

Reviewer 1

I recommend minor revision. The manuscript describes EuroCrops v2.0, a multi-annual harmonized parcel-level crop-type dataset built from publicly available EU Geo-Spatial Application (GSA) declarations (18 Paying Agencies; minimum three consecutive years including 2021), harmonized via the updated HCAT v4 taxonomy and linked to key EU statistical/survey and EO classification systems. The paper is timely and valuable for CAP-related analysis, crop-rotation studies, and as training/validation data for EO products; the processing chain and interoperability ambition are strong

Thank you for the positive comments about the paper.

Handling missing crop codes: you gap-fill codes where possible and otherwise create new unique codes (starting at 10001) “in a random order.” Please clarify the randomization (seed; stability across releases) and provide a quick summary of how frequent this situation is by country/year (e.g., Estonia/Ireland name-only years

We did not randomly assign codes using a random number generator. Instead, we created new unique codes when there was a gap that arose so this was done manually. We have removed ‘in a random order’ from the text and clarified this. We have provided a table in the paper (now Table 1) that summarizes how frequently this situation occurs and is also provided below:

Country/region	# of crop codes	# of missing codes	% of missing codes
AT	246	1	0.41
BE2	362	1	0.28
BE3	-	-	-
BG	272	272	100
CZ	415	1	0.24
DE4	299	29	9.7
DEA	288	1	0.35
DK	508	110	21.65
EE	313	313	100
ES	-	-	-
FI	-	-	-
FR	-	-	-
IE	214	214	100
IT11	-	-	-
NL	492	6	1.22
PT	312	129	41.35
SI	181	2	1.1
SK	282	282	100

Translation + matching quality: you translate via multiple MT APIs (Google/DeepL/OPUS-MT and sometimes ChatGPT), then manually check disagreements, and match via weighted Levenshtein plus rules; you note translation/matching challenges (e.g., ~58% for Finnish, ~51% matched for Brandenburg). Please add a small, systematic quality assessment so users understand typical error modes (spelling mistakes, colloquial names, mixtures)

Before the crop names were translated from their native language to English, a pre-processing step was undertaken to align the crop names across the years because countries provide an updated crop code list on an annual basis and the EuroCropsv2.0 process has been ongoing since 2021. The github now includes a table with the original_code and how these are related to multiple instances of original_name. The process for consolidating the dataset (i.e., having unique pairs of original_code/original_name) involved the use of an LLM (gpt-4o and nous-hermes-2-mixtral-8x7b-dpo), or a manual process. The source of the consolidated original_name is provided in this table here:

https://github.com/Martincccc/EuroCropsV2/blob/main/data/processing/name_mistakes.csv.

The aligned crop names were then translated using three machine learning packages. We have uploaded one table per country (XX_trans_stats.csv) with the original codes and their translations from Google Translate (g_name), DeepL (dpl_name) and OPUS-MT (opus_name) to github, placed in this folder:

https://github.com/Martincccc/EuroCropsV2/tree/main/data/eurocrop_trans, where XX in the filename is the two letter country code or the three letter regional code. Note that these files have been updated in 2026 so that they are reproducible since this process was done over several years (since 2021) as outlined above.

We attempted to provide the best translations possible, which is why a third machine translator was introduced as part of EuroCropsv2.0. At the beginning of EuroCropsv2.0, we discovered some wrong translations by DeepL and Google Translate. Therefore, we added OPUS-MT, which is not as good as the other two translators, accounting for 63-92% of the translation errors, but it served as an additional check. Without the use of OPUS-MT, the number of manual checks would have been reduced significantly, but we ran the risk of wrong translations that would not have been picked up using only DeepL and Google Translate.

The translation accuracy varies by country with the lowest in Spain at 40% to the highest in Belgium at around 72%. Estonia provides English translations of most labels on <https://klassifikaatorid.stat.ee/>, so a simple search there resolved most disagreements. Other countries do not have such convenient portals. Slovenia releases their labels along with the scientific name of the crops while Ireland's labels are in English. Spelling mistakes were rare except for the Spanish, Bulgarian, Czech, Portuguese and Italian GSA, which led to direct failures of translation and no results were provided by Google Translate. In the German GSA, there were many abbreviations and short forms of names used, which hindered the automatic translation. In the GSA of Czechia, there were extremely detailed species level crop descriptions of grasses and legumes as well as the use of colloquial terms. After machine translation, colloquial terms often led to non-crop-like descriptions such as diabetes, quiet, etc. Subsequent Google searches of these terms sometimes turned up empty, whereas for other niche translations, they showed up in Wikipedia. In other cases, we sent the colloquial terms to native speakers to review but this was not possible for all languages.

We have added a new section to the Supplementary Material (called Translation of crop names) and added all of this text. It includes a summary table, Table S1, which summarizes the number and percentage of multiple instances of original_name per country and the translation statistics and errors. The accuracy refers to the percentage of time when all the machine translations were in agreement. Where there was disagreement in the remaining translations, these were all manually checked. Of these manually checked translations, five different cases may have occurred. In the majority of cases, at least one machine translator was correct. Often OPUS-MT was the poorest

performer, but other times, e.g., in Slovakia, Google Translate and DeepL failed but OPUS-MT was correct. The second type of case included the occurrence of niche crop names, e.g., species level trees or legumes, which were confirmed by Wikipedia, or where the translator chose a more frequently used homonym, which were then checked using Wikipedia's disambiguation pages. A third case was the occurrence of colloquial or regional names, which were also checked using Wikipedia. Finally, spelling errors (case 4) and crop names that were abbreviated or truncated (case 5) were also identified through this manual checking process.

Table S1: Summary of instances of multiple original_name for crops and translation statistics

Country or Region	# of multiple original_name [%]	Accuracy	Manual Correction	Poor Machine Translation	Niche Crop/Homonymy	Colloquial/Regional name	Spelling errors	Abbreviation/Truncation
AT	13 [5%]	60.98	39.02	82.29	6.25	1.04	0	10.42
BE2	52 [14%]	71.82	28.18	92.16	6.86	0.98	0	0
BE3	75 [36%]	71.29	28.71	88.33	11.67	0	0	0
BG	-	58.3	41.7	72.57	24.78	1.77	0.88	0
CZ	146 [35%]	45.08	54.92	62.56	33.04	0.88	2.64	0
DE4	89 [30%]	67.69	32.31	77.13	12.77	0	0	10.11
DEA	51 [18%]	-	-	-	-	-	-	-
DK	286 [56%]	63.78	36.22	66.3	27.72	0	0	5.98
EE	-	43.77	55.23	76.14	16.3	0	0	0
ES	-	40.29	59.71	69.14	26.34	2.47	1.65	0.41
FI	-	57.2	42.8	85.71	14.29	0	0	0
FR	-	64.91	35.09	84.87	15.13	0	0	0
IE	-	-	-	-	-	-	-	-
ITI1	1 [0%]	59.35	40.65	75.51	23.81	0	0.68	0
NL	91 [18%]	76.42	23.58	87.93	8.62	0	0	3.45
PT	-	56.73	43.27	77.78	14.81	1.48	5.19	0.74
SI	8 [4%]	-	-	-	-	-	-	-
SK	-	63.48	36.52	88.35	9.71	1.94	0	0

*EE has separable prefixes: 7.57% of the errors because MT cannot handle prefix at the end

Optional to extend the validation side: you may consider briefly positioning EuroCrops alongside global gridded crop-area products such as CROPGRIDS (173 crops circa 2020) as complementary yet solid benchmarks at coarser resolution (useful for sanity-checking aggregated totals, not as parcel-level substitutes)

<https://www.nature.com/articles/s41597-024-03247-7>

<https://openknowledge.fao.org/items/ebc60e53-2b29-4173-b1fa-2320669b1312>

Overall, I see these as documentation and reproducibility refinements rather than methodological blockers

Thank you for this suggestion. However, the purpose was not to evaluate the EurocropsV2 dataset against all potential crop datasets. We compared EuroCropsV2 with what is considered to be the most robust reference, i.e., crop statistics from Eurostat and gridded agricultural census data from the Integrated Farm Statistics (IFS) as these were available at a finer level of granularity than Eurostat NUTS2 level statistics. However, we have added a reference to this paper in the discussion section, citing it as a source for users to access if they want to do a comparison with other datasets.

Reviewer 2

This manuscript presents EuroCrops v2.0, a multi-annual harmonized dataset of parcel-level crop declarations derived from publicly available Geo-Spatial Application (GSA) data across 18 Paying Agencies in the European Union. The dataset is harmonized using an updated Hierarchical Crop and Agriculture Taxonomy (HCAT4), which introduces additional dimensions for seasonality and usage, and is linked to multiple European statistical and Earth Observation classification systems (AGRIPROD, LUCAS, IFS/FSS, FADN, HRL Crop Types). The manuscript fits well within the scope of *Earth System Science Data*. It describes a large, policy-relevant, open-access geospatial dataset with clear potential for reuse in Earth observation, agricultural monitoring, sustainability assessment, and CAP-related evaluation. The extension from a single-year dataset (EuroCrops v1) to a multi-annual harmonized dataset represents a significant and valuable advance. The manuscript is generally well structured, and the authors demonstrate considerable effort in data cleaning, semantic harmonization, and validation. The comparison with Eurostat statistics and gridded IFS data provides useful validation evidence. The discussion appropriately acknowledges limitations.

Thank you for the positive comments about the paper.

However, several aspects require clarification or strengthening, particularly regarding reproducibility of the harmonization workflow, quantification of methodological impacts (e.g., stacking and rasterization), and documentation of uncertainty and metadata.

Major Comments

Handling of Multiple Crops per Parcel

The manuscript describes rules for handling multiple crops listed within a single parcel label (e.g., hierarchy-based grouping, majority crop assignment, first-mentioned crop). However:

The proportion of parcels affected by multi-crop labels is not reported.

The potential bias introduced by assigning the “first mentioned” crop is not assessed.

Country-level differences in this issue are not quantified. Given that some countries (e.g., Portugal, Ireland) exhibit specific declaration complexities, this could introduce systematic distortions.

There are two different issues regarding multiple crops per parcel: (a) multiple crops declared per parcel due to double cropping, intercropping or further subdivision of the parcel into smaller subparcels, and (b) multiple overlapping parcels in the GSA with different crops. In the manuscript, we emphasize the need to process the data set in order to have a single declaration per parcel, which is a decision that we have made.

Regarding the first issue, Portugal allows farmers to declare multiple crops (up to 12 in 2022). The EuroCropsv.2.0 dataset includes only the crop declared in the column C1 of their GSA. In the case of Austria, if two main crops are cropped as a mixture and harvested the same year, they are both declared, separated by a slash but this will never include cover crops. Original codes and names (including the two crops) are preserved in the dataset. The hcat4 is linked to the first mentioned crop. We have added a sentence to the limitations section of the paper to acknowledge this situation.

Regarding the second issue, duplicated parcels can occur, where the reasons vary by country and this can change over time. The main country affected by this issue was Ireland, e.g., a parcel can have multiple owners so is declared multiple times. We made the decision to have only one parcel geometry in the database so we removed duplicated and overlapping parcels.

2. Reproducibility of the Harmonization Workflow

The harmonization process is conceptually well described, including translation, semantic verification, matching to HCAT4, and iterative refinement. However, several components lack sufficient detail for full reproducibility. The manuscript indicates that crop names were translated using Google Translate, DeepL, OPUS-MT, and, where necessary, ChatGPT. Additionally, a large language model (LLM) was used to detect semantic discrepancies in crop code definitions over time. For ESSD, automated or semi-automated semantic decisions that affect taxonomy assignment should be reproducible. The manuscript should clarify:

Which LLM(s) and versions were used?

Were deterministic settings applied?

What prompts or criteria were used to detect semantic discrepancies?

How were disagreements between translation tools resolved?

What proportion of entries required manual intervention?

Without these details, it is difficult for future users to replicate or audit the harmonization procedure.

Reviewer number 1 had a similar query about the translations, which we answered as follows and also addresses the questions above:

Before the crop names were translated from their native language to English, a pre-processing step was undertaken to align the crop names across the years because countries provide an updated crop code list on an annual basis and the EuroCropsv2.0 process has been ongoing since 2021. The github now includes a table with the original_code and how these are related to multiple instances of original_name. The process for consolidating the dataset (i.e., having unique pairs of original_code/original_name) involved the use of an LLM (gpt-4o and nous-hermes-2-mixtral-8x7b-dpo), or a manual process. The source of the consolidated original_name is provided in this table here:

https://github.com/Martincccc/EuroCropsV2/blob/main/data/processing/name_mistakes.csv.

The aligned crop names were then translated using three machine learning packages. We have uploaded one table per country (XX_trans_stats.csv) with the original codes and their translations from Google Translate (g_tname), DeepL (dpl_name) and OPUS-MT (opus_tname) to github, placed in this folder:

https://github.com/Martincccc/EuroCropsV2/tree/main/data/eurocrop_trans, where XX in the filename is the two letter country code or the three letter regional code. Note that these files have been updated in 2026 so that they are reproducible since this process was done over several years (since 2021) as outlined above.

We attempted to provide the best translations possible, which is why a third machine translator was introduced as part of EuroCropsv2.0. At the beginning of EuroCropsv2.0, we discovered some wrong translations by DeepL and Google Translate. Therefore, we added OPUS-MT, which is not as good as the other two translators, accounting for 63-92% of the translation errors, but it served as an additional check. Without the use of OPUS-MT, the number of manual checks would have been reduced significantly, but we ran the risk of wrong translations that would not have been picked up using only DeepL and Google Translate.

The translation accuracy varies by country with the lowest in Spain at 40% to the highest in Belgium at around 72%. Estonia provides English translations of most labels on <https://klassifikaatorid.stat.ee/>, so a simple search there resolved most disagreements. Other countries do not have such convenient portals. Slovenia releases their labels along with the scientific name of the crops while Ireland's labels are in English. Spelling mistakes were rare except for the Spanish, Bulgarian, Czech, Portuguese and Italian GSA, which led to direct failures of translation and no results were provided by Google Translate. In the German GSA, there were many abbreviations and short forms of names used, which hindered the automatic translation. In the GSA of Czechia, there were extremely detailed species level crop descriptions of grasses and legumes as well as the use of colloquial terms. After machine translation, colloquial terms often led to non-crop-like descriptions such as diabetes, quiet, etc. Subsequent Google searches of these terms sometimes turned up empty, whereas for other niche translations, they showed up in Wikipedia. In other cases, we sent the colloquial terms to native speakers to review but this was not possible for all languages.

We have added a new section to the Supplementary Material (called Translation of crop names) and added all of this text. It includes a summary table, Table S1, which summarizes the number and percentage of multiple instances of original_name per country and the translation statistics and errors. The accuracy refers to the percentage of time when all the machine translations were in agreement. Where there was disagreement in the remaining translations, these were all manually checked. Of these manually checked translations, five different cases may have occurred. In the majority of cases, at least one machine translator was correct. Often OPUS-MT was the poorest performer, but other times, e.g., in Slovakia, Google Translate and DeepL failed but OPUS-MT was correct. The second type of case included the occurrence of niche crop names, e.g., species level trees or legumes, which were confirmed by Wikipedia, or where the translator chose a more frequently used homonym, which were then checked using Wikipedia's disambiguation pages. A third case was the occurrence of colloquial or regional names, which were also checked using Wikipedia. Finally, spelling errors (case 4) and crop names that were abbreviated or truncated (case 5) were also identified through this manual checking process.

Table S1: Summary of instances of multiple original_name for crops and translation statistics

Country or Region	# of multiple original_name [%]	Accuracy	Manual Correction	Poor Machine Translation	Niche Crop/Homonymy	Colloquial/Regional name	Spelling errors	Abbreviation/Truncation
AT	13 [5%]	60.98	39.02	82.29	6.25	1.04	0	10.42
BE2	52 [14%]	71.82	28.18	92.16	6.86	0.98	0	0
BE3	75 [36%]	71.29	28.71	88.33	11.67	0	0	0
BG	-	58.3	41.7	72.57	24.78	1.77	0.88	0
CZ	146 [35%]	45.08	54.92	62.56	33.04	0.88	2.64	0
DE4	89 [30%]	67.69	32.31	77.13	12.77	0	0	10.11
DEA	51 [18%]							
DK	286 [56%]	63.78	36.22	66.3	27.72	0	0	5.98
EE*	-	43.77	55.23	76.14	16.3	0	0	0
ES	-	40.29	59.71	69.14	26.34	2.47	1.65	0.41
FI	-	57.2	42.8	85.71	14.29	0	0	0
FR	-	64.91	35.09	84.87	15.13	0	0	0
IE	-	-	-	-	-	-	-	-
ITI1	1 [0%]	59.35	40.65	75.51	23.81	0	0.68	0
NL	91 [18%]	76.42	23.58	87.93	8.62	0	0	3.45
PT	-	56.73	43.27	77.78	14.81	1.48	5.19	0.74
SI	8 [4%]	-	-	-	-	-	-	-
SK	-	63.48	36.52	88.35	9.71	1.94	0	0

*EE has separable prefixes: 7.57% of the errors occurred because the machine translation could not handle the prefix at the end

3. Uncertainty and Quality Indicators

The manuscript presents validation against Eurostat and IFS data, including R^2 and NSE metrics. This is commendable and strengthens confidence in the dataset. However, the dataset itself does not appear to provide:

Harmonization confidence indicators,

Translation reliability scores,

Flags for manually corrected entries,

Country-level reliability summaries.

Given the complexity of semantic harmonization across languages and classification systems, some form of quality metadata would enhance transparency.

We have addressed some of the requests outlined above through new tables that have been uploaded to github and put into the paper (Table 1, Table S1), tackling translation reliability and manually corrected entries. In particular, see the responses provided to reviewer 1 for more details. Harmonization confidence indicators and other country-level reliability summaries are not provided as it is unclear how these would be calculated. If the reviewer has suggestions, we would be happy to consider these.

4. Metadata, FAIR Compliance, and Versioning

The dataset is openly available (with DOI), and code is provided via GitHub. This is highly positive. However, the manuscript would benefit from:

A clear data dictionary describing all attributes in the geoparquet files,

Explicit CRS documentation,

Description of file naming conventions,

A versioning and update policy (e.g., semantic versioning, expected update frequency),

Clarification on long-term repository archiving (e.g., Zenodo snapshot of code).

The dataset structure was already provided in the supplementary material, which describes all the attributes in the geoparquet files. This information has also been added to the JRC data catalog and is also in the [README.md](#) file on the github repository.

This readme file also contains a description of the file naming convention of the yearly GSA file and the stack files.

Regarding the comment on providing explicit CRS documentation, see the response below about the CRS.

There is currently no follow up intended to EuroCrops V2 so this will be a snapshot. However, the data and code are available on Github and this is a community collaboration so updates may be driven by the needs of the community in the future.

Regarding clarification on long-term repository archiving, the current version of the data will be maintained in the JRC catalogue and ftp site similar to other long-term repositories like Zenodo.

Minor Comments

Abstract: Consider explicitly stating the temporal coverage range (e.g., 2008–2023, where applicable) for clarity.

We have added the temporal range explicitly to the abstract.

Consistency of terminology: Ensure consistent referencing of HCAT versions (HCAT2, HCAT3, HCAT4).

We have ensured consistent referencing of the version of HCAT.

CRS specification: Explicitly define the coordinate reference systems used for storage and visualization.

All the original GSA data and the EuroCrops V2 dataset have the same coordinate reference system (CRS), which is the Lambert Azimuthal Equal Area (LAEA) coordinate system using the European Terrestrial Reference System 1989 (ETRS89). The EPSG code for this coordinate system is EPSG:3035. More details can be found at <http://epsg.io/3035>. We have added this detail to the paper although this information was already provided in the Supplementary Material in the section called **Data set structure**.

Validation summary: A concise summary table of validation statistics per crop class would enhance readability.

We have added a summary table of validation statistics per crop class for a selection of crops in Table 3 and provided statistics for the full set of crops in Table S2 of the Supplementary Material, referring to this in the main text.

Additional note about reprocessing

Due to an inconsistency in the algorithm, the data set has been reprocessed as follows: In the previous version, priority was given to the largest parcel but in Spain, there are large parcels with codes containing “unknown” and small overlapping parcels with crops listed. Hence the algorithm was changed so that priority was given to the largest parcel only if it contained an actual crop in the HCAT4 taxonomy. This correction has been applied across all countries and regions and released as V2.01.

Following this correction, additional issues were identified in the stack-layer processing. In the previous version, these processing errors resulted in duplicated crop fields in the output. The algorithm has therefore been further revised to remove such duplications. In parallel, enhancements to the scripting logic have allowed for a reduction of the buffer zone between output polygons, leading to improved spatial precision.

The geoparquet files are released as part of the distribution v2.01 of the dataset in the JRC data catalog (<https://data.jrc.ec.europa.eu/dataset/b9fb9e67-78a9-4327-9d59-39a928d812d3>).