



OPEN

# An experimental dataset on yields of pulses across Europe

DATA DESCRIPTOR

Daniele Antichi<sup>1,2</sup>✉, Silvia Pampana<sup>1,2</sup>, Lorenzo Gabriele Tramacere<sup>1,2</sup>, Véronique Biarnes<sup>3</sup>, Ina Stute<sup>4</sup>, Žydrė Kadžiulienė<sup>5</sup>, Becky Howard<sup>6</sup>, Isabel Duarte<sup>7</sup>, Oskars Balodis<sup>8</sup>, Iris Bertin<sup>9</sup>, David Makowski<sup>10</sup> & Nicolas Guilpart<sup>9</sup>

Future European agriculture should achieve high productivity while limiting its impact on the environment. Legume-supported crop rotations could contribute to these goals, as they request less nitrogen (N) fertilizer inputs, show high resource use efficiency and support biodiversity. However, legumes grown for their grain (pulses) are not widely cultivated in Europe. To further expand their cultivation, it remains crucial to better understand how different cropping and environmental features affect pulses production in Europe. To address this gap, we collected the grain yields of the most cultivated legumes across European countries, from both published scientific papers and unpublished experiments of the European projects LegValue and Legato. Data were integrated into an open-source, easily updatable dataset, including 5229 yield observations for five major pulses: chickpea (*Cicer arietinum* L.), faba bean (*Vicia faba* L.), field pea (*Pisum sativum* L.), lentil (*Lens culinaris* Medik.), and soybean (*Glycine max* (L.) Merr.). These data were collected in 177 field experiments across 21 countries, from 37° N (southern Italy) to 63° N (Finland) of latitude, and from ca. 8° W (western Spain) to 47° E (Turkey), between 1980 and 2020. Our dataset can be used to quantify the effects of the soil, climate, and agronomic factors affecting pulses yields in Europe and could contribute to identifying the most suitable cropping areas in Europe to grow pulses.

## Background & Summary

Grain legumes (also named pulses) are Fabaceae crops sown and harvested for dry grain production and used as feed or food. Nowadays, they are widely recognized as key components of sustainable cropping systems, because, besides their use as food and feed, they may additionally deliver several supporting and regulating ecosystem services<sup>1,2</sup>.

Legumes have the unique ability to establish a symbiosis with rhizobia bacteria to fix atmospheric N<sub>2</sub>, thus providing another source of nitrogen (N) to the plant in addition to that available in soils. Moreover, the mineralization of their N-rich crop residues enlarges soil N availability for the following crops<sup>3,4</sup>. Together, these processes can reduce mineral N fertilizer requirements in cropping systems, and consequently can lower fossil energy use and direct/indirect net greenhouse gas (GHG) emissions in agriculture, associated with the manufacture and field application of mineral N fertilizers<sup>4</sup>. Moreover, increased cultivation of grain legumes would contribute to the diversification of cereal-dominated crop rotations in Europe, with expected positive effects on biodiversity, control of weeds, pests and diseases, and soil structure<sup>5,6</sup>, as well as on crop yields<sup>7</sup>.

Legumes can also have negative outcomes. For example, higher soil N losses to the environment have been observed due to the higher soil N content after legume harvest<sup>8,9</sup>. In comparison to cereals, some legume crops may also require a higher pesticide use intensity (e.g., pea) or more irrigation water (e.g., soybean)<sup>10,11</sup>. However,

<sup>1</sup>Department of Agriculture, Food and Environment, University of Pisa, Via del Borghetto 80, Pisa, 56124, Italy.

<sup>2</sup>Centre for Agri-environmental Research "Enrico Avanzi", University of Pisa, Via Vecchia di Marina 2, San Piero a Grado, 56122, Italy. <sup>3</sup>Terres Inovia, Avenue Lucien Bretignières, Campus de Grignon, Thiverval-Grignon, 78850, France. <sup>4</sup>Fachhochschule Südwestfalen, Lübecker Ring 2, Soest, 59494, Germany. <sup>5</sup>Lithuanian Research Centre for Agriculture and Forestry, Instituto al. 1, Akademija, Kėdainiai, LT-58344, Lithuania. <sup>6</sup>PGRO Research Limited, The Research Station, Great North Road, Thornhaugh, Peterborough, PE8 6HJ, UK. <sup>7</sup>Instituto Nacional de Investigação Agrária e Veterinária, Estrada de Gil Vaz, Apartado 6, 7351-901, Elvas, Portugal. <sup>8</sup>Faculty of Agriculture, Latvia University of Agriculture, Lielā iela 2, Jelgava, LV-3001, Latvia. <sup>9</sup>Université Paris-Saclay, AgroParisTech, INRAE, UMR Agronomie, 91120, Palaiseau, France. <sup>10</sup>University Paris-Saclay, AgroParisTech, INRAE, UMR MIA Paris-Saclay, 91120, Palaiseau, France. ✉e-mail: [daniele.antichi@unipi.it](mailto:daniele.antichi@unipi.it)

despite these potential drawbacks, the overall benefits of increased legume cultivation in the European Union (EU) are still expected to be positive<sup>9,12</sup>. Nevertheless, the presence of legumes in cropping systems of the EU is still scarce (nowadays, about 2% of arable land) so that the EU is highly dependent on imports of feed-legumes and agricultural systems are still heavily reliant on fossil energy used for the manufacture of synthetic nitrogen fertilizers<sup>8,13</sup>.

Several subsidies have been introduced under the EU's Common Agricultural Policy (CAP) and Rural Development Plans (RDPs) to increase the presence of grain legumes in European farming systems, to support both the transition to more sustainable food production<sup>9</sup> and the goals of the European Green Deal "Farm to Fork" strategy<sup>14</sup>. Nevertheless, many socio-economic and agronomic factors have been shown to explain the marginalization of grain legumes in Europe<sup>15,16</sup>. Importantly, annual economic margins of pulses remain too low in comparison to other crops like cereals, due to inferior yields, higher yield instability<sup>17</sup>, and lower market prices<sup>18</sup>. Moreover, investments in research and development (R&D) have been lower for legumes than for other major crops, and legume cultivation is facing many genetic and agronomic challenges that need to be addressed to improve the yield and yield stability of pulses. These challenges include breeding for improved varieties<sup>19</sup>, as well as abiotic stress management<sup>20,21</sup> and the control of pests, diseases, and weeds, which is particularly challenging if synthetic agrochemicals cannot be used like in organic farming systems<sup>22</sup>.

To increase the growth of grain legumes in Europe, an essential first step is to identify the most suitable areas for their cultivation, i.e., regions where high and stable yields can be attained. Indeed, a better knowledge of the most suitable areas for pulses is important to identify which legume species are best adapted to local conditions of climate and soils<sup>23</sup>, identify the most important limiting factors and improve agronomic management of these crops across a wide range of agroclimatic zones in Europe<sup>1,24</sup>, and give guidance for the development of value chains. To date, European regions potentially relevant to grow pulses have been only identified for soybean<sup>25</sup>, but not for other grain legumes. Moreover, many different cultivars were produced for a given species and they could respond differently to different environments and management practices (Genotype x Environment x Management interactions<sup>26</sup>). It is thus important to build large datasets including yield data covering a wide range of cultivars, environments, and management practices to robustly identify suitable areas<sup>23,25,27</sup>.

Some existing datasets provide yield data for grain legumes in Europe, but none of them meets all the characteristics required for our purpose. For example, annual yield data of grain legumes are provided at the country level by the Food and Agriculture Organization (FAO) of the United Nations (see <https://www.fao.org/faostat/en/#data>), but the spatial resolution is too coarse to capture local effects of variations in climate and soils on crop yield. The Global Dataset of Historical Yields for major crops provides annual yield data at a spatial resolution of 0.5° (grid cells of 55 km), but only for soybean<sup>28,29</sup>. The Spatial Production Allocation Model (SPAM) data provide yield data for 42 crops including grain legumes at the global scale at a spatial resolution of 0.083°, with 10-km grid cells, but only for the year 2010<sup>30</sup>. Moreover, available datasets mostly provide yield data where crops are grown, while observed actual yields in areas where grain legumes are not yet grown by farmers would be of interest as well, to identify areas where grain legumes could be potentially grown in the future. A previous dataset of legume yields from field experiments was made available by Cernay and co-authors<sup>27</sup>. It covers the period 1967–2016 and includes records across 41 countries and 18 Köppen-Geiger climatic zones. But this dataset is focused on field experiments comparing several pulse species and does not include yield data collected in experiments including single pulse species. This dataset thus missed numerous experimental pulse yield data obtained in Europe.

To fill this gap, we present a new dataset on grain legume yields, gathering results from field experiments for the five dominant species in Europe: chickpea (*Cicer arietinum* L.), faba bean (*Vicia faba* L.), field pea (*Pisum sativum* L.), lentil (*Lens culinaris* Medik.), and soybean (*Glycine max* (L.) Merr.). This dataset, named the "European Grain Legume Dataset" (EGLD) includes 5229 yield observations from 177 field experiments across 21 countries, from 1980 to 2020. EGLD can prove to be useful for disentangling the effects of soil, climate and agronomic drivers of legume yields in Europe.

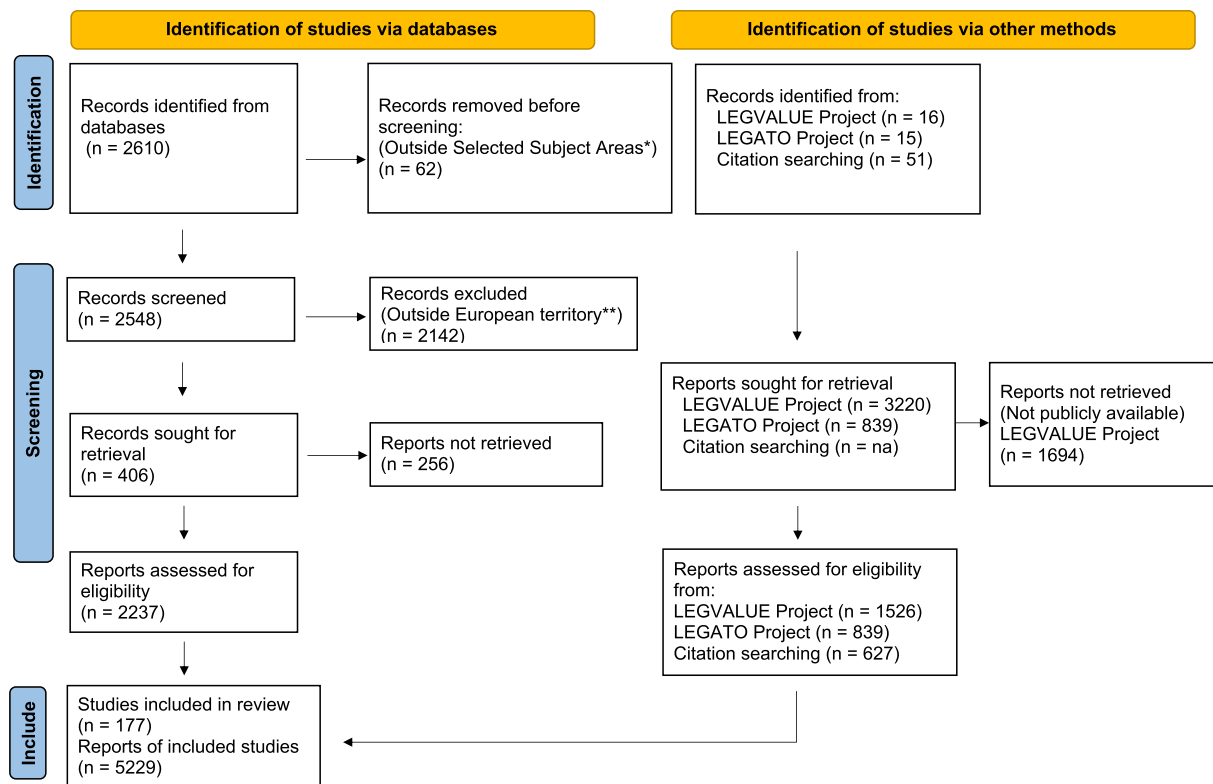
## Methods

Data were collected from three different sources (Fig. 1): (i) published experimental data from an online literature search; (ii) experimental data from trials conducted within the EU-FP7 LEGATO (LEGumes for the Agriculture of TOMorrow) Project (GA nr. 613551); (iii) experimental data from trials conducted within the EU-H2020 LEGVALUE (Fostering sustainable legume-based farming systems and agri-feed and food chains in the EU) Project (GA nr. 727672). Five pulses, namely chickpea, faba bean, field pea, lentil, and soybean have been searched, as the most cultivated in Europe.

**Data collection from scientific papers.** Concerning the first source of information, we used the Scopus collection to retrieve the foremost publications reporting peer-reviewed research on the selected pulse species. The systematic search of peer-reviewed journals was completed in June 2020. The search criteria included the following elements, searched for in the following order within article title, abstract, and keywords:

- Crop\*
- AND one of the following terms: chickpea, faba bean, lentil, pea, soybean
- AND (yield OR 'dry matter' OR biomass)
- AND (compar\* OR assessment OR product\* OR performance\*)
- AND (trial\* OR factorial OR experiment\* OR treatment\* OR condition\*).

At first, we did not set any restrictions on publication date and language and retrieved 2610 papers. The search results were refined to "Agricultural and Biological Sciences" and "Environmental Sciences" as subject



**Fig. 1** PRISMA<sup>186</sup> flow diagram for the systematic review process, including searches of database and of unpublished experiments of LEGVALUE and LEGATO Projects. \*EU-27 countries plus Turkey, Albania, Switzerland, United Kingdom, Norway, Ukraine, Serbia, Montenegro, North Macedonia, Kosovo, Moldova, Iceland, Belarus) \*\* “Agricultural and Biological Sciences” and “Environmental Sciences”.

areas in Scopus. After that, we further restricted the records to affiliation countries in Europe (EU-27 countries plus Turkey, Albania, Switzerland, United Kingdom, Norway, Ukraine, Serbia, Montenegro, North Macedonia, Kosovo, Moldova, Iceland, and Belarus).

At that point, the literature search identified 406 articles of potential interest (Fig. 1).

We also gathered 51 additional papers from the literature review by Ditzler *et al.*<sup>1</sup> and from the references cited in the eligible papers identified above, that had not been identified by the initial search, according to our inclusion criteria.

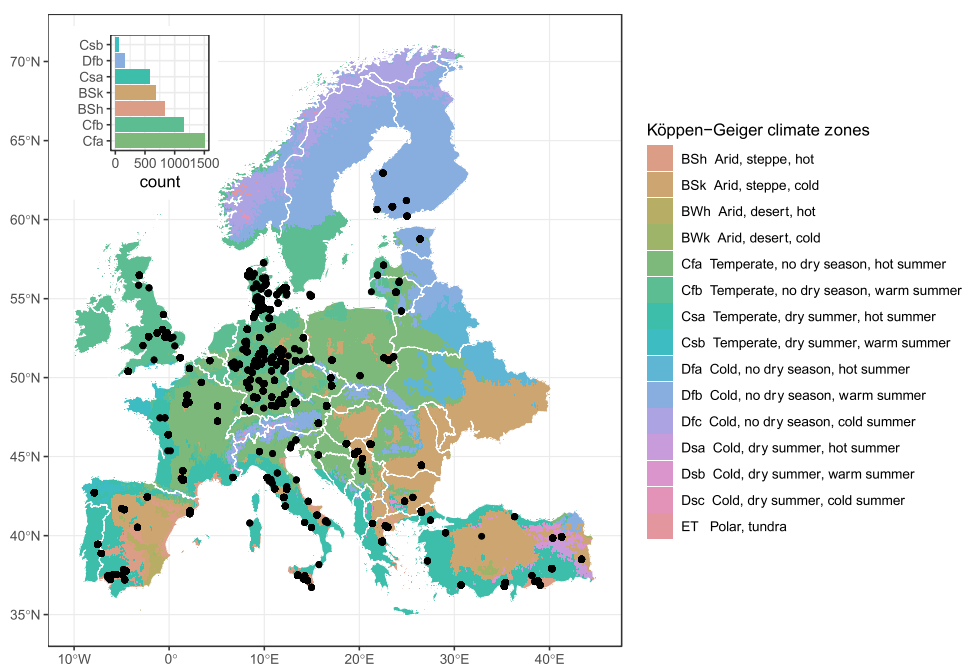
Article titles and abstracts have then been screened for eligibility according to the following criteria: (i) article title and/or article abstract reporting one or more grain legume species grown as sole crop; (ii) article title and/or article abstract reporting at least one experiment at one of the countries listed above; (iii) article title and article abstract published in a peer-reviewed journal. All the eligible full-text articles were thoroughly read at least twice and by two different authors to assess their admissibility according to additional criteria: (i) at least one grain legume grown as sole crop per each experimental site included in the field experiment; (ii) data coming from experimental trials; (iii) study reporting grain yields (i.e., papers reporting only total plant biomass were excluded); (iv) precise information on field site (region and site name and/or latitude and longitude coordinates); (v) main agronomic practices reported. At least the year was considered a mandatory information for sowing and harvest to relate the agronomic data with weather conditions. If the month of sowing/harvest, but not the day, was reported, then the central day of the month was used as a proxy of the actual sowing/harvest dates. Then, if grain yield data were expressed as fresh matter without indication of percent moisture, they were excluded; otherwise, fresh matter data were converted into dry matter and all values were reported as dry weights per unit area.

We finally picked out 146 eligible full-text articles published between 1990 and 2020<sup>31–176</sup>. Data were manually extracted from the tables of the selected papers and the WebPlotDigitizer-4.2.0 app (<https://automeris.io/WebPlotDigitizer>) was used to extract data from figures. All data were converted into a digital format and inserted into a csv file.

**Data collection from experimental trials.** To expand the dataset, we sought out further yield data that met the following criteria: i) data coming from experimental trials; ii) at least one of the selected grain legume species grown as sole crop; iii) experimental site of the field experiment at one of the countries listed above; iv) study reporting grain yields; v) disposal of precise information on field site (region and site name and/or latitude and longitude coordinates); vi) main agronomic practices available and reported. Accordingly, we included also open data provided by the former EU-FP7 project LEGATO “LEGumes for the Agriculture of TOMorrow”

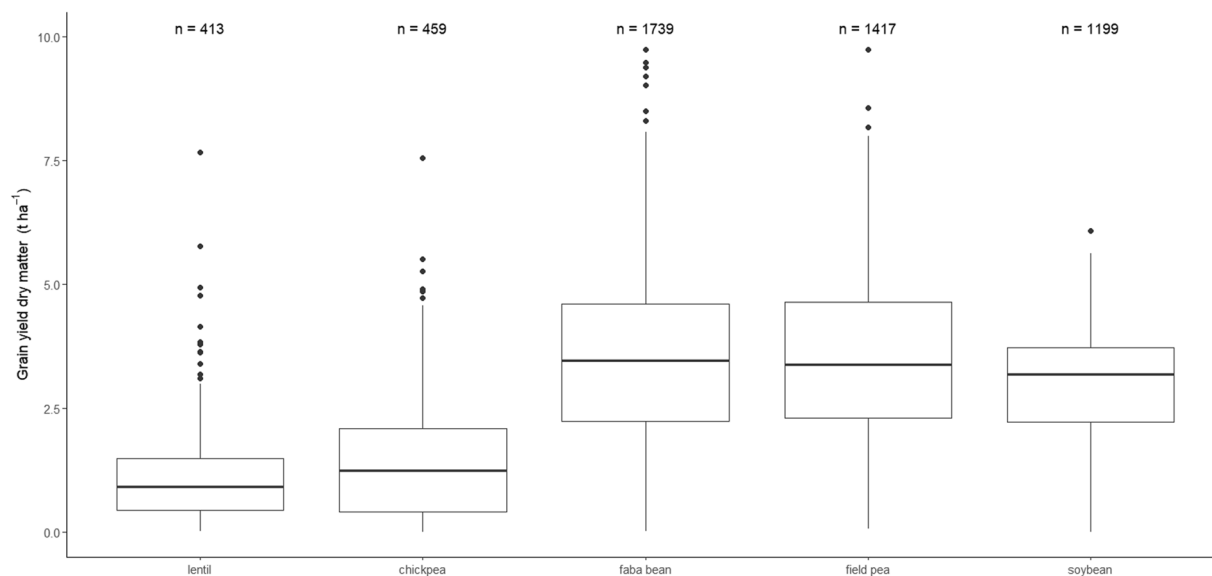
| Country             | Chickpea   |        | Faba bean   |        | Field pea   |        | Lentil     |        | Soybean     |        | Total (country)<br>n |
|---------------------|------------|--------|-------------|--------|-------------|--------|------------|--------|-------------|--------|----------------------|
|                     | n          | (%)    | n           | (%)    | n           | (%)    | n          | (%)    | n           | (%)    |                      |
| Austria             | 2          | (0.4)  | 48          | (2.8)  | 52          | (3.7)  |            |        | 6           | (0.5)  | 108                  |
| Belgium             |            |        |             |        |             |        |            |        | 14          | (1.2)  | 14                   |
| Bulgaria            | 4          | (0.9)  |             |        | 24          | (1.7)  |            |        |             |        | 28                   |
| Croatia             |            |        |             |        |             |        |            |        | 24          | (2.0)  | 24                   |
| Czech Republic      |            |        | 38          | (2.2)  | 66          | (4.7)  |            |        |             |        | 104                  |
| Denmark             | 4          | (0.9)  | 75          | (4.3)  | 113         | (8.0)  | 4          | (1.0)  | 4           | (0.3)  | 200                  |
| Estonia             |            |        | 38          | (2.2)  | 38          | (2.7)  |            |        |             |        | 76                   |
| Finland             |            |        | 38          | (2.2)  | 51          | (3.6)  |            |        |             |        | 89                   |
| France              |            |        | 212         | (12.2) | 150         | (10.6) |            |        | 44          | (3.7)  | 406                  |
| Germany             |            |        | 661         | (38.0) | 331         | (23.4) | 58         | (14.0) | 486         | (40.5) | 1536                 |
| Greece              | 168        | (36.3) | 20          | (1.1)  | 36          | (2.5)  | 137        | (33.2) |             |        | 361                  |
| Italy               | 85         | (18.5) | 186         | (10.7) | 63          | (4.4)  | 25         | (6.1)  | 88          | (7.3)  | 447                  |
| Latvia              |            |        | 5           | (0.3)  | 1           | (0.1)  |            |        | 5           | (0.4)  | 11                   |
| Lithuania           |            |        | 5           | (0.3)  | 31          | (2.2)  |            |        | 24          | (2.0)  | 60                   |
| Poland              |            |        | 7           | (0.4)  | 39          | (2.8)  |            |        | 57          | (4.8)  | 103                  |
| Portugal            | 51         | (11.1) | 20          | (1.1)  | 26          | (1.8)  |            |        |             |        | 97                   |
| Romania             |            |        |             |        | 12          | (0.8)  | 24         | (5.8)  | 82          | (6.8)  | 118                  |
| Serbia              |            |        | 12          | (0.7)  | 34          | (2.4)  |            |        | 147         | (12.3) | 193                  |
| Spain               | 112        | (24.4) | 161         | (9.2)  | 90          | (6.4)  | 49         | (11.9) |             |        | 412                  |
| Turkey              | 33         | (7.2)  |             |        | 70          | (4.9)  | 116        | (28.1) | 218         | (18.2) | 437                  |
| United Kingdom      |            |        | 215         | (12.3) | 190         | (13.4) |            |        |             |        | 405                  |
| <b>Total (crop)</b> | <b>459</b> | —      | <b>1741</b> | —      | <b>1417</b> | —      | <b>413</b> | —      | <b>1199</b> | —      | <b>5229</b>          |

**Table 1.** Distribution of dataset entries among European Union countries for the five pulse species included in EGLD. n = number of entries and (%) = percent of entries by country.



**Fig. 2** Geographic distribution of the experimental sites reporting the yields of the five pulse species. The insert provides the number of data in the seven most represented Koppen Geiger climate zones<sup>186,187</sup>.

(2014–2017), which has delivered the results of the 2-yr experimental activities performed on field pea, faba bean, chickpea, and lentil. The methodological aspects and the experimental details of the trials are openly available at [https://intranet.iamz.ciheam.org/forms/Legato/WP6/files/Field\\_Trial\\_Protokol\\_5.1.2016.pdf](https://intranet.iamz.ciheam.org/forms/Legato/WP6/files/Field_Trial_Protokol_5.1.2016.pdf). The dataset was downloaded from <https://intranet.iamz.ciheam.org/forms/Legato/WP6/index.php> and eligible data, according to the criteria, were extracted and manually added to the dataset. Overall, data were produced in field experiments



**Fig. 3** Box plots showing the distributions of grain yield data for the five pulse species. Values are reported as dry weights per unit area ( $\text{t ha}^{-1}$ ). Main body of the boxplot shows the interquartile range ( $\text{IQR} = \text{Q3} - \text{Q1}$ ), and the central line the median ( $\text{Q2}$ ). Whiskers (bars) represent  $\text{Q1} - 1.15 \text{ IQR}$  (lower) and  $\text{Q3} + 1.15 \text{ IQR}$  (upper), and dots indicate outliers.

| Source              | Soil classification | Soil texture | Cultivar   | Tillage     | Plant density | Sowing density | Row spacing |
|---------------------|---------------------|--------------|------------|-------------|---------------|----------------|-------------|
| LEGATO Experiment   | 801                 | 357          | 0          | 581         | 139           | 173            | 839         |
| LEGVALUE Experiment | 878                 | 202          | 100        | 1162        | 1316          | 460            | 1038        |
| Paper               | 1563                | 1161         | 264        | 961         | 2432          | 974            | 839         |
| <i>Total</i>        | <i>3242</i>         | <i>1720</i>  | <i>364</i> | <i>2704</i> | <i>3887</i>   | <i>1607</i>    | <i>2716</i> |

**Table 2.** Number of entries with missing data about soil classification, texture, cultivar, plant and sowing density, and row spacing (per source of information and in total).

in compliance with a common protocol. The field experiments were conducted on plots of  $10 \text{ m}^2$  and replicated four times in space, according to a completely randomized design (CRD). Plant samples for grain yield assessment were collected on  $1 \text{ m}^2$  sampling areas when the crops reached harvest maturity.

The dataset was also complemented with the European H2020 LEGVALUE Task 1.2 Partners ([www.legvalue.eu](http://www.legvalue.eu)) own data, which were generated both on-station and on-farm but always under well-determined field conditions. The authors, as effective partners of the Project could deliver their not already published results. Data that met the eligibility criteria were added by each partner to the original dataset, following the related instructions (see Data records section). The trials were conducted according to site-specific experimental protocols, that were documented by the providing partners and are available in the dataset (Supplementary Table S1).

Those details on geographical site, pedoclimatic conditions, treatment replications, and agronomic management (i.e., cultivar choice, tillage, plant density, fertilization, crop protection, growing cycle length) have been included in the dataset for each entry coming from both LEGATO and LEGVALUE experiments.

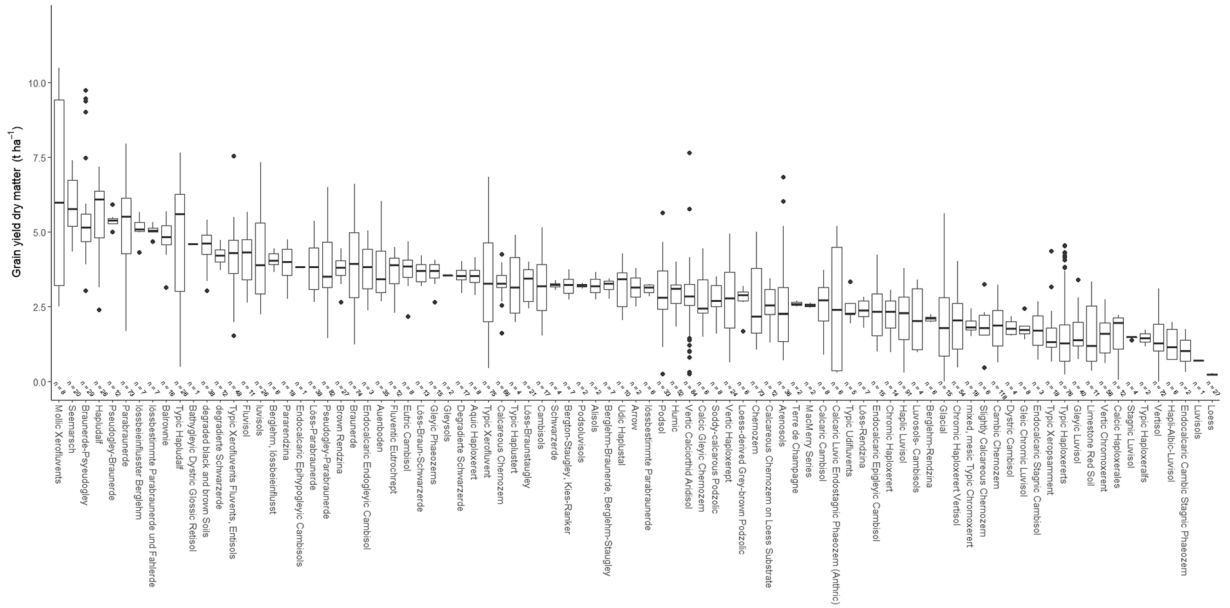
Additional information on the experimental methodology of the trials of both LEGATO and LEGVALUE are summarized in Supplementary Table S2.

### Data Records

The dataset created with all the extracted data, named “European Grain Legumes Dataset” (EGLD) is available at figshare<sup>177</sup>.

**Overview of data files.** In the figshare repository, the following files are provided<sup>177</sup>: (i) “European Grain Legume Dataset.csv” that contains the data; (ii) “EGLD instructions.pdf” and “EGLD instructions.csv” that report the list and the definition of each column in the dataset (meta-data); (iii) “List of sources.pdf” and “List of sources.csv” that provide the list of the sources where the data were collected from.

In “EGLD instructions” files we reported a brief description and the assumptions for each column (category) for the reader’s guidance, to facilitate data entry operations as well as to inform the proper interpretation of the data.



**Fig. 4** Box plots of grain yields dry matter ( $t\ ha^{-1}$ ), ranked by soil type. Main body of the boxplot shows the interquartile range (IQR), and the central line the median. Whiskers (bars) represent  $Q1-1.15\ IQR$  (lower) and  $Q3+1.5\ IQR$  (upper), with dots to indicate outliers.  $n =$  Number of entries. Values are pooled across the five species.

The files “List of sources” complement information about the founts from which the data has been acquired, together with the web link at which they are accessible and the responsible institution.

**Overview of the structure of the dataset.** The “European Grain Legume Dataset.csv” file contains all the data for the five selected pulse species, with variables as columns and entries as rows. As several papers and experiments reported multiple yield assessments, a unique ID was assigned to a single combination of year x site x crop species x experimental treatment level and was reported as a single entry (i.e., a single row) in the dataset.

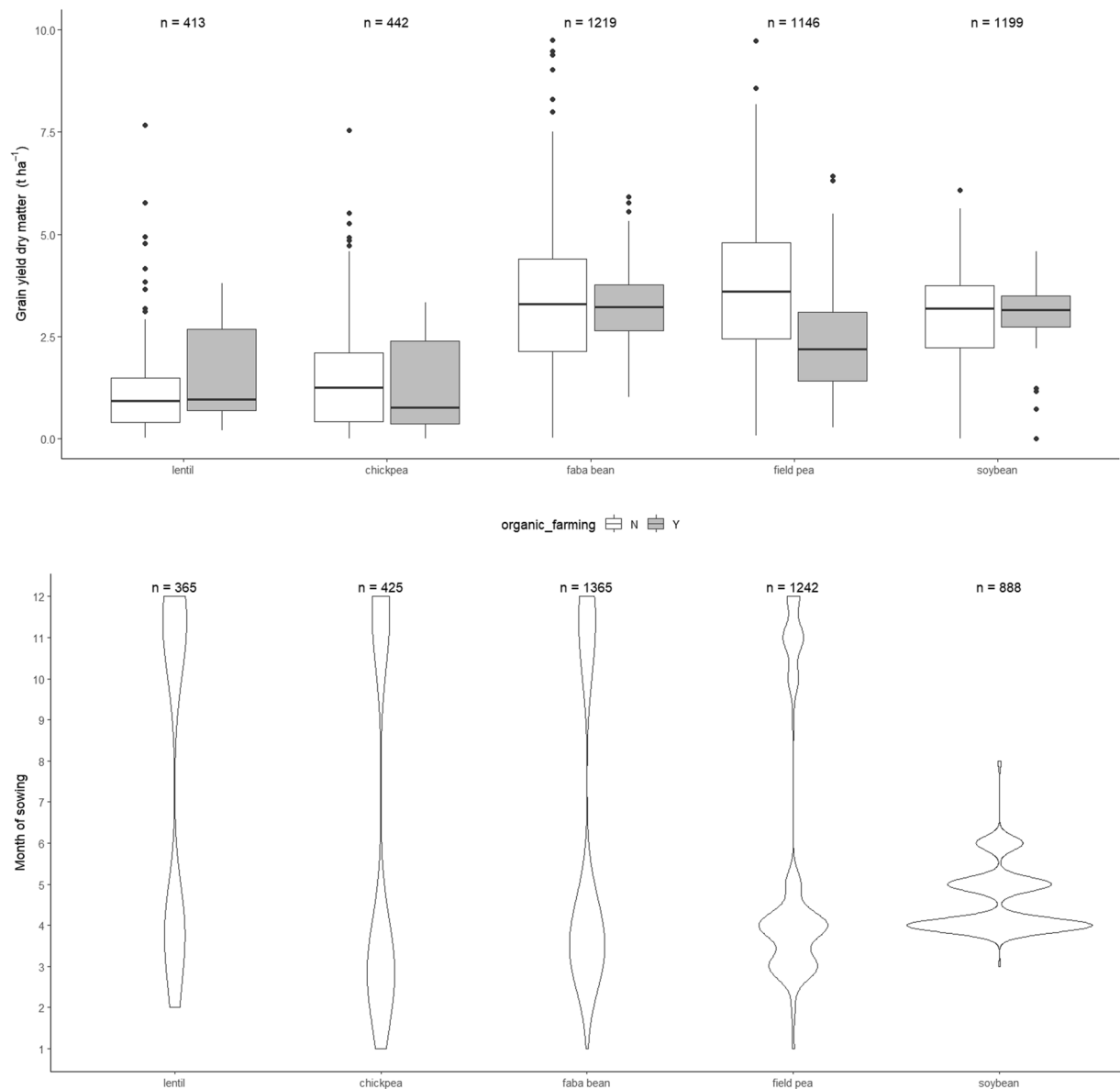
This process led to a total of 5229 yield data for the five selected pulses all over Europe (Table 1 and Fig. 2), of which 2864 were collected from published papers, 1526 from non-published experimentations of the LEGVALUE project, and 839 from the LEGATO project. Records identified from published papers or the LEGATO Project or other experiments of the LEGVALUE Project were differently labeled in the database in the column “source” as Paper, Experiment LEGATO or Experiment LEGVALUE, respectively. The coordinates of each yield data were extracted and included in “European Grain Legume Dataset.csv”.

All the sources from which data were recovered are reported in the file named “List of Sources.pdf”, available on the figshare repository<sup>177</sup>. For data obtained from the LEGATO project or from experiments of the LEGVALUE project, a brief description of the experiment is delineated, while for published papers the full reference is reported. We also provide a web address where the sources can be accessed (through the digital identification number (doi) if available). Most of the published papers are publicly available under Open Access licenses. A few can be accessed by personal or institutional subscription, depending on each journal policy.

All records enclose 72 columns that describe a corresponding number of variables with different types of information. These variables can be grouped into four categories: the first provides information on the source of the data, the second on the experiment, the third on agricultural management practices, and the fourth category conveys figures about crop yield. The full list of variables included in the dataset, together with their detailed description and unit has been reported in Supplementary Table 1.

**Overview of the data.** The EGLD<sup>177</sup> contains yield data from 21 countries, from ca. 37° N (southern Italy) to 63° N (Finland) of latitude, and from ca. 8° W (western Spain) to 47° E (Turkey) (Fig. 2), thus capturing a wide range of pedoclimatic conditions. With about one third (29%) of the data, Germany is the country with the highest number of observations, followed by Italy, Turkey, Spain, France, the UK, and Greece, each of these countries representing about 8% of the dataset (Table 1). Other countries have lower numbers of observations, with Bulgaria, Croatia, Belgium, and Latvia having less than 30 observations each.

Faba bean, pea, and soybean are well represented in the dataset as they account for 33%, 27%, and 23% of total number of observations, and are present in 19, 16, and 13 of the 21 countries, respectively. On the other hand, chickpea and lentil are much less represented as they account for only 9% and 8% of the whole dataset, and are present in 8 and 7 countries, respectively. Belgium and Croatia provide information only for one pulse (soybean), while Italy and Denmark have information about all the five species. The dataset captures a wide range of grain yield values, from complete crop failure (yield almost null) to very high yield (for all the species, the maximum yield was higher than  $6\ t\ ha^{-1}$  of dry matter) (Fig. 3). The most represented agronomic practice that has



**Fig. 5** Two examples of management practices included in the dataset: organic cultivation and month of sowing. **(a)** Box plot of the effect of organic cultivation (grey) vs conventional cultivation (white) on the yield of the five pulse species. Main body of the boxplots shows the interquartile range (IQR), and the central line the median. Whiskers (bars) represent  $Q1 - 1.15 \text{ IQR}$  (lower) and  $Q3 + 1.5 \text{ IQR}$  (upper), with dots to indicate outliers. **(b)** Violin plot of the month of sowing for the five pulse species. The width of the areas represents the proportion of data located.  $n$  = Number of entries.

been evaluated in the experimental trials was found to be the genotype (cultivar) as it represents more than 50% of all the entries of the dataset. Tillage and weed control are the next, with 8 and 6% of total entries, respectively.

Missing data are indicated as “NA” (not available) cells in the dataset. For example, cultivar was missing in 7% of entries, soil texture in 33% of the total, while the soil classification was not reported for more than half entries (about 62%), and tillage was missing in half (52%) of entries (Table 2).

### Technical Validation

After the data extraction, the quality check was further carried out by comparing the entire set of collected data against the corresponding original source (papers or experiments) to make sure the data were digitalized correctly. The formats of each column (numerical or string) were checked to correct misprinting. We visualized the data distribution for each numerical column and detected outliers, that were manually checked and validated by comparing them with the values reported in the original papers or experiments. Moreover, the values of crop yield reported in the database are consistent with the results of published research<sup>1,17,23,178–180</sup>. Graphical exploration of the data was performed by using the package ggplot2 with the ggplot function of the statistical software R, version 4.1.2.

## Usage Notes

The yield dataset can be used for several purposes. First it is useful to improve our knowledge of the relationships between pedoclimatic factors, agronomic practices, and legume yield. The wide range of soil types (Fig. 4), climate zones (Fig. 1), and management practices (Fig. 5) covered by our dataset provides a unique opportunity to identify the main factors influencing yield of pulses<sup>8,181,182</sup>. A better understanding of GxExM interactions impacts of the yield for grain legumes would contribute to improving our understanding of the determinants of productive performances of the cultivars tested in Europe and identify the most promising in different regions<sup>15,16</sup>.

Second, the dataset can be used to map achievable yields of grain legumes over Europe under current and future climate scenarios. This can be done by fitting statistical or machine learning models to predict grain yield from climate (and possibly other factors like soil and management practices) inputs, as done by Guilpart *et al.* for soybean<sup>25</sup>. Thanks to the experiment geographical coordinates, it is possible to associate climate and soil information with the yield data included in our dataset<sup>183–185</sup>. It would be then possible to train data-driven yield forecasting models and produce yield maps for the five pulse species under various hypothetical climate and management scenarios, as done for soybean<sup>25</sup>. Such yield maps may be highly relevant to: (i) identify suitable areas for pulses cultivation in Europe under both current and future climate; (ii) simulate the impact of an increase of pulses growing area on the pulse production in Europe; (iii) support the definition of market scenarios as well as policies on protein/starch production in the EU, highlighting potentialities, environmental barriers, and constraints; (iv) identify pulses species and cultivars best adapted to local pedoclimatic conditions over Europe.

Notably, the EGLD dataset could be easily updated using data retrieved from recently published papers and data recorded in new experiments. However, at the figshare accession data is that peer reviewed in 2023<sup>177</sup> and this version will be maintained.

## Code availability

Prisma Flow Diagram ([http://www.prisma-statement.org/documents/PRISMA\\_2020\\_flow\\_diagram\\_new\\_SRs\\_v2.docx](http://www.prisma-statement.org/documents/PRISMA_2020_flow_diagram_new_SRs_v2.docx)) was used to generate Fig. 1.

The statistical software R (version 4.1.2, 2021-11-01-“Bird Hippie” Copyright © 2021. The R Foundation for Statistical Computing) was used to produce Figs. 2–5. R scripts used to produce figures are openly available on the figshare repository.

Received: 25 April 2023; Accepted: 28 September 2023;

Published online: 17 October 2023

## References

- Ditzler, L. *et al.* Current research on the ecosystem service potential of legume inclusive cropping systems in Europe. A review. *Agron. Sustain. Dev.* **41**, 26 (2021).
- World Resources Institute. *Millennium Ecosystem Assessment Ecosystems and Human Well-being: General Synthesis*. Island Press, Washington DC (2005).
- Peoples, M. B. *et al.* The contributions of nitrogen-fixing crop legumes to the productivity of agricultural systems. *Symbiosis* **48**, 1–17 (2009).
- Jensen, E. S. *et al.* Legumes for mitigation of climate change and the provision of feedstock for biofuels and biorefineries. A review. *Agron. Sustain. Dev.* **32**, 329–364 (2012).
- Tamburini, G. *et al.* Agricultural diversification promotes multiple ecosystem services without compromising yield. *Sci. Adv.* **6** (2020).
- Beillouin, D., Ben-Ari, T., Malézieux, E., Seufert, V. & Makowski, D. Positive but variable effects of crop diversification on biodiversity and ecosystem services. *Glob. Chang. Biol.* **27**, 4697–4710 (2021).
- Zhao, J. *et al.* Global systematic review with meta-analysis reveals yield advantage of legume-based rotations and its drivers. *Nat. Commun.* **13**, 4926 (2022).
- Watson, C. A. *et al.* Grain Legume Production and Use in European Agricultural Systems. *Adv. Agron.* **144**, 235–303 (2017).
- Foyer, C. H. *et al.* Neglecting legumes has compromised human health and sustainable food production. *Nat. Plants* **2**, 16112 (2016).
- Urruty, N., Deveaud, T., Guyomard, H. & Boiffin, J. Impacts of agricultural land use changes on pesticide use in French agriculture. *Eur. J. Agron.* **80**, 113–123 (2016).
- Rüdelsheim, P. L. J. & Smets, G. Baseline information on agricultural practices in the EU Soybean (*Glycine max* (L.) Merr.). Available at: <http://www.europabio.org/baseline-information-agricultural-practices-eu-soybean-glycine-max-l-merr> (2012).
- Cusworth, G., Garnett, T. & Lorimer, J. Legume dreams: The contested futures of sustainable plant-based food systems in Europe. *Glob. Environ. Change* **69**, 102321 (2021).
- Harchaoui, S. & Chatzimpiros, P. Energy, nitrogen, and farm surplus transitions in agriculture from historical data modeling. *France, 1882–2013. J. Ind. Ecol.* **23**, 412–425 (2019).
- European Union. Farm to Fork Strategy. [https://food.ec.europa.eu/horizontal-topics/farm-fork-strategy\\_en](https://food.ec.europa.eu/horizontal-topics/farm-fork-strategy_en) (2020).
- Magrini, M.-B. *et al.* Why are grain-legumes rarely present in cropping systems despite their environmental and nutritional benefits? Analyzing lock-in in the French agrifood system. *Ecol. Econ.* **126**, 152–162 (2016).
- Zander, P. *et al.* Grain legume decline and potential recovery in European agriculture: a review. *Agron. Sustain. Dev.* **36**, 26 (2016).
- Cernay, C., Pelzer, E. & Makowski, D. A global experimental dataset for assessing grain legume production. *Sci. Data* **3**, 160084 (2016).
- de Visser, C. L. M., Schreuder, R. & Stoddard, F. The EU's dependency on soya bean import for the animal feed industry and potential for EU produced alternatives. *OCL* **21**, D407 (2014).
- Rubiales, D., Annicchiarico, P., Vaz Patto, M. C. & Julier, B. Legume breeding for the agroecological transition of global agri-food systems: a European perspective. *Front. Plant Sci.* **12** (2021).
- Zahran, H. H. Rhizobium–Legume symbiosis and nitrogen fixation under severe conditions and in an arid climate. *Microbiol. Mol. Biol. Rev.* **63**, 968–989 (1999).
- Pampana, S., Masoni, A. & Arduini, I. Response of cool-season grain legumes to waterlogging at flowering. *Can. J. Plant Sci.* **96** (2016).



22. Ghosh, P. K. *et al.* Grain legume inclusion in cereal–cereal rotation increased base crop productivity in the long run. *Exp. Agric.* **56**, 142–158 (2020).
23. Pelzer, E. *et al.* Design, assessment, and feasibility of legume-based cropping systems in three European regions. *Crop Pasture Sci.* **68**, 902 (2017).
24. Voisin, A. S. *et al.* Legumes for feed, food, biomaterials and bioenergy in Europe: a review. *Agron. Sustain. Dev.* **34**, 361–380 (2014).
25. Guilpart, N., Iizumi, T. & Makowski, D. Data-driven projections suggest large opportunities to improve Europe's soybean self-sufficiency under climate change. *Nat. Food.* **3**, 255–265 (2022).
26. Cooper, M., Voss-Fels, K. P., Messina, C. D., Tang, T. & Hammer, G. L. Tackling  $G \times E \times M$  interactions to close on-farm yield-gaps: creating novel pathways for crop improvement by predicting contributions of genetics and management to crop productivity. *Theor. Appl. Genet.* **134**, 1625–1644 (2021).
27. Cernay, C., Ben-Ari, T., Pelzer, E., Meynard, J.-M. & Makowski, D. Estimating variability in grain legume yields across Europe and the Americas. *Sci. Rep.* **5**, 11171 (2015).
28. Iizumi, T. & Sakai, T. The global dataset of historical yields for major crops 1981–2016. *Sci. Data* **7**, 97 (2020).
29. Iizumi, T. *et al.* Historical changes in global yields: major cereal and legume crops from 1982 to 2006. *Glob. Ecol. Biogeogr.* **23**, 346–357 (2014).
30. Yu, Q. *et al.* A cultivated planet in 2010 – Part 2: The global gridded agricultural-production maps. *Earth Syst. Sci. Data* **12**, 3545–3572 (2020).
31. Acikgoz, E. *et al.* Genotype x environment interaction and stability analysis for dry matter and seed yield in field pea (*Pisum sativum* L.). *Span. J. Agric. Res.* **7**, 96 (2009).
32. Amalfitano, C. *et al.* Production system effects on growth, pod yield and seed quality of organic faba bean in southern Italy. *Folia Hortic.* **30**, 375–385 (2018).
33. Annicchiarico, P., Nazzicari, N., Pecetti, L., Romani, M. & Russi, L. Pea genomic selection for Italian environments. *BMC Genom.* **20**, 603 (2019).
34. Arlauskienė, A. *et al.* Competitiveness and productivity of organically grown pea and spring cereal intercrops. *Žemdirbystė (Agriculture)* **98**, 339–348 (2011).
35. Assaead, A. M., McGowan, M., Hebblethwaite, P. D. & Brereton, J. C. Effect of soil compaction on growth, yield and light interception of selected crops. *Ann. Appl. Biol.* **117**, 653–666 (1990).
36. Avola, G., Tuttobene, R., Gresta, F. & Abbate, V. Weed control strategies for grain legumes. *Agron. Sustain. Dev.* **28**, 389–395 (2008).
37. Avola, G., Riggi, E., Gresta, F., Sortino, O. & Onofri, A. Random effects models, BLUPs and redundancy analyses for grain legume crops in semi-arid environments. *Eur. J. Agron.* **93**, 18–26 (2018).
38. Badagliacca, G. *et al.* Long-term no-tillage application increases soil organic carbon, nitrous oxide emissions and faba bean (*Vicia faba* L.) yields under rain-fed Mediterranean conditions. *Sci. Total Environ.* **639**, 350–359 (2018).
39. Baldivieso-Freitas, P. *et al.* Crop yield, weed infestation, and soil fertility responses to contrasted ploughing intensity and manure additions in a Mediterranean organic crop rotation. *Soil Tillage Res.* **180**, 10–20 (2018).
40. Barrios, A., Aparicio, T., Rodríguez, M. J., Pérez de la Vega, M. & Caminero, C. Winter sowing of adapted lines as a potential yield increase strategy in lentil (*Lens culinaris* Medik.). *Span. J. Agric. Res.* **14**, e0702 (2016).
41. Berti, A., Bravin, F. & Zanin, G. Application of decision-support software for postemergence weed control. *Weed Sci.* **51**, 618–627 (2003).
42. Bouniols, A. *et al.* Simulation of soybean nitrogen nutrition for a silty clay soil in southern France. *Field Crops Res.* **26**, 19–34 (1991).
43. Cabelguenne, M., Debaeke, P. & Bouniols, A. EPICphase, a version of the EPIC model simulating the effects of water and nitrogen stress on biomass and yield, taking account of developmental stages: validation on maize, sunflower, sorghum, soybean and winter wheat. *Agric. Syst.* **60**, 175–196 (1999).
44. Caruso, C., Maucieri, C., Cavallaro, V., Borin, M. & Barbera, A. C. Olive mill wastewater spreading and AMF inoculation effects in a low-input semi-arid Mediterranean crop succession. *Arch. Agron. Soil Sci.* **64**, 2060–2074 (2018).
45. Casa, R. & Cascio, L. B. Soil conservation tillage effects on yield and water use efficiency on irrigated crops in Central Italy. *J. Agron. Crop Sci.* **194**, 310–319 (2008).
46. Celik, I. *et al.* Impacts of different tillage practices on some soil microbiological properties and crop yield under semi-arid Mediterranean conditions. *Int. J. Plant Prod.* **5**, 237–254 (2011).
47. Celik, I. *et al.* Long-term tillage and residue management effect on soil compaction and nitrate leaching in a Typic Haploxerert soil. *Int. J. Plant Product.* **11**, 131–149 (2017).
48. Cociu, A. I. & Cizmas, G. D. Effects of stabilization period of conservation agriculture practices on winter wheat, maize and soybean crops, in rotation. *Roman. Agric. Res.* **30**, 171–181 (2013).
49. Cociu, A. I. Long-term effects of tillage systems on winter wheat, maize and soybean grain yield and yield stability under rainfed conditions in the eastern romanian Danube plain. *Roman. Agric. Res.* **22**, 91–109 (2016).
50. Colomb, B., Debaeke, P., Jouany, C. & Nolot, J. M. Phosphorus management in low input stockless cropping systems: Crop and soil responses to contrasting P regimes in a 36-year experiment in southern France. *Eur. J. Agron.* **26**, 154–165 (2007).
51. Colson, J., Bouniols, A. & Jones, J. W. Soybean reproductive development: adapting a model for European cultivars. *Agron. J.* **87**, 1129–1139 (1995).
52. Çopur, O., Gur, M. A., Ufuk Demirel, U. & Karakus, M. Performance of some soybean [*Glycine max*(L.) Merr.] genotypes double cropped in semi-arid conditions. *Not. Bot. Hort. Agrobot. Cluj.* **37**, 85–91 (2009).
53. Corre-Hellou, G., Fustec, J. & Crozat, Y. Interspecific competition for soil n and its interaction with n<sub>2</sub> fixation, leaf expansion and crop growth in pea–barley intercrops. *Plant Soil* **282**, 195–208 (2006).
54. de Pascale, S. & Barbieri, G. Effects of soil salinity and top removal on growth and yield of broadbean as a green vegetable. *Sci Hortic.* **71**, 147–165 (1997).
55. Debaeke, P. & Hilaire, A. Production of rainfed and irrigated crops under different crop rotations and input levels in southwestern France. *Can. J. Plant Sci.* **77**, 539–548 (1997).
56. Dhima, K., Vasilakoglou, I., Paschalidis, K. A., Gatsis, T. & Keco, R. Productivity and phytotoxicity of six sunflower hybrids and their residues effects on rotated lentil and ivy-leaved speedwell. *Field Crops Res.* **136**, 42–51 (2012).
57. di Paolo, E., Garofalo, P. & Rinaldi, M. Irrigation and nitrogen fertilization treatments on productive and qualitative traits of broad bean (*Vicia faba* var. *minor* L.) in a Mediterranean environment. *Leg. Res.* **38**, 209 (2015).
58. Diacono, M., Baldivieso-Freitas, P. & Sans Serra, F. Nitrogen utilization in a cereal–legume rotation managed with sustainable agricultural practices. *Agronomy* **9**, 113 (2019).
59. Dogan, E., Kirnak, H. & Copur, O. Effect of seasonal water stress on soybean and site specific evaluation of CROPGRO-Soybean model under semi-arid climatic conditions. *Agric. Water Manag.* **90**, 56–62 (2007).
60. Dogan, E., Kirnak, H. & Copur, O. Deficit irrigations during soybean reproductive stages and CROPGRO-soybean simulations under semi-arid climatic conditions. *Field Crops Res.* **103**, 154–159 (2007).
61. Dolijanovic, Z. *et al.* Dependence of the productivity of maize and soybean intercropping systems on hybrid type and plant arrangement pattern. *Genetika* **45**, 135–144 (2013).
62. Dozet, G. *et al.* The effect of preceding crops nitrogen fertilization and cobalt and molybdenum application on yield and quality of soybean grain. *Roman. Agr. Res.* **33**, 1–11 (2016).

63. Dragicevic, V. *et al.* Effect of the maize–soybean intercropping system on the potential bioavailability of magnesium, iron and zinc. *Crop Pasture Sci.* **66**, 1118 (2015).
64. Đukić, V. *et al.* Pulsed electromagnetic field – a cultivation practice used to increase soybean seed germination and yield. *Zemdirbyste-Agriculture* **104**, 345–352 (2017).
65. Elkoca, E., Kantar, F. & Zengin, H. Weed control in lentil (*Lens culinaris* L.) in eastern Turkey. *N. Z. J. Crop Hort. Sci.* **33**, 223–231 (2005).
66. Elkoca, E. & Kantar, F. Response of pea (*Pisum sativum* L.) to mepiquat chloride under varying application doses and stages. *J. Agron. Crop Sci.* **192**, 102–110 (2006).
67. Fernández-García, P., López-Bellido, L., Muñoz-Romero, V. & López-Bellido, R. J. Chickpea water use efficiency as affected by tillage in rainfed Mediterranean conditions. *Agric. Water Manag.* **129**, 194–199 (2013).
68. Flores, F. *et al.* Faba bean adaptation to autumn sowing under European climates. *Agron. Sustain. Dev.* **32**, 727–734 (2012).
69. Fotiadis, S., Koutroubas, S. D. & Damalas, C. A. Nitrogen utilization and yield determination of spring mediterranean chickpea as influenced by planting date and environmental conditions. *Int. J. Plant Prod.* **13**, 59–72 (2019).
70. Gajić, B., Kresović, B., Tapanarova, A., Životić, L. & Todorović, M. Effect of irrigation regime on yield, harvest index and water productivity of soybean grown under different precipitation conditions in a temperate environment. *Agric. Water Manag.* **210**, 224–231 (2018).
71. Gaurilėikiene, I., Janušauskaite, D., Česnulevičienė, R. & Ramanauskienė, J. The suppression of stem base and root rot diseases of pea as affected by fungicidal seed treatment. *Zemdirbyste-Agriculture* **95**, 50–57 (2008).
72. Gawęda, D., Haliniarz, M., Cierpiało, R. & Klusek, I. Yield, weed infestation and seed quality of soybean (*Glycine max* (L.) Merr.) under different tillage systems. *J. Agric. Sci.* **23**, 268–275 (2015).
73. Ghaouti, L., Vogt-Kaute, W. & Link, W. Development of locally-adapted faba bean cultivars for organic conditions in Germany through a participatory breeding approach. *Euphytica* **162**, 257–268 (2008).
74. Giambalvo, D. *et al.* Faba bean grain yield, N<sub>2</sub> fixation, and weed infestation in a long-term tillage experiment under rainfed Mediterranean conditions. *Plant Soil* **360**, 215–227 (2012).
75. Grevsen, K. weed competitive ability of green peas (*Pisum sativum* L.) affected by seeding rate and genotype characteristics. *Biol. Agric. Hort.* **21**, 247–261 (2003).
76. Gronle, A., Heß, J. & Böhm, H. Effect of intercropping normal-leaved or semi-leafless winter peas and triticale after shallow and deep ploughing on agronomic performance, grain quality and succeeding winter wheat yield. *Field Crops Res.* **180**, 80–89 (2015).
77. Gulluoglu, L., Arioglu, H. & Kurt, C. Adaptability and stability of new soybean cultivars under double cropped conditions of Turkey. *African J. Agric. Res.* **6**, 3320–3325 (2011).
78. Gunes, A. *et al.* Mineral nutrition of wheat, chickpea and lentil as affected by mixed cropping and soil moisture. *Nutr. Cycl. Agroecosyst.* **78**, 83–96 (2007).
79. Halwani, M., Reckling, M., Schuler, J., Bloch, R. & Bachinger, J. Soybean in No-Till Cover-Crop Systems. *Agronomy* **9**, 883 (2019).
80. Hauggaard-Nielsen, H., Ambus, P. & Jensen, E. S. Interspecific competition, N use and interference with weeds in pea–barley intercropping. *Field Crops Res.* **70**, 101–109 (2001).
81. Hauggaard-Nielsen, H., Ambus, P. & Jensen, E. S. The comparison of nitrogen use and leaching in sole cropped versus intercropped pea and barley. *Nutr. Cycl. Agroecosyst.* **65**, 289–300 (2003).
82. Hauggaard-Nielsen, H., Andersen, M. K., Jørnsgaard, B. & Jensen, E. S. Density and relative frequency effects on competitive interactions and resource use in pea–barley intercrops. *Field Crops Res.* **95**, 256–267 (2006).
83. Hauggaard-Nielsen, H., Jørnsgaard, B., Kinane, J. & Jensen, E. S. Grain legume–cereal intercropping: the practical application of diversity, competition and facilitation in arable and organic cropping systems. *Renew. Agric. Food Syst.* **23**, 3–12 (2008).
84. Hauggaard-Nielsen, H., Mundus, S. & Jensen, E. S. Grass-clover undersowing affects nitrogen dynamics in a grain legume–cereal arable cropping system. *Field Crops Res.* **136**, 23–31 (2012).
85. Hawes, C. *et al.* Plant responses to an integrated cropping system designed to maintain yield whilst enhancing soil properties and biodiversity. *Agronomy* **8**, 229 (2018).
86. Helenius, J. & Jokinen, K. Yield advantage and competition in intercropped oats (*Avena sativa* L.) and faba bean (*Vicia faba* L.): application of the hyperbolic yield-density model. *Field Crops Res.* **37**, 85–94 (1994).
87. Horneburg, B. & Becker, H. C. Crop adaptation in on-farm management by natural and conscious selection: a case study with lentil. *Crop Sci.* **48**, 203–212 (2008).
88. Iliadis, G. C., Roupakias, D. G. & Goulas, C. K. Effectiveness of honeycomb selection for yield superiority at three interplant distances: a field simulation study using chickpea (*Cicer arietinum* L.) inbred lines. *Euphytica* **133**, 299–311 (2003).
89. Jannoura, R., Bruns, C. & Joergensen, R. G. Organic fertilizer effects on pea yield, nutrient uptake, microbial root colonization and soil microbial biomass indices in organic farming systems. *Eur. J. Agron.* **49**, 32–41 (2013).
90. Jensen, E. S. Grain yield, symbiotic N<sub>2</sub> fixation and interspecific competition for inorganic N in pea–barley intercrops. *Plant Soil* **182**, 25–38 (1996).
91. Jovanovic, Ž., Djalovic, I., Komljenovic, I., Kovacevic, V. & Cvijović, M. Influences of liming on vertisol properties and yields of the field crops. *Cereal Res. Commun.* **34**, 517–520 (2006).
92. Jug, D., Stipesevic, B. & Zucek, I. Effects of conventional and reduced tillage systems in winter wheat — soybean crop rotation on crops biomass development. *Cereal Res. Commun.* **34**, 1137–1143 (2006).
93. Jurado-Exposito, M., Garcia-Torres, L. & Castejon-Munoz, M. Broad bean and lentil seed treatments with imidazolinones for the control of broomrape (*Orobanche crenata*). *J. Agric. Sci.* **129**, 307–314 (1997).
94. Jurjescu, A., David, G., Paul, P. & Simion, A. The behaviour of some varieties of pea, soybean and lentils concerning the relationship bacterization-sowing period on the yield and quality in Timiș plain. *Res. J. Agric. Sci.* **42**, 82–87 (2010).
95. Kadziulienė, Z. *et al.* Qualitative effects of pea and spring cereals intercrop in the organic farming systems. *Agron. Res.* **7**, 606–611 (2009).
96. Kahraman, A., Khan, M. K., Pandey, A. & Dogan, E. Effect of Supplemental Irrigation on Lentil Yield and Growth in Semi-Arid Environment. *Not. Bot. Horti. Agrobot. Cluj Napoca* **44**, 237–244 (2016).
97. Kalapchieva, S., Topalova, E. & Petkova, V. Morphological, physiological and productivity response in garden pea genotypes during high temperature stress. *Genetika* **51**, 417–428 (2019).
98. Kaul, H.-P., Aufhammer, W. & Wäagner, W. Dry matter and nitrogen accumulation and residues of oil and protein crops. *Eur. J. Agron.* **5**, 137–147 (1996).
99. Klimek-Kopyra, A., Kulig, B., Oleksy, A. & Zajac, T. Agronomic performance of naked oat (*Avena nuda* L.) and faba bean intercropping. *Chil. J. Agric. Res.* **75**, 168–173 (2015).
100. Knudsen, T. M., Hauggaard-Nielsen, H., Jørnsgaard, B. & Jensen, E. S. Comparison of interspecific competition and N use in pea–barley, faba bean–barley and lupin–barley intercrops grown at two temperate locations. *J. Agric. Sci.* **142**, 617–627 (2004).
101. Kocira, S. Effect of amino acid biostimulant on the yield and nutraceutical potential of soybean. *Chil. J. Agric. Res.* **79**, 17–25 (2019).
102. Kontturi, M. *et al.* Pea–oat intercrops to sustain lodging resistance and yield formation in northern European conditions. *Acta Agric. Scand. B Soil Plant Sci.* **61**, 612–621 (2011).
103. Kunz, C., Weber, J. F., Peteinatos, G. G., Sökefeld, M. & Gerhards, R. Camera steered mechanical weed control in sugar beet, maize and soybean. *Precis. Agric.* **19**, 708–720 (2018).

104. Lizarazo, C. I., Isotalo, J. & Lindfors, A. v. & Stoddard, F. L. Progress towards flowering of faba bean (*Vicia faba* L.) is more than photothermal. *J. Agron. Crop Sci.* **203**, 385–396 (2017).
105. López-Bellido, R. J., López-Bellido, L., López-Bellido, F. J. & Castillo, J. E. Faba Bean (*Vicia faba* L.) response to tillage and soil residual nitrogen in a continuous rotation with wheat (*Triticum aestivum* L.) under rainfed Mediterranean conditions. *Agron. J.* **95**, 1253–1261 (2003).
106. López-Bellido, L., López-Bellido, R. J., Castillo, J. E. & López-Bellido, F. J. Chickpea response to tillage and soil residual nitrogen in a continuous rotation with wheat. *Field Crops Res.* **88**, 191–200 (2004).
107. López-Bellido, R. J. *et al.* Chickpea and faba bean nitrogen fixation in a Mediterranean rainfed Vertisol: Effect of the tillage system. *Eur. J. Agron.* **34**, 222–230 (2011).
108. Madeira, A. C., Clark, J. A. & Rossall, S. Growth and light interception in field bean (*Vicia faba*) infected by *Ascochyta fabae*. *J. Agric. Sci.* **123**, 225–232 (1994).
109. Martínez, J., Vidal, D. & Simón, E. Nitrogenase activity, nodulation and seed production in *Vicia faba* as affected by methabenzthiazuron. *J. Agric. Sci.* **127**, 319–324 (1996).
110. Martiniello, P. & Ciola, A. Dry matter and seed yield of mediterranean annual legume species. *Agron. J.* **87**, 985–993 (1995).
111. Mazzoncini, M. *et al.* Rainfed wheat and soybean productivity in a long-term tillage experiment in central Italy. *Agron. J.* **100**, 1418–1429 (2008).
112. Monti, M., Pellicanò, A., Santonoceto, C., Preiti, G. & Pristeri, A. Yield components and nitrogen use in cereal-pea intercrops in Mediterranean environment. *Field Crops Res.* **196**, 379–388 (2016).
113. Moschini, V. *et al.* Performance of organic grain legumes in Tuscany. *Ital. J. Agron.* **9**, 38 (2014).
114. Muñoz-Romero, V., López-Bellido, L. & López-Bellido, R. J. Faba bean root growth in a Vertisol: Tillage effects. *Field Crops Res.* **120**, 338–344 (2011).
115. Naudin, C., Corre-Hellou, G., Pineau, S., Crozat, Y. & Jeuffroy, M.-H. The effect of various dynamics of N availability on winter pea–wheat intercrops: Crop growth, N partitioning and symbiotic N<sub>2</sub> fixation. *Field Crops Res.* **119**, 2–11 (2010).
116. Neugschwandtner, R. W., Wichmann, S., Gimplinger, D. M., Wagentristl, H. & Kaul, H. P. Chickpea performance compared to pea, barley and oat in central Europe: growth analysis and yield. *Turk. J. Field Crops.* **18**, 179–184 (2013).
117. Neugschwandtner, R., Ziegler, K., Krieger, S., Wagentristl, H. & Kaul, H.-P. Nitrogen yield and nitrogen fixation of winter faba beans. *Acta Agric. Scand. B Soil Plant Sci.* **65**, 658–666 (2015).
118. Neugschwandtner, R. W. *et al.* Agronomic potential of winter grain legumes for central Europe: development, soil coverage and yields. *Field Crops Res.* **241**, 107576 (2019).
119. Neumann, A., Schmidtke, K. & Rauber, R. Effects of crop density and tillage system on grain yield and N uptake from soil and atmosphere of sole and intercropped pea and oat. *Field Crops Res.* **100**, 285–293 (2007).
120. Niță, S. *et al.* Results obtained for soybean, pea and lentils crops on a cambic chernozem in the Banat's plain during 2008–2010. *Roman. Agr. Res.* **29**, 155–162 (2012).
121. Ogórek, R., Lejman, A. & Sobkowicz, P. effect of the intensity of weed harrowing with spike-tooth harrow in barley-pea mixture on yield and mycobiota of harvested grains. *Agronomy* **9**, 103 (2019).
122. Olle, M., Williams, I. H. & Rosa, E. Selecting appropriate faba bean var. minor varieties for production under Northern European environmental conditions. *Acta Agric. Scand. B Soil Plant Sci.* **69**, 432–438 (2019).
123. Onat, B., Bakal, H., Gulluoglu, L. & Arioglu, H. The effects of high temperature at the growing period on yield and yield components of soybean [*Glycine max* (L.) Merr] varieties. *Turk. J. Field Crops* <https://doi.org/10.17557/tjfc.356210> (2017).
124. Ortas, I. The effect of mycorrhizal fungal inoculation on plant yield, nutrient uptake and inoculation effectiveness under long-term field conditions. *Field Crops Res.* **125**, 35–48 (2012).
125. Pampana, S., Masoni, A., Mariotti, M., Ercoli, L. & Arduini, I. Nitrogen fixation of grain legumes differs in response to nitrogen fertilisation. *Exp. Agric.* **54** (2018).
126. Pannacci, E. & Tei, F. Effects of mechanical and chemical methods on weed control, weed seed rain and crop yield in maize, sunflower and soybean. *Crop Prot.* **64**, 51–59 (2014).
127. Pannacci, E., Tei, F. & Guiducci, M. Evaluation of mechanical weed control in legume crops. *Crop Prot.* **104**, 52–59 (2018).
128. Pannecoucq, J. *et al.* Screening for soybean varieties suited to Belgian growing conditions based on maturity, yield components and resistance to *Sclerotinia sclerotiorum* and *Rhizoctonia solani* anastomosis group 2-IIIB. *J. Agric. Sci.* **156**, 342–349 (2018).
129. Pappa, V. A., Rees, R. M., Walker, R. L., Baddeley, J. A. & Watson, C. A. Legumes intercropped with spring barley contribute to increased biomass production and carry-over effects. *J. Agric. Sci.* **150**, 584–594 (2012).
130. Perego, A. *et al.* Agro-environmental aspects of conservation agriculture compared to conventional systems: A 3-year experience on 20 farms in the Po valley (Northern Italy). *Agric. Syst.* **168**, 73–87 (2019).
131. Pietola, L. & Elonen, P. Sprinkler irrigation of field crops during rainy growing season. *Ann. Agric. Fenn.* **30**, 359–373 (1991).
132. Pilbeam, C. J., Akatse, J. K., Hebblethwaite, P. D. & Wright, S. D. Yield production in two contrasting forms of spring-sown faba bean in relation to water supply. *Field Crops Res.* **29**, 273–287 (1992).
133. Plaza-Bonilla, D., Nolot, J.-M., Passot, S., Raffaillac, D. & Justes, E. Grain legume-based rotations managed under conventional tillage need cover crops to mitigate soil organic matter losses. *Soil Tillage Res.* **156**, 33–43 (2016).
134. Podgórska-Lesiak, M. & Sobkowicz, P. Prevention of pea lodging by intercropping barley with peas at different nitrogen fertilization levels. *Field Crops Res.* **149**, 95–104 (2013).
135. Pöttsch, F., Lux, G. & Schmidtke, K. Sulphur demand, uptake and fertilization of *Vicia faba* L. under field conditions. *Field Crops Res.* **228**, 76–83 (2018).
136. Pöttsch, F., Lux, G., Lewandowska, S., Bellingrath-Kimura, S. D. & Schmidtke, K. Optimizing relative seed frequency of intercropped pea and spring barley. *Eur. J. Agron.* **105**, 32–40 (2019).
137. Radicetti, E., Mancinelli, R. & Campiglia, E. Combined effect of genotype and inter-row tillage on yield and weed control of chickpea (*Cicer arietinum* L.) in a rainfed Mediterranean environment. *Field Crops Res.* **127**, 161–169 (2012).
138. Radicetti, E., Mancinelli, R. & Campiglia, E. The competitive ability of different chickpea (*Cicer arietinum*) genotypes against *Polygonum aviculare* under field conditions. *Crop Prot.* **42**, 312–319 (2012).
139. Rauber, R., Schmidtke, K. & Kimpel-Freund, H. Konkurrenz und ertragsvorteile in gemengen aus erbsen (*Pisum sativum* L.) und hafer (*Avena sativa* L.). *J. Agron. Crop Sci.* **185**, 33–47 (2000).
140. Rodiño, A. P., Hernández-Nistal, J., Hermida, M., Santalla, M. & de Ron, A. M. Sources of variation for sustainable field pea breeding. *Euphytica* **166**, 95–107 (2009).
141. Rubio, J., Flores, F., Moreno, M. T., Cubero, J. I. & Gil, J. Effects of the erect/bushy habit, single/double pod and late/early flowering genes on yield and seed size and their stability in chickpea. *Field Crops Res.* **90**, 255–262 (2004).
142. Ruisi, P. *et al.* Tillage effects on yield and nitrogen fixation of legumes in Mediterranean conditions. *Agron. J.* **104**, 1459–1466 (2012).
143. Santín-Montanyá, M. I., Zambrana, E., Fernández-Getino, A. P. & Tenorio, J. L. Dry pea (*Pisum sativum* L.) yielding and weed infestation response, under different tillage conditions. *Crop Prot.* **65**, 122–128 (2014).
144. Sarunaite, L., Deveikyte, I. & Kadziulienė, Z. Intercropping spring wheat with grain legume for increased production in an organic crop rotation. *Zemdirbyste-Agriculture* **97**, 51–58 (2010).
145. Sau, F. & Mínguez, I. M. Adaptation of indeterminate faba beans to weather and management under a Mediterranean climate. *Field Crops Res.* **66**, 81–99 (2000).

146. Schröder, D. & Köpke, U. Faba bean (*Vicia faba* L.) intercropped with oil crops – a strategy to enhance rooting density and to optimize nitrogen use and grain production? *Field Crops Res* **135**, 74–81 (2012).
147. Šiaudinis, G. The Effect of bacterial application on the productivity of faba bean (*Vicia faba* L.) and its mixtures with spring wheat (*Triticum aestivum* L.) under two agroclimatic conditions in Lithuania. *Appl. Ecol. Environ. Res.* **15**, 2011–2021 (2017).
148. Siczek, A., Horn, R., Lipiec, J., Usowicz, B. & Łukowski, M. Effects of soil deformation and surface mulching on soil physical properties and soybean response related to weather conditions. *Soil Tillage Res* **153**, 175–184 (2015).
149. Silva, L. L., Duarte, I., Lourenço, E., Simões, N. & Chaves, M. M. Yield and water productivity of five chickpea varieties under supplemental irrigation in contrasting years. *Irrig. Sci.* **32**, 393–403 (2014).
150. Sobkowicz, P. Competition between triticale (*Triticosecale* Witt.) and field beans (*Vicia faba* var. *minor* L.) in additive intercrops. *Plant Soil Environ.* **52**, 47–54 (2006).
151. Stipesevic, B., Jug, D., Jug, I., Tolimir, M. & Cvijovic, M. Winter wheat and soybean zinc uptake in different soil tillage systems. *Cereal Res. Commun.* **37**, 305–310 (2009).
152. Stützel, H., Aufhammer, W. & Löber, A. Effects of sowing technique on yield formation of *Vicia faba* as affected by population density, sowing date and plant type. *J. Agric. Sci.* **122**, 255–264 (1994).
153. Sulas, L., Roggero, P. P., Canu, S. & Seddaiu, G. Potential nitrogen source from field bean for rainfed mediterranean cropping systems. *Agron J.* **105**, 1735–1742 (2013).
154. Szparaga, A. *et al.* Modification of growth, yield, and the nutraceutical and antioxidative potential of soybean through the use of synthetic biostimulants. *Front. Plant Sci.* **9** (2018).
155. Tekin, S., Yazar, A. & Barut, H. Comparison of wheat-based rotation systems and monocropping systems under dryland Mediterranean conditions. *Int. J. Agric. Biol. Eng.* **10**, 203–213 (2017).
156. Tepe, I., Erman, M., Yazlik, A., Levent, R. & Ipek, K. Comparison of some winter lentil cultivars in weed–crop competition. *Crop Prot.* **24**, 585–589 (2005).
157. Theologidou, G. S., Tsialta, J. T., Kaloumenos, N. & Eleftherohorinos, I. G. From Petri dish to field: testing Greek lentil accessions for imazamox tolerance. *Int. J. Plant Prod.* **10**, 265–274 (2016).
158. Toleikienė, M. *et al.* The introduction of soybean in an organic crop rotation in the Nemoral zone: the impact on subsequent spring wheat productivity. *Zemdirbyste-Agriculture* **106**, 321–328 (2019).
159. Turk, Z. & Kendal, E. The practice of AMMI and GGE Biplot Analysis of lentil genotypes assessment in multi-environment trials. *Philippine J. Crop Sci.* **42**, 39–48 (2017).
160. Urbatzka, P., Graß, R., Haase, T., Schüler, C. & Heß, J. Influence of different sowing dates of winter pea genotypes on winter hardiness and productivity as either winter catch crop or seed legume. *Eur. J. Agron.* **40**, 112–119 (2012).
161. Vasilakoglou, I., Vlachostergios, D., Dhima, K. & Lithourgidis, A. Response of vetch, lentil, chickpea and red pea to pre- or post-emergence applied herbicides. *Span. J. Agric. Res.* **11**, 1101 (2013).
162. Vlachostergios, D. N., Lithourgidis, A. S. & Roupakias, D. G. Effectiveness of single-plant selection at low density under organic environment: a field study with lentil. *Crop Sci.* **51**, 41–51 (2011).
163. Vlachostergios, D. N. *et al.* Single-plant selection within lentil landraces at ultra-low density: a short-time tool to breed high yielding and stable varieties across divergent environments. *Euphytica* **214**, 58 (2018).
164. Vocanson, A. & Jeuffroy, M. Agronomic performance of different pea cultivars under various sowing periods and contrasting soil structures. *Agron. J.* **100**, 748–759 (2008).
165. Vollmann, J., Wagentrust, H. & Hartl, W. The effects of simulated weed pressure on early maturity soybeans. *Eur. J. Agron.* **32**, 243–248 (2010).
166. Volpi, I. *et al.* Minimum tillage mitigated soil N<sub>2</sub>O emissions and maximized crop yield in faba bean in a Mediterranean environment. *Soil Tillage Res.* **178**, 11–21 (2018).
167. Wang, L., Gruber, S. & Claupein, W. Effects of woodchip mulch and barley intercropping on weeds in lentil crops. *Weed Res.* **52**, 161–168 (2012).
168. Xu, J. *et al.* Effects of 16-Year woodchip mulching on weeds and yield in organic farming. *Agron. J.* **110**, 359–368 (2018).
169. Zeid, M., Schön, C.-C. & Link, W. Hybrid performance and AFLP- based genetic similarity in faba bean. *Euphytica* **139**, 207–216 (2004).
170. Zhelyazkova, T., Pavlov, D., Delchev, G. & Stoyanova, A. Productivity and yield stability of six grain legumes in the moderate climatic conditions of Bulgaria. *Sci. Papers Ser. A. Agron.* **478–487** (2016).
171. Zimmer, S. *et al.* Effects of soybean variety and *Bradyrhizobium* strains on yield, protein content and biological nitrogen fixation under cool growing conditions in Germany. *Eur. J. Agron.* **72**, 38–46 (2016).
172. Živanov, D. *et al.* Intercropping of field pea with annual legumes for increasing grain yield production. *Zemdirbyste-Agriculture* **105**, 235–242 (2018).
173. Urbatzka, P. *et al.* Grain yield and quality characteristics of different genotypes of winter pea in comparison to spring pea for organic farming in pure and mixed stands. *Org. Agric.* **1**, 187–202 (2011).
174. Neumann, A., Werner, J. & Rauber, R. Evaluation of yield–density relationships and optimization of intercrop compositions of field-grown pea–oat intercrops using the replacement series and the response surface design. *Field Crops Res.* **114**, 286–294 (2009).
175. Jannoura, R., Joergensen, R. G. & Bruns, C. Organic fertilizer effects on growth, crop yield, and soil microbial biomass indices in sole and intercropped peas and oats under organic farming conditions. *Eur. J. Agron.* **52**, 259–270 (2014).
176. Vasilakoglou, I. & Dhima, K. Leafy and semi-leafless field pea competition with winter wild oat as affected by weed density. *Field Crops Res.* **126**, 130–136 (2012).
177. Antichi, D. *et al.* An experimental dataset on yields of pulses across Europe., *Figshare*, <https://doi.org/10.6084/m9.figshare.c.6573760.v1> (2023).
178. Mariotti, M., Andreuccetti, V., Arduini, I., Minieri, S. & Pampana, S. Field bean for forage and grain in short-season rainfed Mediterranean conditions. *Ital. J. Agron.* **13**, 208–215 (2018).
179. Pampana, S., Masoni, A. & Arduini, I. Grain legumes differ in nitrogen accumulation and remobilisation during seed filling. *Acta Agric. Scand. B Soil Plant Sci.* **66** (2016).
180. Dutta, A., Trivedi, A., Nath, C. P., Gupta, D. & Hazra, K. K. A comprehensive review on grain legumes as climate-smart crops: challenges and prospects. *Environ. Challenges* **7**, 100479 (2022).
181. Darch, T. *et al.* The effect of soil type on yield and micronutrient content of pasture species. *PLoS One* **17**, e0277091 (2022).
182. Weindl, I. *et al.* Sustainable food protein supply reconciling human and ecosystem health: a Leibniz position. *Glob. Food Sec.* **25**, 100367 (2020).
183. Muñoz-Sabater, J. *et al.* ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. *Earth Syst. Sci. Data* **13**, 4349–4383 (2021).
184. Hengl, T. *et al.* SoilGrids250m: Global gridded soil information based on machine learning. *PLoS One* **12**, e0169748 (2017).
185. Iizumi, T., Okada, M. & Yokozawa, M. A meteorological forcing data set for global crop modeling: development, evaluation, and intercomparison. *J. Geophys. Res.* **119**, 363–384 (2014).
186. Page, M. J. *et al.* The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* **n71** <https://doi.org/10.1136/bmj.n71> (2021).
187. Beck, H. E. *et al.* Present and future Köppen–Geiger climate classification maps at 1-km resolution. *Sci. Data* **5**, 180214 (2018).

## Acknowledgements

This work was done in the frame of Task 1.2 of LEGVALUE project, which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement Nr. 727672. The Authors acknowledge all the LEGVALUE partners which contributed by sharing their experimental data: INRAE, TERRES INOVIA, UNIPI, WU, PGRO, WR, SSSA, INIAV, AICF, SEGES, LLKC, FH-SWF, CRAN, LAMMC, FiBL. The authors are also grateful to the partners of the LEGATO Project "LEGumes for the Agriculture of TOMorrow", project number 613551 funded by the European Union under the FP7 Programme for having kindly provided their experimental data for the database.

## Author contributions

D.A. and N.G. conceived and supervised the work. D.M. contributed to design the work. I.S., V.B., and O.B. contributed data/evidence collection. D.A., S.P., L.T. performed the literature search, read the papers, and extracted data. D.A., S.P., L.T., N.G. compiled data records, performed standardization and technical validation of the database. D.A., S.P., L.T. supervised and validated the findings. D.A., S.P., L.T., I.D., Z.K., B.H., N.G. contributed to the analysis of the results. D.A., S.P., N.G. wrote the manuscript. D.M., S.P. revised the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41597-023-02606-0>.

**Correspondence** and requests for materials should be addressed to D.A.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023