural Sciences

Licence/Conditions

d'utilisation des données



etalab 2.0

Fichiers

Métadonnées

Conditions

Versions

Fichiers

Métadonnées

Conditions

Versions

Chercher dans ce jeu de données...



+ Téléverser des fichiers

Filtrer par

Type de fichier : Tout

Accès : Tout

Trier

1 à 4 de 4 Fichiers

Modifier les fichiers

Télécharger



LCIA_fermented_products.tab

Données tabulaires - 56.1 Ko

Publié 25 mars 2023

0 téléchargement

26 Variables, 135 Observations UNF:6:azW0...r3w==

Life Cycle Impact Assessment results obtained for 1 kg of each ready-to-eat product for all the environmental indicators evaluated by the "EF 3.0 Method (adapted) V1.00 / EF 3.0 normalization and weighting set": Climate change, Ozone depletion, Ionizing radiation, Photochemical ozone formation, Particulate matter, Human toxicity (non-cancer), Human toxicity (cancer), Acidification, Eutrophication (freshwater), Eutrophication (marine), Eutrophication (terrestrial), Ecotoxicity (freshwater), Land use, Water use, Resource use (fossils), Resource use (mineral and metals). Results are given by sub-step of the process as well as in total for each indicator.



LCI_fermented_products.tab

Données tabulaires - 43.0 Ko

Publié 25 mars 2023

0 téléchargement

10 Variables, 198 Observations UNF:6:6QX4...JWg==

Life Cycle Inventory for the four fermented products. All the mass and energy flows are indicated, grouped by step and sub-step of the production process. The measured flows refer to the total mass of the final product. These data need to be combined with information from the LCI_parameters dataset to obtain the exact LCI flows for each food product. The names of all the data used for the inventories and the associated database are also given. Some comments are added to help understand the meaning of the formulas describing the amounts of data.



LCI_others.tab

Données tabulaires - 3.3 Ko

Publié 25 mars 2023

0 téléchargement

8 Variables, 24 Observations UNF:6:PEKb...5IA==

Life Cycle Inventories of frozen lactic acid bacteria, obtained from [2] and adjusted based on the data used and the associated database. This file also contains the Life Cycle Inventory of household waste used in this study, together with the name of the data used and the associated database.



LCI_parameters.tab

Données tabulaires - 823 o

Publié 25 mars 2023

0 téléchargement

11 Variables, 8 Observations UNF:6:n+1J...S1g==

All complementary parameters needed to complete the LCI_fermented_products data (Pea_ratio, Milk_Ratio, Nb_molds, Losses_molding, Losses_demolding, Mass_product, Total_mass_final_product).



Publier les données sur un entrepôt

• Métadonnées

Ce jeu de données a été configuré pour utiliser la langue **English** pour toutes les entrées de métadonnées.

Métadonnées générales

Identifiant pérenne

doi:10.57745/X4QWKZ

Date de publication

2023-03-25

Titre

Life Cycle Assessment of new fermented food products mixing cow milk and pea protein sources

Point de contact

Utiliser le bouton de courriel ci-dessus pour joindre la personne-contact.

Auteur

Pénicaud, Caroline (INRAE)

Pénicaud, Caroline (Université Paris-Saclay, INRAE, AgroParisTech, UMR SayFood, 91120, Palaiseau, France) - ORCID: 0000-0002-1161-0272

Huguet, Juliette (Université Paris-Saclay, INRAE, AgroParisTech, UMR SayFood, 91120, Palaiseau, France)

Chassard, Christophe (Université Clermont Auvergne, INRAE, Vetagro Sup, UMRF, 15000 Aurillac, France)

Lavigne, René (Université Clermont Auvergne, INRAE, Vetagro Sup, UMRF, 15000 Aurillac, France)

Irlinger, Françoise (Université Paris-Saclay, INRAE, AgroParisTech, UMR SayFood, 91120, Palaiseau, France)

Souchon, Isabelle (Avignon Université, INRAE, UMR SQPOV, 84000 Avignon, France)

Marette, Stephan (Université Paris-Saclay, INRAE, AgroParisTech, UMR PSAE, 91120, Palaiseau, France)

Saint-Eve, Anne (Université Paris-Saclay, INRAE, AgroParisTech, UMR SayFood, 91120, Palaiseau, France)

Description

Life Cycle Assessments (LCA) were performed to assess the environmental performance of 4 new fermented food products that mix animal (milk) and plant (pea) protein sources in different ratios (100% Pea, 75% Pea-25% Milk, 50% Pea-50% Milk, 25% Pea-75% Milk). The system perimeter goes from the agricultural production of the ingredients to the ready-to-eat products. Environmental impact results were obtained for 1 kg of ready-to-eat product, for all the environmental indicators calculated by the EF 3.0 Method on the SimaPro software. Life cycle inventories included the different flows in the LCA (raw materials, energy, water, cleaning products, packaging, transport, wastes). Foreground data have been acquired on the manufacturing site, background data were taken from the Ecoinvent 3.6 database. The dataset contains details on products, processes, equipment, infrastructures, mass and energy flows, Life Cycle Inventories (LCI) and Life Cycle Impact Assessment (LCIA) results. English

Sujet

Earth and Environmental Sciences; Engineering; Agricultural Sciences

Type de données

Dataset

Informations sur la subvention

ANR: ANR-17-CE21-0003

Informations sur le projet

DIETPLUS (Effets des changements de régimes alimentaires sur l'équilibre des marchés, le partage de la valeur dans les filières, la santé publique, l'environnement et l'usage des sols)

Déposant

Pénicaud, Caroline

Date de dépôt

2022-09-06

• Prévisualisation des données .tab

Prévisualisation

Métadonnées

Versions

Explorer avec View Data

	Product	Impactcategory	Unit	Total	Peasuspensionproduction	Reconsti
1	100% Pea	Climate change	kg CO2 eq	6.299494086157211	0.43064306904326777	0.0
2		Ozone depletion	kg CFC11 eq	1.8503831566129986E-6	5.809575107035818E-8	0.0
3		Ionising radiation	kBq U-235 eq	0.1446579360020255	0.14478346983097584	0.0
4		Photochemical ozone formation	kg NMVOC eq	0.014386308781728606	0.0012961334401196273	0.0
5		Particulate matter	disease inc.	2.43578831741705E-7	3.693851525241374E-8	0.0
6		Human toxicity, non-cancer	CTUh	1.0056235391977985E-7	2.482781660292215E-8	0.0
7		Human toxicity, cancer	CTUh	4.164645042536197E-9	4.872134282337577E-10	0.0
8		Acidification	mol H+ eq	0.030254375991145305	0.005276050178073429	0.0
9		Eutrophication, freshwater	kg P eq	0.0015702867533810387	1.0927702069018702E-4	0.0
10		Eutrophication, marine	kg N eq	0.011309749663984093	0.005949271054574707	0.0
11		Eutrophication, terrestrial	mol N eq	0.06800542245541753	0.02045048601762971	0.0
12		Ecotoxicity, freshwater	CTUe	155.07048424443235	15.468212386872015	0.0
13		Land use	Pt	62.0723805359548	39.72927599462802	0.0
14		Water use	m3 depriv.	3.675494469732781	0.3891013164944962	0.0
15		Resource use, fossils	MJ	325.30984495045584	6.548437616694237	0.0
16		Resource use, minerals and metals	kg Sb eq	8.36878036123882E-5	8.353455764244581E-6	0.0
17		Climate change - Fossil	kg CO2 eq	6.2385525004934435	0.4081343704025707	0.0
18		Climate change - Biogenic	kg CO2 eq	0.05604581173192133	0.022271572456327157	0.0
19		Climate change - Land use and LU change	kg CO2 eq	0.00489577393184089	2.3712618437201415E-4	0.0
20		Human toxicity, non-cancer - organics	CTUh	4.305222455624948E-9	3.2743651054667495E-10	0.0
21		Human toxicity, non-cancer - inorganics	CTUh	9.408909437487095E-9	1.4945623245081021E-9	0.0
22		Human toxicity, non-cancer - metals	CTUh	8.734126813513596E-8	2.3027314704333402E-8	0.0
23		Human toxicity, cancer - organics	CTUh	1.557835489021798E-9	8.88936658840935E-11	0.0
24		Human toxicity, cancer - inorganics	CTUh	0.0	0.0	0.0
25		Human toxicity, cancer - metals	CTUh	2.6068095535143995E-9	3.983197623496653E-10	0.0
26		Ecotoxicity, freshwater - organics	CTUe	4.466182591017426	1.3638145020858938	0.0

Droits réservés © 2023, Recherche Data Gouv | Gestion du suivi d'audience | Conditions générales d'utilisation

Fourni par  v. 5.12.1 build 1122-cf90431

Générer un data paper depuis l'entrepôt...



Recherche Data Gouv [Génération datapaper](#)

(Recherche Data gouv)

Création d'un Data Paper

Cet outil vous permet de créer une ébauche de data paper (publication scientifique décrivant un jeu de données) à partir du DOI d'un jeu de données déposé dans l'entrepôt/catalogue entrepot.recherche.data.gouv.fr

Statistiques 556 943 téléchargements

L'entrepôt pluridisciplinaire *Recherche Data Gouv* est une solution souveraine pour le partage et l'ouverture des données de recherche produites par des d'un entrepôt disciplinaire reconnu. Il est basé sur le logiciel libre [Dataverse](#). Le dépôt des données doit se faire dans l'espace institutionnel attribué aux contributeurs. Un espace générique est dédié aux données produites par les établissements ne disposant pas d'espace dédié. Les tests sont à retrouver les actualités et événements de la plateforme [Recherche Data Gouv](#).

The *Recherche Data Gouv* multidisciplinary repository is a sovereign solution for sharing and opening up data produced by communities that do not have a recognized disciplinary repository. It is based on the [Dataverse](#) software. Data should be deposited in the space assigned to an institution which one of the contributor belongs to. A generic space is available for data produced by institutions which do not such a space yet themselves. Tests should be performed in the [sandbox](#).

The latest news and events on the [Recherche Data Gouv](#) platform.

Modèle

DOI

[Créer](#)

Modèles disponibles :

- Recherche Data Gouv
- Data In Brief



... ou le rédiger par soi-même

- Template word fourni par Data In Brief



[Browse journals](#) > [Data in Brief](#) > [Guide for authors](#)

Guide for Authors

 [Download Guide for Authors in PDF](#)

[Aims and scope](#) +

[Visit journal homepage](#) >

Submit your own Data in Brief

To submit a Data in Brief article directly to Data in Brief:

- 1 Fill in this [template](#).
- 2 Submit the document along with any computer code used to analyze the data through the journal's [online submission system](#). ⚡ Important: choose 'Data Article' as the article type when submitting directly, even if your data article is linked to another publication.

Template varie très régulièrement...
Bien vérifier à chaque nouveau data paper

- **Cas du data paper associé à un article : différentes expériences**
 - Co-soumission avec l'article principal
 - Soumission indépendante de l'article principal = Soumission sans article principal
 - Soumission après parution de l'article principal
- **Cas d'une soumission en lien avec article, surprenant pour moi de devoir publier d'abord l'article puis les données**
 - Données fournies pour le reviewing de l'article mais
 - Découpage article / data paper pas toujours bien compris par les reviewers
 - Confusion entre data paper et supplementary material
 - Même pour certains journaux !

Processus de publication

- Très similaire à un article classique
 - Editeur prend en charge le data paper
 - Envoi à un/des reviewers
 - Cycle(s) de révisions plus ou moins conséquentes
 - Acceptation
 - Publication

Et après ?

- Article publié comme un article classique
- Déjà quelques retours sur les données
 - Mise en évidence d'un oubli / une erreur
 - Versionnage permet de compléter/corriger sans perdre la traçabilité
 - Demandes de précisions
 - Demandes de données complémentaires

Données consultées avec une volonté de les réutiliser !!!

Ce que j'ai appris de ce format de publication

- A mettre en forme mes données
- A suffisamment documenter mes données avec des métadonnées
- A accepter qu'après publication il peut y avoir besoin de revenir sur les données pour corriger ou compléter les choses : « service après-vente » à assurer

- **Les données sont cruciales**
 - Permettre de les réutiliser pour aller plus loin
 - Y compris soi-même...
 - Apporter de la transparence pour mettre de la confiance
- **Publier un data paper : bénéfices**
 - Assurer une revue critique des données et métadonnées, au moins de leur mise en forme
 - Donner de la visibilité à ses données
 - Donner de la crédibilité à ses travaux
- **A quel coût ?**
 - En temps : modeste, format assez léger
 - Financier : environ 300 € pour Data In Brief (accord entre nos établissements et Elsevier)