Data in Brief 46 (2023) 108883

Contents lists available at ScienceDirect

Data in Brief

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

Data Article

# Dataset to monitor regionalised environmental impacts of the main agricultural products in Spain

Nelson Kevin Sinisterra-Solís[a,b,*], Neus Sanjuán[a], Javier Ribal[b], Vicent Estruch[b], Gabriela Clemente[a]

[a] ASPA Group. Dept. of Food Technology, Building 3F, Universitat Politècnica de València, Camí de Vera s/n, 46022 València, Spain
[b] Dept. of Economics and Social Sciences, Building 3P, Universitat Politècnica de València, Camí de Vera s/n, 46022 València, Spain

ARTICLE INFO

*Article history:*
Received 26 December 2022
Revised 29 December 2022
Accepted 2 January 2023
Available online 6 January 2023

Dataset link: R project to assess the environmental impacts of the main Spanish crops at a regional level (Reference data)

*Keywords:*
Agricultural LCA
Inventory data
Environmental inventory
NUTS 2
R programming

ABSTRACT

Estimating the average environmental impacts of a representative crop in a specific region is a helpful starting point from which to propose improvements in the agricultural sector. However, data collection from official representative sources is complex, and often they require subsequent treatment to be transformed into meaningful inventory data. This article shows a comprehensive dataset for obtaining inventory data and developing an environmental life cycle impact assessment of representative agricultural production corresponding to reference holdings at a regional level (NUTS 2) in Spain. The dataset comprises Excel files with the data compiled from secondary sources to be used in the assessment and the R code scripts to transform the data into relevant inventory data to estimate the environmental impacts of the reference holdings. This dataset is a reliable tool for researchers and other potential users to be used as a secondary information source for further studies. It can also be used to estimate the environmental impacts of the farming activity of agri-food products in other regions or countries by collecting similar data for the specific region and adjusting the R code.

DOI of original article: 10.1016/j.scitotenv.2022.158909
* Corresponding author.
*E-mail address:* nelsiso@doctor.upv.es (N.K. Sinisterra-Solís).
*Social media:* @sinisterrank (N.K. Sinisterra-Solís)

https://doi.org/10.1016/j.dib.2023.108883

## Specifications Table

| | |
|---|---|
| Subject | Agricultural Science<br>Environmental Science |
| Specific subject area | Agricultural Economics<br>Environmental engineering |
| Type of data | Tables in Excel files<br>R project including the abovementioned code scripts and folder structure |
| How the data were acquired | Data gathered from Spanish official statistics from 2010 to 2017, specialised literature.<br>Unit impacts from the unit processes estimated in GaBi software v10.6.1.35<br>Original R code developed to estimate the environmental impacts of representative agricultural production in Spain at a NUTS 2 level. |
| Data format | Raw data: xlsx<br>Processed (data conversion and data enrichment, network calculation): (xlsx)<br>Code: R files, R project file |
| Description of data collection | The collected data correspond to the parameters needed to estimate the resource consumption and environmental emissions of the main farming activities. Specifically, we collected data on material consumption and emissions from infrastructure building and management, fertiliser and pesticide production, on-field fertiliser and pesticide emissions, fuel consumption from machinery use, and irrigation. These data, together with the environmental impacts of the unit processes (e.g. electricity mix, production of agricultural inputs) are used to assess the environmental impacts from a data panel of reference holdings at the main Spanish NUTS 2. The data were compiled in a set of Excel sheets and subsequently operated on (tidying, transforming, visualisation and modelling operations) by developing a set of scripts using R programming language and RStudio interface. |
| Data source location | Raw data were extracted from the sources listed below. In the "Data description" section, these sources are related to the corresponding raw data of each excel file.<br>• Agriculture, Ecosystems & Environment Journal, https://doi.org/10.1016/J.AGEE.2016.10.006<br>• Department of Agriculture of the United State of America (USDA), https://www.usda.gov/<br>• European Environmental Agency (EEA), https://www.eea.europa.eu/<br>• European Soil Data Centre (ESDAC), https://esdac.jrc.ec.europa.eu/<br>• Eurostat, https://ec.europa.eu/eurostat/web/main/home<br>• Food and Agriculture Organization of the United Nations, https://www.fao.org/home/en<br>• GaBi software, https://sphera.com/<br>• INE-Instituto Nacional de Estadísticas (Spanish National institute of statistics), https://www.ine.es/en/index.htm<br>• International Journal of Life Cycle Assessment, https://doi.org/10.1007/s11367-013-0607-z<br>• Intergovernmental Panel on Climate Change (IPCC), https://www.ipcc.ch/<br>• Journal of Cleaner Production, https://doi.org/10.1016/J.JCLEPRO.2020.121656<br>• Environmental Science & Technology Journal,<br>• MAPA-Ministerio de Agricultura, Pesca y Alimentación (Spanish Ministry of Agriculture, Fisheries and Food), https://www.mapa.gob.es/es/<br>• MITECO-Ministerio para la Transición Ecológica y el Reto Demográfico (Spanish Ministry for the Ecological Transition and Demographic Challenge), https://www.miteco.gob.es/en/<br>• Nutrient Cycling in Agroecosystems Journal, https://doi.org/10.1007/S10705-006-9000-7<br>• PestLCI Consensus, https://pestlciweb.man.dtu.dk/<br>• SIAR-Sistema de Información Agroclimática para el Regadío (Spanish Agricultural INformation System for Irrigation), https://eportal.mapa.gob.es/websiar/NecesidadesHidricas.aspx<br>• USEtox software, https://usetox.org/<br>• WULCA working group, https://wulca-waterlca.org/ |

(*continued on next page*)

| | |
|---|---|
| Data accessibility | Repository name: Mendeley Data<br>Data identification number: [1] 10.17632/dd49c8y2cc.2<br>Direct URL to data: https://data.mendeley.com/datasets/dd49c8y2cc |
| Related research article | N. Sinisterra-Solís, N. Sanjuán, J. Ribal, V. Estruch, G. Clemente, An approach to regionalise the life cycle inventories of Spanish agriculture: Monitoring the environmental impacts of orange and tomato crops, Science of The Total Environment. 856 (2023) 158909. https://doi.org/10.1016/J.SCITOTENV.2022.158909. |

## Value of the Data

- This dataset provides a comprehensive approach to assess the environmental impacts of the main agricultural commodities corresponding to reference holdings at the Spanish NUTS 2 level.
- The dataset can be relevant for researchers and decision-makers who want to study the environmental impacts of the farming stage of agricultural commodities.
- The dataset can be used as secondary information source for other studies, and to assess the environmental impacts of the farming stage of agri-food products at the NUTS 2 level in Spain.
- The R script can be adapted to assess the environmental impacts of the farming stage in other regions or countries.

## Objective

This dataset was generated to provide the input data and the computational code to estimate the environmental impact of representative agricultural production in Spain at a NUTS 2 level. The methodology has been applied to monitor the environmental impacts of orange and tomato crops [2].

## 1. Data Description

### *1.1. Excel files*

The input data collected to develop the life cycle inventory data to assess the environmental impacts of the reference holdings were gathered in 21 excel files, which are described below.
*Dbi1_ECREA*: This dataset gathers information from the annual studies on costs and incomes of agricultural holdings of the Spanish Ministry of Agriculture, Fisheries and Food, also known by their acronym in Spanish ECREA. ECREA is a type of report of the Spanish farm accountancy data network (from now on ECREA-FADN). *Dbi1_ECREA* refers to 26 ECREA-FADN reports, available on 22 April 2022 [3]. The data file provides information of a set of reference holdings at the NUTS 2 level in Spain sampled in the ECREA-FADN reports considering different crops (e.g. tomato, orange, olive, etc.) and farm systems (greenhouse, open-field irrigated or open-field rainfed) in the years 2010 to 2017; specifically, the yield ($kg{\cdot}ha^{-1}$), the average surface of the holdings ($ha{\cdot}holding^{-1}$), expenses on fertilisers and fuel ($€{\cdot}ha^{-1}$), and macronutrients supplied ($kg\ macronutrient{\cdot}ha^{-1}$) are complied . Data corresponding to orange and tomato references holdings are provided in the supplementary material (SM-1) of the related research article [2]; therefore, to estimate life cycle inventory and the environmental impacts of orange and tomato reference holdings, this data should be taken from SM-1 of [2] and should be included in the corresponding cells of this file (*Dbi1_ECREA)* before running the R code. The columns “Id_holding” and “Id_holding_yr” are added as identification variables of the reference holdings and the corresponding year, respectively; whereas, “Key1”, “Key2” and “Key3” are added as key variables that help to relate this dataset with others. The dataset has been filtered to show only the data for tomato and orange reference holdings, which are the object of the case study [1].

Macronutrient data for some reference holdings in some years were not reported in the corresponding ECREA-FADN report. Therefore, these data are imputed (Table 1) by applying a rule of three, based on the yield . The data imputed are calculated in the Excel file.

*Dbi2_subnational_WS*: subnational water scarcity characterisation factors ($m^3$ world eq.·$m^3$ of water consumed in agriculture$^{-1}$) for Spanish NUTS 2 from [4].

*Dbi3_unit_impact*: environmental impact scores of the background unit processes and characterisation factors of the on-field emissions from fertilisers used to evaluate the environmental impacts of the reference holdings. Original data are taken from Ecoinvent v3.8 [5] and GaBi DB [6,7] databases, by using GaBi software [8]. To avoid copyright issues, the original values were replaced by 1 (non-real value); therefore, to replicate the estimation of the environmental impacts, those values should be replaced by the real ones by accessing the databases mentioned above.

*Dbi4_input_infrastructure*: material for building the greenhouse [9] and irrigation systems [10,11], and electricity consumption (kWh·ha$^{-1}$) to operate the vents of the greenhouse system [11].

*Dbi5_fertiliser_product*: data on the purchase value of the fertiliser products (€·kg$^{-1}$), without indirect tax, available in Spain in the years 2010 to 2017 [12,13] and their respective macronutrient content (kg of macronutrient · kg of fertiliser $^{-1}$) [14].

*Dbi6_pest_dose*: data on the dose of pesticide products used on different crops (kg of active substance·ha$^{-1}$) and surface areas (ha) with at least one application per year [15]. In addition, available USEtox characterisation factors (CFs), both recommended and interim, were complied for the midpoint human toxicity (total, cancer and non-cancer) and freshwater ecotoxicity impacts of the active substances [16,17].

*Dbi7_fertiliser_ef*: data on on-field emission factors (EF) to air from the application of nitrogen fertilisers. Specifically, Tier 2 EF for ammonia ($NH_3$) [18,19], Tier 1 EF for nitrogen oxides (NOx) [19,20], Tier 2 EF for direct nitrous oxide ($N_2O$) [21] and Tier 1 EF for indirect $N_2O$ from $NH_3$ volatilisation and $NO_3^-$ leaching [22] are shown.

*Dbi8_n_and_p_balances1*: data on supply of nitrogen and phosphorus and their balance at the end of 2016 for different crops at the Spanish NUTS 2 level [23,24].

*Dbi9_n_and_p_balances2*: data on the total supply of nitrogen and phosphorus in Spanish agriculture and their balances from 2010 to 2017 [25,26].

*Dbi10_pestl_consensus_sim*: results of primary pesticide distribution to air, soil and crop for a set of crops, modelled with PestLCI Consensus v1.0 [27] and considering the different settings of the technological input parameters for each crop in Spanish agriculture.

*Dbi11_land_surface*: annual time series (from 2010 to 2017) of the total, agricultural and freshwater surface area (ha) of the Spanish NUTS 2 [28].

*Dbi12_pest_CF_recipe*: ReCiPe CFs for the midpoint human toxicity (total, cancer and non-cancer) and freshwater ecotoxicity for the on-field emissions of the active substances and considering the three-time perspectives of the impacts, namely individualist, hierarchist and egalitarian [8].

*Dbi13_fuel_for_machinery*: monthly time series (2010 to 2017) of the purchase value of type B diesel at a the Spanish NUTS 2 [29].

*Dbi14_surface_irrigated*: annual time series (2010 to 2017) of agricultural surface area (ha) irrigated by furrow, sprinkler and drip methods at the Spanish NUTS 2 [28].

*Dbi15_water_source_and_use*: annual time series (2000 to 2017) of the quantity of water ($m^3$) irrigated by furrow, sprinkler and drip methods, as well as groundwater, surface water and other types of freshwater (reclaimed and desalinated) available at the Spanish NUTS 2 [30,31].

*Dbi16_ECREA_sample*: Contribution of NUTS3 to the number of holdings surveyed in the ECREAs for each crop at the Spanish NUTS 2 [3].

*Dbi17_NUTS*: Name of the NUTS 2 considered in ECREA-FADN reports and their respective NUTS 3 [32].

*Dbi18_water_irri1*: data on sowing and harvesting dates of the most representative crops of the NUTS 3 in the ECREA-FADN sample [33]. Dichotomous variables were created for the sowing and harvesting date, where 1 means that the respective month is part of the crop season, otherwise, it is 0. The dataset also provides information on the minimum and maximum rooting depth (m) and soil water depletion fraction (dimensionless) for the different crops [34].

**Table 1**
Reference holdings in which macronutrient data were imputed.

| Crop | System | Year | NUTS 2 |
|---|---|---|---|
| Almond | Irrigated open field | 2015 | Comunidad Valenciana |
| Almond | Irrigated open field | 2016 | Comunidad Valenciana |
| Almond | Irrigated open field | 2017 | Comunidad Valenciana |
| Almond | Rainfed open field | 2016 | Aragón |
| Almond | Rainfed open field | 2017 | Aragón |
| Almond | Rainfed open field | 2013 | Castilla-La Mancha |
| Almond | Rainfed open field | 2017 | Castilla-La Mancha |
| Almond | Rainfed open field | 2016 | Región de Murcia |
| Almond | Rainfed open field | 2017 | Región de Murcia |
| Almond | Rainfed open field | 2016 | Comunidad Valenciana |
| Almond | Rainfed open field | 2017 | Comunidad Valenciana |
| Apple | Irrigated open field | 2016 | Aragón |
| Apple | Irrigated open field | 2017 | Aragón |
| Apricot | Irrigated open field | 2016 | Región de Murcia |
| Apricot | Irrigated open field | 2017 | Región de Murcia |
| Persimon | Irrigated open field | 2016 | Comunidad Valenciana |
| Persimon | Irrigated open field | 2017 | Comunidad Valenciana |
| Lemon | Irrigated open field | 2016 | Región de Murcia |
| Lemon | Irrigated open field | 2017 | Región de Murcia |
| Lemon | Irrigated open field | 2016 | Comunidad Valenciana |
| Lemon | Irrigated open field | 2017 | Comunidad Valenciana |
| Mandarin | Irrigated open field | 2016 | Comunidad Valenciana |
| Mandarin | Irrigated open field | 2017 | Comunidad Valenciana |
| Nectarine | Irrigated open field | 2016 | Aragón |
| Nectarine | Irrigated open field | 2017 | Aragón |
| Nectarine | Irrigated open field | 2016 | Extremadura |
| Nectarine | Irrigated open field | 2017 | Extremadura |
| Nectarine | Irrigated open field | 2016 | Región de Murcia |
| Nectarine | Irrigated open field | 2017 | Región de Murcia |
| Nectarine | Irrigated open field | 2016 | Comunidad Valenciana |
| Nectarine | Irrigated open field | 2017 | Comunidad Valenciana |
| Orange* | Irrigated open field | 2016 | Andalucía |
| Orange | Irrigated open field | 2017 | Andalucía |
| Orange | Irrigated open field | 2016 | Región de Murcia |
| Orange | Irrigated open field | 2017 | Región de Murcia |
| Orange | Irrigated open field | 2016 | Comunidad Valenciana |
| Orange | Irrigated open field | 2017 | Comunidad Valenciana |
| Peach | Irrigated open field | 2016 | Aragón |
| Peach | Irrigated open field | 2017 | Aragón |
| Peach | Irrigated open field | 2016 | Extremadura |
| Peach | Irrigated open field | 2017 | Extremadura |
| Peach | Irrigated open field | 2016 | Región de Murcia |
| Peach | Irrigated open field | 2017 | Región de Murcia |
| Peach | Irrigated open field | 2016 | Comunidad Valenciana |
| Peach | Irrigated open field | 2017 | Comunidad Valenciana |
| Pear | Irrigated open field | 2016 | Aragón |
| Pear | Irrigated open field | 2017 | Aragón |
| Pear | Irrigated open field | 2016 | Región de Murcia |
| Plum | Irrigated open field | 2016 | Extremadura |
| Plum | Irrigated open field | 2017 | Extremadura |
| Plum | Irrigated open field | 2016 | Región de Murcia |
| Plum | Irrigated open field | 2017 | Región de Murcia |
| Alfalfa | Irrigated open field | 2016 | Aragón |
| Alfalfa | Irrigated open field | 2016 | Castilla y León |
| Alfalfa | Rainfed open field | 2016 | Castilla y León |
| Barley | Irrigated open field | 2016 | Aragón |
| Barley | Irrigated open field | 2016 | Castilla y León |
| Barley | Irrigated open field | 2016 | Castilla-La Mancha |
| Barley | Rainfed open field | 2016 | Aragón |
| Barley | Rainfed open field | 2016 | Castilla y León |
| Barley | Rainfed open field | 2016 | Castilla-La Mancha |

(*continued on next page*)

**Table 1** (*continued*)

| Crop | System | Year | NUTS 2 |
|---|---|---|---|
| Barley | Rainfed open field | 2016 | Extremadura |
| Sugar beet | Irrigated open field | 2016 | Andalucía |
| Sugar beet | Irrigated open field | 2016 | Castilla y León |
| Corn | Irrigated open field | 2016 | Andalucía |
| Corn | Irrigated open field | 2016 | Castilla y León |
| Corn | Irrigated open field | 2016 | Extremadura |
| Dried peas | Rainfed open field | 2010 | Aragón |
| Dried peas | Rainfed open field | 2016 | Aragón |
| Dried peas | Rainfed open field | 2012 | Castilla y León |
| Dried peas | Rainfed open field | 2016 | Castilla y León |
| Dried peas | Rainfed open field | 2016 | Castilla-La Mancha |
| Durum wheat | Irrigated open field | 2016 | Aragón |
| Durum wheat | Rainfed open field | 2016 | Andalucía |
| Durum wheat | Rainfed open field | 2016 | Aragón |
| Fodder corn | Irrigated open field | 2016 | Castilla y León |
| Forage vetch | Rainfed open field | 2016 | Castilla y León |
| Lentils | Rainfed open field | 2016 | Castilla y León |
| Oat | Rainfed open field | 2016 | Andalucía |
| Oat | Rainfed open field | 2016 | Castilla y León |
| Oat | Rainfed open field | 2016 | Castilla-La Mancha |
| Oat | Rainfed open field | 2016 | Extremadura |
| Rapeseed | Irrigated open field | 2016 | Castilla y León |
| Rapeseed | Rainfed open field | 2016 | Castilla y León |
| Ryegrass | Irrigated open field | 2016 | Aragón |
| Rice | Irrigated open field | 2016 | Andalucía |
| Rice | Irrigated open field | 2016 | Extremadura |
| Rye | Rainfed open field | 2016 | Aragón |
| Rye | Rainfed open field | 2016 | Castilla y León |
| Sunflower | Irrigated open field | 2016 | Castilla y León |
| Sunflower | Rainfed open field | 2016 | Andalucía |
| Sunflower | Rainfed open field | 2013 | Castilla y León |
| Sunflower | Rainfed open field | 2014 | Castilla y León |
| Sunflower | Rainfed open field | 2016 | Castilla y León |
| Sunflower | Rainfed open field | 2016 | Castilla-La Mancha |
| Soft wheat | Irrigated open field | 2016 | Aragón |
| Soft wheat | Irrigated open field | 2016 | Castilla y León |
| Soft wheat | Irrigated open field | 2016 | Extremadura |
| Soft wheat | Rainfed open field | 2016 | Andalucía |
| Soft wheat | Rainfed open field | 2016 | Aragón |
| Soft wheat | Rainfed open field | 2016 | Castilla y León |
| Soft wheat | Rainfed open field | 2016 | Castilla-La Mancha |
| Soft wheat | Rainfed open field | 2016 | Extremadura |
| Triticale | Rainfed open field | 2016 | Castilla-La Mancha |
| Olive for oil | Irrigated open field | 2016 | Andalucía |
| Olive for oil | Irrigated open field | 2017 | Andalucía |
| Olive for oil | Irrigated open field | 2016 | Aragón |
| Olive fot oil | Irrigated open field | 2017 | Aragón |
| Olive for oil | Irrigated open field | 2016 | Castilla-La Mancha |
| Olive for oil | Irrigated open field | 2017 | Castilla-La Mancha |
| Olive for oil | Irrigated open field | 2016 | Extremadura |
| Olive for oil | Irrigated open field | 2017 | Extremadura |
| Olive for oil | Rainfed open field | 2016 | Andalucía |
| Olive for oil | Rainfed open field | 2017 | Andalucía |
| Olive for oil | Rainfed open field | 2016 | Aragón |
| Olive for oil | Rainfed open field | 2017 | Aragón |
| Olive for oil | Rainfed open field | 2016 | Castilla-La Mancha |
| Olive for oil | Rainfed open field | 2017 | Castilla-La Mancha |
| Olive for oil | Rainfed open field | 2016 | Extremadura |
| Olive for oil | Rainfed open field | 2017 | Extremadura |
| Olive | Irrigated open field | 2016 | Andalucía |

(*continued on next page*)

**Table 1** (*continued*)

| Crop | System | Year | NUTS 2 |
|---|---|---|---|
| Olive | Irrigated open field | 2017 | Andalucía |
| Olive | Rainfed open field | 2016 | Andalucía |
| Olive | Rainfed open field | 2017 | Andalucía |
| Olive | Rainfed open field | 2016 | Extremadura |
| Olive | Rainfed open field | 2017 | Extremadura |
| Artichoke | Irrigated open field | 2016 | Región de Murcia |
| Broccoli | Irrigated open field | 2016 | Región de Murcia |
| Cantaloupe | Irrigated greenhouse | 2015 | Andalucía |
| Cantaloupe | Irrigated open field | 2016 | Castilla-La Mancha |
| Cantaloupe | Irrigated open field | 2016 | Región de Murcia |
| Celery | Irrigated open field | 2016 | Comunidad Valenciana |
| Chard | Irrigated greenhouse | 2016 | Comunidad Valenciana |
| Cucumber | Irrigated greenhouse | 2015 | Andalucía |
| Cucumber | Irrigated greenhouse | 2016 | Andalucía |
| Lettuce | Irrigated open field | 2016 | Región de Murcia |
| Onion | Irrigated open field | 2016 | Castilla-La Mancha |
| Onion | Irrigated open field | 2016 | Comunidad Valenciana |
| Pepper | Irrigated greenhouse | 2015 | Andalucía |
| Pepper | Irrigated greenhouse | 2016 | Andalucía |
| Pepper | Irrigated greenhouse | 2016 | Región de Murcia |
| Pepper | Irrigated greenhouse | 2016 | Comunidad Valenciana |
| Pepper for paprika | Irrigated open field | 2016 | Extremadura |
| Extra early potato | Irrigated greenhouse | 2016 | Comunidad Valenciana |
| Extra early potato | Irrigated open field | 2016 | Comunidad Valenciana |
| Mid-season potato | Irrigated open field | 2016 | Castilla y León |
| Strawberry | Irrigated greenhouse | 2015 | Andalucía |
| Strawberry | Irrigated greenhouse | 2016 | Andalucía |
| Tomato industry | Irrigated open field | 2016 | Andalucía |
| Tomato industry | Irrigated open field | 2016 | Extremadura |
| Tomato | Irrigated greenhouse | 2015 | Andalucía |
| Tomato | Irrigated greenhouse | 2016 | Andalucía |
| Tomato | Irrigated greenhouse | 2015 | Región de Murcia |
| Tomato | Irrigated greenhouse | 2016 | Región de Murcia |
| Tomato | Irrigated greenhouse | 2016 | Comunidad Valenciana |
| Watermelon | Irrigated greenhouse | 2015 | Andalucía |
| Watermelon | Irrigated greenhouse | 2016 | Andalucía |
| Watermelon | Irrigated open field | 2010 | Andalucía |
| Watermelon | Irrigated open field | 2015 | Andalucía |
| Watermelon | Irrigated open field | 2016 | Andalucía |
| Watermelon | Irrigated open field | 2016 | Castilla-La Mancha |
| Watermelon | Irrigated open field | 2016 | Región de Murcia |
| Watermelon | Irrigated open field | 2016 | Comunidad Valenciana |
| Courgette | Irrigated greenhouse | 2015 | Andalucía |
| Courgette | Irrigated greenhouse | 2016 | Andalucía |
| Wine grape | Irrigated open field | 2016 | Andalucía |
| Wine grape | Irrigated open field | 2017 | Andalucía |
| Wine grape | Irrigated open field | 2016 | Castilla-La Mancha |
| Wine grape | Irrigated open field | 2017 | Castilla-La Mancha |
| Wine grape | Rainfed open field | 2016 | Aragón |
| Wine grape | Rainfed open field | 2017 | Aragón |
| Wine grape | Rainfed open field | 2016 | Castilla y León |
| Wine grape | Rainfed open field | 2017 | Castilla y León |
| Wine grape | Rainfed open field | 2016 | Castilla-La Mancha |
| Wine grape | Rainfed open field | 2017 | Castilla-La Mancha |
| Wine grape | Rainfed open field | 2016 | Extremadura |
| Wine grape | Rainfed open field | 2017 | Extremadura |

*Dbi19_water_irri2_soil*: data on the composition of a sample of soil (clay, sand and silt content) at the Spanish NUTS 3, obtained from Lucas Topsoil 2015 DB [35,36]. By using "Soil water characteristics" software v6.02 [37], those soil composition data were used to define the class texture of the soil (e.g. loam, sandy loam, clay, etc.) and its respective available water was estimated (mm).

*Dbi20_water_irri3_Eto*: monthly time series (from 2010 to 2017) of precipitation and reference evapotranspiration in the NUTS 3 considered in the ECREA-FADN sample. To gather these data, four weather stations were selected randomly in each NUTS 3. For Ávila (one of the NUTS 3 of Castilla y León), due to data limitations, only three weather stations were considered [38].

*Dbi21_water_irri4_Kc*: values of the crop evapotranspiration coefficients for different crops (dimensionless). Data were obtained from the monthly time series (2010-2017) reported by the weather stations referenced in *Dbi20_water_irri3_Eto* [38]. A variable was created to count the number of times that each value appears in the time series.

### *1.2. R objects*

R scripts were developed to estimate the environmental impacts from the reference holdings studied, using as functional unit (FU) 1 kilogram of commercial product from each reference holding. The code was developed through nine scripts and the setup R file. The scripts include an explanation of the detailed procedure; however, a brief description is made below.

*SR0_library*: this script calls the extra packages that, together with R base functions, run the code. In addition, it creates the paths of the R project (*Project_path*) and the input data files (*InputData_path*).

*SR1_base_script*: code that calls the transversal input datasets to be used in the other scripts, such as "Dbi1_ECREA", "Dbi2_subnational_WS", and "Dbi3_unit_impact" files, and creates two more parameters, "n_simulation", which represents the number of simulations considered when developing the estimations, and "Diesel_density", which indicates the density of the diesel fuel. Through this script, tidying and transformation operations are applied to obtain suitable "ECREA" and "unit_process_impacts" Tibble objects to be related to the other scripts.

*SR2_infrastructure*: code to estimate the inventory data and environmental impacts ofthe greenhouse and the irrigation system infrastructure (e.g. plastics, steel, etc.), and managing the greenhouse (electricity consumption to operate the vents for greenhouse crops) of each reference holding.

*SR3_fertiliser_consumption*: code to estimate the inventory data and environmental impacts from the production of the fertilisers consumed in each reference holding.

*SR4_pesticide_consumption*: code to estimate the inventory data and environmental impacts from the production of the pesticide active substances consumed in each reference holding.

*SR5_on_field_emissions*: code to estimate the inventory data (on-field emissions) and environmental impacts from the application of fertiliser and pesticide products. In particular, emissions from fertiliser ($NH_3$, $NO_x$, $N_2O$, $NO_3^-$ and $PO_4^{3-}$) and pesticide application were calculated.

*SR6_machinery_use*: code to estimate the inventory data and environmental impacts from machinery use, particularly from the fuel consumption, in on-field operations, except those from irrigation.

*SR7_irrigation*: code to estimate the inventory data and environmental impacts from irrigation, taking into account water and energy consumption.

*SR8_impacts_agg*: code to join the environmental impacts estimated to the other scripts.

## 2. Experimental Design, Materials and Methods

From the ECREA-FADN reports, a reference dataset was created (*Dbi1_ECREA*) in which the reference holdings studied were defined. Additional information, as specific as possible, from

other secondary sources, was gathered to develop the life cycle inventory data to assess the environmental impacts of the reference holdings. This analysis is restricted to the farming stage, and boundaries are, thus, set at the farm gate including all the relevant stages: material consumption and emissions from infrastructure building and management, fertiliser and pesticide production, on-field fertiliser and pesticide emissions, fuel consumption from machinery use, and irrigation. Most of the collated data were compiled in Excel files; only some parameters for the estimation of energy for irrigation were included directly in the R code. By and large, in the R code, functions such as filter, select, join, summarise, simulate, etc., were applied to obtain the inventory data and the environmental impact scores of the reference holdings from the input data. For a correct code execution, file names and variable names should be kept as defined in the data files and R scripts. The file location and folder structure should remain the same, too.

### *2.1. Software*

| Name | Type | Source |
|---|---|---|
| "R" version 4.1.4 | Programming language | [39] |
| "RStudio" version 2022.2.3.492 | Programming interface | [40] |
| "Tidyverse" version 1.3.1 | Package | [41] |
| "Openxlsx" version 4.2.4 | Package | [42] |
| "Linprog" version 0.9-2 | Package | [43] |
| "Triangle" version 0.12 | Package | [44] |
| "Feather" version 0.3.5 | Package | [45] |

## Ethics Statements

The authors declare that this paper complies with the ethical requirements for publication in Data in Brief, namely:

- Authorship of the paper: Authorship is limited to those who have made a significant contribution to the conception, design, execution, or interpretation of the reported study.
- Originality and plagiarism: The authors ensure that they have written entirely original works, and in the case where words of others have been used, that this has been appropriately cited.
- Data access and retention: When required, authors are available to provide the raw data in connection with a paper for editorial review, and provide public access to such data.
- Multiple, redundant or concurrent publication: this manuscript has not been presented for publication in other journals.
- Acknowledgement of sources: In this study, there is no work of other authors to be acknowledged.
- Disclosure and conflicts of interest: this submission includes disclosure of all relationships that could be viewed as presenting a potential conflict of interest.
- Fundamental errors in published works: when required, the authors will notify the journal editor or publisher and cooperate with the editor to retract or correct the paper, in case of a significant error or inaccuracy in the published work.

Besides, the authors confirm that this study does not involve human subjects, animal experiments and data collected from social media platforms

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper/article.

## Data Availability

R project to assess the environmental impacts of the main Spanish crops at a regional level (Reference data) (Mendeley Data).

## CRediT Author Statement

**Nelson Kevin Sinisterra-Solís:** Conceptualization, Methodology, Software, Formal analysis, Data curation, Investigation, Writing – original draft; **Neus Sanjuán:** Conceptualization, Methodology, Formal analysis, Investigation, Writing – review & editing; **Javier Ribal:** Conceptualization, Methodology, Software, Writing – review & editing, Supervision; **Vicent Estruch:** Methodology, Validation, Formal analysis, Investigation; **Gabriela Clemente:** Conceptualization, Methodology, Investigation, Investigation, Writing – review & editing.

## Acknowledgements

The authors gratefully acknowledge both the Universitat Politècnica de València for providing the funds for N.K. Sinisterra-Solís's research contract through Subprogram 1 (PAID-01-18) and also the Colombian Government for their funding through COLCIENCIAS.

## References

[1] N.K. Sinisterra-Solís, N. Sanjuan, J. Ribal, V. Estruch, G. Clemente, R project to assess the environmental impacts of the main Spanish crops at a regional level - Mendeley Data, (n.d.). https://doi.org/10.17632/dd49c8y2cc.2.
[2] N. Sinisterra-Solís, N. Sanjuán, J. Ribal, V. Estruch, G. Clemente, An approach to regionalise the life cycle inventories of Spanish agriculture: Monitoring the environmental impacts of orange and tomato crops, Sci. Total Environ. 856 (2023) 158909, doi:10.1016/J.SCITOTENV.2022.158909.
[3] Ministerio de Agricultura Pesca y Alimentación (MAPA), Estudios de costes de explotaciones agrícolas, (2022). https://www.mapa.gob.es/es/ministerio/servicios/analisis-y-prospectiva/ECREA_Informes-Agricolas.aspx (accessed April 27, 2022).
[4] A.M. Boulay, L. Lenoir, Sub-national regionalisation of the AWARE indicator for water scarcity footprint calculations, Ecol Indic 111 (2020) 106017, doi:10.1016/J.ECOLIND.2019.106017.
[5] G. Wernet, C. Bauer, B. Steubing, J. Reinhard, E. Moreno-Ruiz, B. Weidema, The ecoinvent database version 3 (part I): overview and methodology, Int. J. Life Cycle Assess. 21 (2016) 1218–1230, doi:10.1007/S11367-016-1087-8.
[6] SPHERA, GaBi Databases, GaBi Databases & Modelling Principles 2022. (2022). https://gabi.sphera.com/international/databases/gabi-databases/(accessed April 28, 2022).
[7] SPHERA, GaBi Databases & Modelling Principles 2022, (2022). https://gabi.sphera.com/support/gabi/gabi-modelling-principles/(accessed April 28, 2022).
[8] SPHERA, GaBi Software, (2022). https://gabi.sphera.com/international/software/(accessed April 28, 2022).
[9] A. Antón, M. Torrellas, V. Raya, J.I. Montero, Modelling the amount of materials to improve inventory datasets of greenhouse infrastructures, Int. J. Life Cycle Assess. 19 (2013) 29–41, doi:10.1007/s11367-013-0607-z.
[10] B. Martin-Gorriz, B. Gallego-Elvira, V. Martínez-Alvarez, J.F. Maestre-Valero, Life cycle assessment of fruit and vegetable production in the Region of Murcia (south-east Spain) and evaluation of impact mitigation practices, J. Clean Prod. 265 (2020) 121656, doi:10.1016/J.JCLEPRO.2020.121656.
[11] A. Antón, M. Torrellas, M. Núñez, E. Sevigné, M.J. Amores, P. Muñoz, J.I. Montero, Improvement of agricultural life cycle assessment studies through spatial differentiation and new impact categories: case study on greenhouse tomato production, ACS Publications 48 (2014) 9454–9462, doi:10.1021/es501474y.
[12] MAPA, Ministerio de Agricultura, Pesca y Alimentación de España, ÍNDICES Y PRECIOS PAGADOS AGRARIOS. RESULTADOS EN BASE 2010. (2022). https://www.mapa.gob.es/es/estadistica/temas/estadisticas-agrarias/economia/precios-percibidos-pagados-salarios/precios-pagados-por-los-agricultores-y-ganaderos/default.aspx (accessed May 14, 2022).
[13] MAPA, Ministerio de Agricultura, Pesca y Alimentación de España, ÍNDICES Y PRECIOS PAGADOS AGRARIOS. RESULTADOS EN BASE 2005. (2022). https://www.mapa.gob.es/es/estadistica/temas/estadisticas-agrarias/economia/precios-percibidos-pagados-salarios/precios-pagados-por-los-agricultores-y-ganaderos/default.aspx (accessed May 14, 2022).
[14] MARM, Ministerio de Medio Ambiente, y Medio Rural y Marino. Guía práctica de la fertilización racional de los cultivos en España. Parte 1., 2010. https://www.mapa.gob.es/es/agricultura/publicaciones/Publicaciones-fertilizantes.aspx (accessed June 17, 2022).
[15] Ministerio de Agricultura Pesca y Alimentación (MAPA), Estadística anual de consumo de productos fitosanitarios y Estadística quinquenal de utilización de productos fitosanitarios en la Agricultura, Encuesta de Utilización de Productos Fitosanitarios, Resultados 2019. (2021). https://www.mapa.gob.es/es/estadistica/temas/estadisticas-agrarias/agricultura/estadisticas-medios-produccion/fitosanitarios.aspx (accessed April 27, 2022).

[16] R.K. Rosenbaum, T.M. Bachmann, L.S. Gold, M.A.J. Huijbregts, O. Jolliet, R. Juraske, A. Koehler, H.F. Larsen, M. MacLeod, M. Margni, T.E. McKone, J. Payet, M. Schuhmacher, D. van de Meent, M.Z. Hauschild, USEtox - The UNEP-SETAC toxicity model: Recommended characterisation factors for human toxicity and freshwater ecotoxicity in life cycle impact assessment, Int. J. Life Cycle Assess. 13 (2008) 532–546, doi:10.1007/S11367-008-0038-4.
[17] UNEP/SETAC, USEtox, v2.12, (2019). https://usetox.org/model/download (accessed April 28, 2022).
[18] J. Kuenen, C. Dore, Uncertainties, in: EMEP/EEA Air Pollutant Emission Inventory Guidebook 2019: Technical Guidance to Prepare National Emission Inventories, Office of the European Union, Luxembourg, 2019.
[19] N. Hutchings, J. Weeb, B. Amon, Crop production and agricultural soils 2019, in: EMEP/EEA Air Pollutant Emission Inventory Guidebook 2019, Luxembourg, 2019.
[20] E. Stehfest, L. Bouwman, N2O and NO emission from agricultural fields and soils under natural vegetation: summarizing available measurement data and modeling of global annual emissions, Nutrient Cycling in Agroecosystems 2006 74:3. 74 (2006) 207–228. https://doi.org/10.1007/S10705-006-9000-7.
[21] M.L. Cayuela, E. Aguilera, A. Sanz-Cobena, D.C. Adams, D. Abalos, L. Barton, R. Ryals, W.L. Silver, M.A. Alfaro, V.A. Pappa, P. Smith, J. Garnier, G. Billen, L. Bouwman, A. Bondeau, L. Lassaletta, Direct nitrous oxide emissions in Mediterranean climate cropping systems: Emission factors based on a meta-analysis of available measurement data, Agric. Ecosyst. Environ. 238 (2017) 25–35, doi:10.1016/J.AGEE.2016.10.006.
[22] K. Hergoualc'h, H. Akiyama, M. Bernoux, N. Chirinda, N2O Emissions from Managed Soils, and CO2 Emissions from Lime and Urea Application, in: 2019 Refinement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories, Switzerland, 2019. https://cgspace.cgiar.org/bitstream/handle/10568/107142/N2O_CO2.pdf?sequence=1 (accessed April 28, 2022).
[23] MAPA, Ministerio de Agricultura, Pesca y Alimentación de España, Productos FertilizantesBalance de Nutrientes. Balance de Fósforo En La Agricultura Española, Serie Histórica 1990-2016, 2018 https://www.mapa.gob.es/es/agricultura/temas/medios-de-produccion/productos-fertilizantes/ (accessed May 14, 2022).
[24] MAPA, Ministerio de Agricultura, Pesca y Alimentación de España, Productos FertilizantesBalance de Nutrientes. Balance de Nitrógeno En La Agricultura Española, Serie Histórica 1990-2016, 2018 https://www.mapa.gob.es/es/agricultura/temas/medios-de-produccion/productos-fertilizantes/ (accessed May 14, 2022).
[25] MAPA, Ministerio de Agricultura, Pesca y Alimentación de España, Balance de Fosforo En La Agricultura Española, Serie Histórica 1990-2016. (2018). https://www.mapa.gob.es/es/agricultura/temas/medios-de-produccion/productos-fertilizantes/(accessed May 18, 2022).
[26] MAPA, Ministerio de Agricultura, Pesca y Alimentación de España, Balance de Nitrógeno En La Agricultura Española, Serie Histórica 1990-2016. (2018). https://www.mapa.gob.es/es/agricultura/temas/medios-de-produccion/productos-fertilizantes/(accessed May 18, 2022).
[27] C. Melero, C. Gentil, N. Peña, C. Renaud-Gentié, P. Fantke, PestLCI Consensus, (2020). https://pestlciweb.man.dtu.dk/(accessed May 13, 2022).
[28] Ministerio de Agricultura Pesca y Alimentación, Encuesta sobre Superficies y Rendimientos Cultivos (ESYRCE): resultados de años anteriores, Boletines Monográficos. (2022). https://www.mapa.gob.es/es/estadistica/temas/estadisticas-agrarias/agricultura/esyrce/resultados-de-anos-anteriores/(accessed April 28, 2022).
[29] MITECO, Ministerio para la Transición Ecológica y el Reto Demográfico de España, Histórico de PreciosSeries Históricas de Precios Consultadas, 2022 https://sedeaplicaciones.minetur.gob.es/shpcarburantes/ (accessed May 14, 2022).
[30] INE, Instituto Nacional de Estadísticas, Spain. Encuesta sobre el uso del agua en el sector agrario. Serie 2000-2018, Disponibilidad y Origen Del Agua Por Comunidad Autónoma y Periodo. (2022). https://www.ine.es/jaxi/Tabla.htm?path=/t26/p067/p03/serie/l0/&file=02001.px&L=0 (accessed May 13, 2022).
[31] INE, Instituto Nacional de Estadísticas, Spain. Encuesta sobre el uso del agua en el sector agrario. Serie 2000-2018, Distribución de Agua a Las Explotaciones Agrícolas Por Comunidad Autónoma, Tipos de Cultivos/Técnicas de Riego y Periodo. (2022). https://www.ine.es/jaxi/Tabla.htm?path=/t26/p067/p03/serie/l0/&file=02003.px&L=0 (accessed May 13, 2022).
[32] Eurostat, Eurostat, NUTS - Nomenclature of Territorial Units for Statistics. (2021). https://ec.europa.eu/eurostat/web/nuts/background (accessed May 18, 2022).
[33] MAPA, Ministerio de Agricultura, Pesca y Alimentación de España, Calendarios Siembra, Recolección y Comercialización 2014-2016. (2021). https://www.mapa.gob.es/es/estadistica/temas/estadisticas-agrarias/agricultura/calendarios-siembras-recoleccion/(accessed May 14, 2022).
[34] R. Allen, L. Pereira, D. Raes, M. Smith, Crop evapotranspiration-Guidelines for computing crop water requirements, FAO, Rome, 1998 https://www.scscourt.org/complexcivil/105cv049053/volume3/172618e_5xagwax8.pdf (accessed May 13, 2022).
[35] A. Jones, O. Fernandez Ugalde, S. Scarpa, LUCAS 2015 Topsoil Survey, (2020) 1–75. https://doi.org/10.2760/616084.
[36] ESDAC, European Soil Data Centre (ESDAC), esdac.jrc.ec.europa.eu, European Commission, Joint Research Centre, LUCAS2015_SOIL Data. (2020). https://esdac.jrc.ec.europa.eu/content/lucas2015-topsoil-data (accessed May 13, 2022).
[37] USDA, Agricultural Research Service. U. S. Department of Agriculture: Soil Water Characteristics, (2022). https://www.ars.usda.gov/research/software/download/?softwareid=492 (accessed May 15, 2022).
[38] SIAR, Ministerio de Agricultura, Pesca y Alimentación. Subdirección General de Regadíos e Infraestructuras Rurales. Sistema de Información Agroclimática para el Regadío (SIAR), Necesidades Netas. (2022). https://eportal.mapa.gob.es/websiar/NecesidadesHidricas.aspx (accessed May 14, 2022).
[39] R Core Team, R: a language and environment for statistical computing, (2021). https://www.R-project.org/.
[40] RStudio Team, RStudio: Integrated Development Environment for R, (2022). http://www.rstudio.com/.
[41] H. Wickham, M. Averick, J. Bryan, W. Chang, L.D. McGowan, R. François, G. Grolemund, A. Hayes, L. Henry, J. Hester, M. Kuhn, T.L. Pedersen, E. Miller, S.M. Bache, K. Müller, J. Ooms, D. Robinson, D.P. Seidel, V. Spinu, K. Takahashi, D. Vaughan, C. Wilke, K. Woo, H. Yutani, Welcome to the tidyverse, J. Open Source Softw. 4 (2019) 1686, doi:10.21105/joss.01686.
[42] P. Schauberger, A. Walker, Openxlsx: read, write and edit xlsx files, (2021). https://CRAN.R-project.org/package=openxlsx.

[43] A. Henningsen, Linprog: linear programming /optimization, (2012). https://CRAN.R-project.org/package=linprog.
[44] R. Carnell, Triangle: provides the standard distribution functions for the triangle distribution, (2019). https://CRAN.R-project.org/package=triangle.
[45] H. Wickham, feather: R Bindings to the Feather "API," (2019). https://CRAN.R-project.org/package=feather.