

# WorldCereal: a dynamic open-source system for global-scale, seasonal, and reproducible crop and irrigation mapping

Kristof Van Tricht, Jeroen Degerickx, Sven Gilliams, Daniele Zanaga, Marjorie Battude, Alex Grosu, Joost Brombacher, Myroslava Lesiv, Juan Carlos Laso Bayas, Santosh Karanam, Steffen Fritz, Inbal Becker-Reshef, Belén Franch, Bertran Mollà-Bononad, Hendrik Boogaard, Arun Kumar Pratihast, Benjamin Koetz, Zoltan Szantoi

## Response to reviewers

### Reviewer Comment 1

<https://doi.org/10.5194/essd-2023-184-AC1>

**I have read the preprint with great interest, and I immediately realized the amount of work involved in producing the WorldCereal data product. At the same time, however, I am not an expert in this type of data processing, and I may not be able to provide an expert opinion on the workflow and methods used during development. While I am sorry for this, I still wanted to provide some general feedback, which I detail below.**

**1) I found that overall description of methods would be more accessible if there were a few firm formulations, at list for the basic steps undertaken to produce calculations and assess the main crop distribution. Additionally, and in support to this, a schematics of the workflow would be very useful.**

We agree with the reviewer that a flowchart of the WorldCereal production workflow would greatly improve the readability and ease of interpretation of the methods section. We have therefore added a new figure (Figure 1) showing the generic workflow which allows generation of the different products, combining processing task definition, pre-processing, feature computation and classification. Additionally, we extended the introduction contained within the methods section to clarify the structure of the section and explain the linkages between the different sub-sections.

The new figure has been added as a supplement to this comment.

*Figure 1: WorldCereal production flowchart detailing the steps required to generate the WorldCereal products for a user-defined area and year of interest. This schematic does not include training of the classification model (detector).*

Text added at the start of Section 3:

*Figure 1 illustrates how product generation is achieved by the WorldCereal classification system starting from a user-defined area and year of interest. The number and exact timing of the maize and cereals seasons to be processed are derived from the global crop calendars (Sect. 2.2), which were stratified into uniform zones to enable fast processing at large spatial scales (agro-ecological zones, see Sect. 3.3). Based on this information, the appropriate time series of raw EO data are extracted (see also Figure 6) and pre-processed for each individual growing season (detailed processing steps further specified in Sect. 3.2). Next, the prepared inputs are condensed into product-specific sets of classification features (Sect. 3.4), which directly feed into the respective classification models for temporary crops, maize, cereals and irrigation detection (Sect. 3.5). Following model inference, some postprocessing steps were applied to enhance individual product quality and inter-product consistency (Sect. 3.6). Note that Figure 1 does not include the training aspect of the different classification models applied. Sect. 3.1 further details the reference data used for training our global models, whereas Sect. 3.5 describes the model architecture and training procedure. Finally, Sect. 3.7 specifies how the WorldCereal products were validated.*

**2) Some difficulties in image recognition when plants/crops show stress may arise from nutrient availability rather than water scarcity, or in conjunction with water scarcity. In fact, nutrient management (e.g., fertilizers) may need to be included as covariates for crop recognition.**

Thank you for your suggestion. We agree that there are multiple environmental and management factors which could have a large impact on how a certain crop is behaving throughout the growing season, giving rise to considerable variability in its temporal profile as picked up by the satellite sensors and in turn complicating crop type recognition. Due to our specific model training approach, i.e. combining data from multiple years and multiple regions into a single model, this potential variability is, for a large part, inherently included in our training data. Although adding more covariates to the model should in theory help boost its performance, the main problem here is that we simply do not have the required information on e.g. nutrient management across the globe. As our models need to be able to cover the entire globe, we therefore opted to account for potential variability in crop type signals by gathering sufficient training data originating from a diverse collection of locations with significantly different environmental and management conditions. Note that this also means that if a certain year or region you want to map is behaving completely different compared to what has been covered by our training

data, new reference data collection should be organized in order to cover these changing conditions.

### **3) Definition of irrigation was not very clear (Section 2.4).**

We understand the definition of irrigation is not very clear. We now have addressed this issue in the manuscript by emphasizing that the WorldCereal system is mostly trained using data from (semi-)arid climates and that therefore the focus of the irrigated area mask is on extensive irrigation practices (in volume and in duration) which are needed to ensure significant crop growth. Since we did not have access to training data in more temperate climates, the model is less suited to detect incidental irrigation which is only applied during the start of the growing season or irrigation that is applied during a short dry spell.

The definition now reads:

*In the WorldCereal system, therefore, areas are defined as irrigated agriculture if only due to extensive irrigation over a prolonged period of time significant crop yield can be reached. Other types of irrigation, such as incidental irrigation during the sowing period or during short term droughts, are not the main focus of the irrigated area mask. This primarily excludes irrigation in more temperate climates, where irrigation is mostly applied to enhance crop yield instead of preventing crop failure.*

**4) In general, I also could not fully understand why it was required to define active and non-active crops. In fact, providing, seasonal maps of main crops, wouldn't it already imply that time slices indicate already where a crop is and where is not in that specific time? I may understand that, within the workflow proposed, tagging active and non-active crops may be useful for calculation. And, is this information distributed too? Yet, eventually, what is the scope of the information that is passed to the user of the data product by means of "active/non-active" data?**

The definition of active cropland maps, along with the reason why such active cropland maps would be useful to a user, are detailed in section 2.5 of the manuscript. To monitor global food security, it is important to know the percentage of temporary cropland being effectively cultivated and harvested in a given growing season. As we have seen recently, external factors such as extreme climatic events, a global pandemic or armed conflicts can cause farmers to abandon their fields and complete harvests to fail, putting a lot of pressure on local to global food systems. By producing these active cropland maps, we thus provide helpful insights related to food production and land abandonment, particularly in areas affected by such factors. Important to keep in mind here is that our seasonal crop type maps currently only cover maize and cereals (wheat, barley and rye) and certainly not all major crop types across the globe. We agree with the reviewer that if we were to cover many more crop types, the significance of this active cropland layer would

decrease. Nevertheless, we will never be able to cover all crops grown in a certain region so we still believe this concept of an active cropland marker holds value on the longer term. We have added a sentence in section 2.5 to better justify this:

*External pressures such as natural disasters, global pandemics and armed conflicts may cause farmers to abandon their fields or complete harvests to fail, thereby significantly impacting local to global food security. To gain a better understanding on local food production, the WorldCereal active cropland products indicates...*

**5) Confidence in Eq. 1 may not be a confidence, but rather a quality index of the training. However, the data quality maps distributed as “\_confidence” would rather better report quality against validation, or a combination of confidence and validation, so that an overall quality index is available and distributed as maps.**

The confidence maps presented and distributed in the WorldCereal system are directly derived from the prediction probabilities as provided by the CatBoost binary classification models and as such reflect how certain the model is of its prediction, based on what it has learned from the training data. It simply provides a user with an idea of how difficult it was for our trained model to classify the pixel in one or the other class, or better yet, how different the pixel was behaving with regard to the training data seen by the model. The reviewer is correct that it does currently not relate to the quality of the product in terms of an independent validation. Due to the spatially biased distribution of validation data, particularly for the crop type and irrigation products, we deem it at this point not feasible to include a confidence score with regard to validation at 10 m resolution and global scale. Therefore, we suggest to replace “confidence” by “model confidence” throughout the entire manuscript to make this more clear and have added an additional sentence in section 3.5 to better explain this:

*“Note that this model confidence score simply reflects how certain the model is of its prediction, based on what it has learned from the training data and does not reflect real accuracy based on independent validation data.”*

**6) Overall, and as a non-expert in the methods deployed, I would rather provide a feedback to the authors from the perspective of a potential user of the data. In particular, I would like to suggest to add a short section, “Data usage”, that provides details on how to use these data. For example, the area fraction cultivated with maize in a grid cell is quite straightforward to use. However, for irrigation, I would rather be interested in how many cubic meters per hectare and per day are used and in which period, but if I understand well, this is not the information provided with the data layers marked with “irrigation”.**

Thank you for this valuable feedback. We have added a small section on “Data usage and future prospects”, providing examples on how the generated classification maps can be used in downstream applications. Important to note here is that the low resolution fraction maps we are showing are purely for visualization purposes and do not represent proper area statistics. We refer to Olofsson et al. (2014) for guidelines on how to derive regional statistics from pixel-based classification maps.

Olofsson, P., Foody, G.M., Herold, M., Stehman, S.V., Woodcock, C.E., Wulder, M.A.: Good practices for estimating area and assessing accuracy of land change, *Remote Sens. Environ.*, 148, 42-57, <https://doi.org/10.1016/j.rse.2014.02.015>, 2014.

The part of the new section that answers the reviewer’s comment now reads:

*The WorldCereal project has generated a suite of binary classification maps at 10 m resolution and global scale for the year 2021, which can act as an important starting point towards a dynamic (seasonal) global-scale crop and irrigation monitoring framework (See et al., in press). The maps can in the first place be used at their native resolution to identify hotspots of temporary crop/cereals/maize production and irrigation practices at regional scale, in turn allowing better planning of agricultural field data collection campaigns and improving our understanding of local cultivation practices. Additionally, the data can be downscaled and as such prove useful to enhance and complement subnational to national agricultural/water use statistics (e.g. FAOSTAT and AQUASTAT, both produced by FAO). Important to note here is that our low resolution fraction maps (Figures 7, 8 and 9) have been generated purely for visualization purposes and do not represent proper area statistics. We refer to Olofsson et al. (2014) for detailed guidelines on deriving regional statistics from pixel-based classification maps. The WorldCereal products are being evaluated for integration in various platforms dealing with food security and agricultural water management, including GEOGLAM’s CropMonitor (Becker-Reshef et al., 2023), FAO GIEWS and FAO WaPOR database, where the products will contribute towards improved frequent crop condition reporting, crop production/failure early warning and forecasting, crop-specific assessment of impacts of extreme weather events and agricultural policy changes on planted area and production, and season-specific irrigation monitoring.*

*See, L. Gilliams, S. Conchedda, G., Degerickx, J., Van Tricht, K., Fritz, S., Lesiv, M., Laso Bayas, J.C., Rosero, J., Tubiello, F., Szantoi, Z.: Realizing a vision for dynamic global-scale crop and irrigation monitoring, *Nature Food Commentary*, in press.*

*Becker-Reshef, I., Barker, B., Whitcraft, A., Oliva, P., Mobley, K., Justice, C., Sahajpal, R.: Crop Type Maps for Operational Global Agricultural Monitoring, *Sci Data* 10, 172, <https://doi.org/10.1038/s41597-023-02047-9>, 2023.*

With regard to estimating quantities of water used for irrigation, the reviewer is correct that this is well beyond the scope of our study. To detect irrigated areas

from space, we rely on features that are sensitive to irrigation in a physiological sense (precipitation deficit, evapotranspiration, soil moisture content) and in a spectral sense (various indices that rely on distinct spectral bands, such as the NDVI, NDWI, etc.). Combining these features helps locating irrigation practices, but does not indicate how much water is actually used for irrigation. In areas where more water is used for irrigation, the model tends to be a bit more confident, but understanding how much water is actually used for irrigation requires more in-depth knowledge on the water fluxes of each pixel. A lot of research is conducted on this topic at a smaller catchment/region level, mainly focusing on remote sensing based soil moisture estimates in combination with a land surface model, or using energy balance models that use remote sensing data and meteorological models to accurately estimate the amount of water that is consumed by the crop. Assessment of evapotranspiration from remote sensing data at larger spatial scales is for instance done in the FAO WaPOR project (<https://www.fao.org/in-action/remote-sensing-for-water-productivity/wapor-data/en>).

## **7) Finally, would it be possible to extend the methods to other crops?**

Our approach for crop type identification can definitely be extended to other crops. This would require (1) sufficient high-quality reference data for the crop of interest in the regions where the crop should be mapped and (2) knowledge on the timing of the growing season(s) during which the crop of interest is cultivated for all regions of interest. We have already tested this in a case study in Ukraine where we additionally mapped sunflower and rapeseed and are planning to extend this effort also to the global scale. We added two sentences to the newly created “Data usage and future prospects” section to clarify this:

*Although WorldCereal specifically focused on maize and cereals, the crop type identification system presented here represents a generic framework for crop type mapping and can be easily extended towards other crop types. To do so, one would require (1) high-quality reference data for the crop of interest covering all regions of interest and (2) knowledge about the timing of the growing season(s) in which the crop is cultivated (cf. crop calendars and agro-ecological zones as presented in Sect. 2.2 and 3.3 respectively). Our harmonized in-situ reference database (Sect. 3.1; Boogaard et al., 2023) already contains data on many other crop types and can serve as a starting point, which can be further complemented by user-provided reference data.*

## Reviewer Comment 2

<https://doi.org/10.5194/essd-2023-184-AC2>

Firstly, I want to highlight and applaud the immense amount of work that went into the creation of the system presented in this paper. Readers not experienced in agriculture mapping or at least land cover mapping may not appreciate how much time, thought, computational resource, and effort from many diverse people are required to create the products detailed in this paper.

Overall, I thought the manuscript was well written and easy to read and understand. There are some minor comments/suggestions detailed below, most of which are related to method clarifications or drawing lasting findings and observations from the experience of creating the products than the products themselves.

In the Introduction, the authors mention that there are other global land cover maps that include a cropland class as well as GCEP30 and GLAD maps that focus only on mapping cropland globally. The authors state that these products have at least one of a list of limitations, one of which is “they have limited local applicability in areas with less training data”. This seems to be a limitation of WorldCereal too. There is no discussion of how WorldCereal’s cropland extent product compares to other available maps, which leaves users wondering if the WorldCereal cropland extent map might be better for their use case than the existing ones or not. At minimum, a discussion of the reported accuracies in the GCEP and GLAD maps compared to those reported for WorldCereal would be helpful (or a discussion of how users might decide to use one vs another).

Thank you for the suggestion. This was indeed missing in our manuscript. In order to address this issue we added a small discussion in section 5 (Product validation) to (1) compare the accuracy numbers of the WorldCereal cropland product with other recent global cropland layers and (2) to comment on the application specific appropriateness of the product, also in relation to other global cropland layers:

*It has been shown that global single cropland maps tend to have a much higher accuracy than the cropland class of land cover maps. For example, the cropland class of WorldCover has a user’s accuracy of 80.6 (+/- 1.5) and a producer’s accuracy of 79.3 (+/- 1.5) compared to 88.5% (+/- 0.5) and 92.1% (+/- 0.4) respectively for WorldCereal. The single layer map from the University of Maryland (UMD croplands) has comparable accuracy numbers as the WorldCereal map with an overall accuracy of 97.2 (+/- 0.3) and a user’s and producer’s accuracy of 88.5% and 86.4% (+/-1.9). In contrast, the older single*

*layer GCEP product has substantially lower accuracies: 91.7 % overall, 78.3 % user's and 83.4 % producer's accuracy. We can therefore conclude that from a global perspective the most recent global cropland maps (WorldCereal and UMD croplands) are both very high quality products. Selection of the most appropriate product to use for a given application will both depend on the nature of the application (different products adopt slightly different definitions of cropland / seasonal versus multi-year products) and the region (one product might have had more/better quality training data for a particular region compared to others). The use of the WorldCereal temporary crops product would be more appropriate for applications interested in active croplands for the specific growing seasons ending in 2021 and would benefit from the increased spatial resolution (10 m versus 30 m), whereas UMD croplands might be better suited for applications which need to also include the fallow class and consider a longer period of time (stable cropland area).*

**Lines 74-75 state that maps “can be locally finetuned if users add their own training data,” but there is no discussion on how this could be done. Can the authors add information about how users could finetune the models/maps, or otherwise rephrase the statement?**

While our pixel-based Catboost models do support local finetuning based on additional training data, the current WorldCereal system is mainly tuned towards the production of global maps and as such is not easily customizable by users to train new models or finetune existing models. We are therefore integrating the WorldCereal system into openEO platform, which will simplify this entire workflow and introduce more flexibility to the user. This will include the possibility for a user to apply the existing models to another year (allowing change detection applications), add independent training data, combine these data with relevant data from our public reference database and train new models for custom regions, growing seasons and crop types. We have added a statement in the new Section 6 on “Data usage and future prospects” to clarify this:

*The WorldCereal system will be integrated in the openEO platform processing environment, allowing any user to easily interact with the system and launch customized model training and processing tasks for specific years, growing seasons, locations and crop types based on public and user-provided reference data. Users will have the opportunity to use the existing trained models or train dedicated models for their application. We advise to add application-specific training data to the system in order to guarantee high-quality outputs, especially when environmental conditions in the area and period of interest highly differ from the conditions currently captured by the available reference data. Fully opening up the system to the broader user community in this way will (1) allow for a continuous expansion of harmonized crop type and irrigation reference data, (2) improvement of the products based on local user knowledge and provided training data, (3) ensure the system can meet the (changing) needs of the*



community and (4) allow for new applications of the generated global products (e.g. serving as a baseline for change detection analysis at regional to global scales).

**It's not clear if the vision of the dynamic open-source WorldCereal system is that any user would run the system, or that the authors would be continuously funded to run the system and produce the associated products (lines 71-81). Can the authors clarify this?**

WorldCereal is in the first place an ESA-funded project which aimed to demonstrate the feasibility of global crop (type) and irrigation mapping at 10 m resolution, without any operational mandate. The WorldCereal consortium hence does not receive continuous funding to run the system operationally at global scale. As stated in our answer to the previous question, the idea is that the system itself is fully open-source, allowing any user to run customized training and/or processing tasks.

**In Section 2.2-2.3, it might be helpful to include a small figure illustrating the timeline of the crop seasons that are used by the system.**

Thank you for this suggestion to improve interpretability of this section. Obviously the timeline of the growing seasons is different depending on the AEZ. We have therefore included a new figure (Figure 6) showing the extent of the growing seasons in two significantly different AEZ, i.e. one AEZ with two maize seasons and one AEZ where the distinction is made between winter and spring cereals. We refer to this new figure in several sections throughout the manuscript to provide the necessary clarifications. To support the new Figure 6, the AEZ used in this Figure were identified in Figure 3 (Figure 2 in previous version) and product abbreviations were defined in Table 4.

The new and updated figures are attached to this comment as a supplement.

*Figure 6: Demonstration of WorldCereal 2021 product generation and timing for two distinctive AEZ's. (a) AEZ located in central U.S., where only one maize season occurs and spring cereals are mapped jointly with maize and (b) AEZ located in Somalia, where two maize seasons occur and no spring cereals are mapped. Product abbreviations are explained in Table 4, whereas locations of AEZ are highlighted in Figure 3.*

**Lines 346-347 states that the goal "was to train generalized models across multiple years that do not specifically require new training data in unseen years". However, the validation was performed for only 2021 which was a seen year (included in the training data). The results do not show that new training data is not required for unseen years. Do the authors have some other experiments to demonstrate that the model does generalize well to unseen years? When the system is run in future years, will more training data need to be added each year?**

Thank you for this excellent remark. We have indeed conducted specific experiments in Ukraine where we have generated crop extent, crop type and irrigation maps for the years 2018-2021. These maps were validated based on independent datasets provided by the National Technical University of Ukraine, Kyiv Polytechnic Institute (NTUU "KPI"). The accuracy of the maps has been shown to remain relatively stable across all years, despite the fact that we only had (limited) training data for 2018-2019 originating from this particular country. Overall accuracies and ranges of user's and producer's accuracies for the different years amounted to: 92.5%, 85.5-97.9%, 84.2-99.6% (2018); 92.7%, 88.9-97.9%, 86.9-99.3% (2019), 85%, 78.5-94.7%, 64.5-99% (2020), 93.9%, 88.2-99.3%, 86.1-99.5% (2021). Validation of the year 2022 (during which we additionally mapped sunflower and rapeseed) showed similar results, and no data for 2022 was used in training: OA 91.5%, UA range 86.5-99.7%, PA range 65.5-99%. Recognizing that this is a demonstration on one specific country it still demonstrates the power of the system. A separate publication on this topic is under preparation. In general, we do recognize that having access to year-specific data for training will always result in the best mapping accuracies. In case such data is not available, the accuracy of the maps for the new year would highly depend on how the environmental and management conditions differ from those already covered in the training database. For instance, in case of exceptional drought, the capability of our models to correctly identify cereals and maize will likely decrease. It will be up to the user of the WorldCereal system to identify the need for new training data depending on the targeted application and the available reference data in our harmonized reference database.

We have added the following paragraph in Section 5 on crop type product validation:

*In order to demonstrate the temporal robustness of the WorldCereal models an additional validation effort was done for Ukraine based on an independent dataset obtained from the the National Technical University of Ukraine, Kyiv Polytechnic Institute (Kussul et al., in preparation). Country-wide crop type maps were generated for the period 2018-2021. Overall, users's and producer's accuracies were found to remain stable across the years (OA of 92.5%, 92.7%, 85% and 93.9% respectively), despite only limited training data (2018-2019) being available for this particular country.*

*Kussul, N. et al.: Cropland maps validation for Ukraine. In preparation.*

We have also added a sentence in the new Section 6 to comment on the need for users to add application-specific training data:

*We advise to add application-specific training data to the system in order to guarantee high-quality outputs, especially when environmental conditions in the area and period of interest highly differ from the conditions currently captured by the available reference*

*data.*

**Line 190: what is meant by “an equivalent number of looks of 3”? Equivalent to what?**

The equivalent number of looks (ENL) is a standard term used in multilook synthetic aperture radar (SAR) image processing and refers to the degree of averaging applied to the SAR measurements to get rid of speckle (noise) in the resulting images.

**Lines 194-197: was GDD normalization not applied to the Landsat thermal data for crop type? Can you clarify what was done for the 16-day median filter, since Landsat observations are 16-day frequency?**

Landsat thermal features have only been used in the irrigation model, not for training crop extent nor crop type models (cf. Tables 1 and 2). GDD normalization was not deemed essential nor practically feasible for irrigation mapping as this is a crop type a-specific mapping procedure. GDD normalization however requires crop type specific parameters to be set (minimum and maximum temperature boundaries defining the temperature range in which the crop is actively growing).

The 16-day median filter was applied to ensure a consistent time series (only one observation within each 16-day period was retained). We agree that the impact of this operation is limited and restricted to regions which potentially show overlap between different acquisitions.

**Line 200: Why not 10m to align with Sentinel-2 tile grid, especially since it’s later resampled to 10m anyway?**

In order to optimize both memory and computational efficiency of our workflows (allowing good scalability), all data sources are processed at their native resolution as far down the pipeline as possible. For instance, Sentinel-2 20 m bands are processed separately from the 10 m bands and the resulting features are resampled and merged at the very end. The same is true for the Copernicus DEM: elevation and slope features are derived at a resolution of 20 m and are resampled to 10 m at the very end of the feature computation workflow.

**Line 233: there is a reference error.**

Thank you for noticing, this has been corrected in the manuscript.

**Line 249: Can you clarify what is meant by “equidistant time series descriptors”? Does this mean spread evenly throughout the time series length, or at specific intervals?**

We indeed mean “spread evenly throughout the time series length”. We have adjusted the sentence to make this more clear:

*For NDVI in specific, the temporal profile was captured in more detail by sampling the time series at six positions spread evenly throughout its length (resulting in six additional features, ts0-ts5) and by computing 12 of the temporal features based on the work by Valero et al. (2016).*

**Line 278: How was the random perturbation degree chosen? Is the model performance sensitive to this parameter? It seems surprisingly large. Are the lat/long variables input as continuous or categorical features to the model? It seems that continuous features could cause the model to learn spurious patterns (or prevent it from learning useful patterns) since the coordinates are rotational and do not wrap around.**

The random perturbation degree of lat/lon coordinates was empirically determined in a sensitivity analysis where a trade-off was found between (a) a model that is too sensitive to knowing exact locations (limited to no perturbation), which resulted in artefacts, especially in regions with considerable lack of training data, and (b) a model that no longer uses the localization information (too much perturbation). We agree with the reviewer that the resulting degree of perturbation is somewhat large. This could be diminished in future versions in case more training data would become available, especially over data-sparse regions.

With regard to the implementation of the lat/lon coordinates themselves we did not experience this as a limitation and currently do not expect a large impact on product quality as we have relatively little data near the extreme positive or negative coordinates. Nevertheless, we do believe the reviewer made a valuable point here. To address this, a future version of the system could adopt the use of a cartesian coordinate system, as demonstrated in Tseng et al. (2022).

Tseng, G., Kerner, H., Rolnick, D.: TIML: Task-Informed Meta-Learning for Agriculture, arXiv:2202.02124, <https://doi.org/10.48550/arXiv.2202.02124>, 2022.

**Line 430: Why was CGLS chosen instead of something else like WorldCover for filtering tiles? We have found CGLS to underestimate crop extent in previous studies (e.g. <https://arxiv.org/pdf/2006.16866.pdf>). Could this lead to more tiles being filtered out of the product than should be?**

The step for filtering tiles that do not require any processing by the WorldCereal system was mainly done to avoid inclusion of large zones where crop growth is virtually impossible (deserts) or can be safely ruled out (middle of tropical rainforests, northern latitudes in Canada and Russia). Using a 10m resolution product to perform this filtering would have required quite some processing and was deemed unnecessary for the targeted application. At the time when this

filtering operation was performed, the WorldCover product was either not available yet at global scale or showed similar issues as the CGLS product (note that quite some substantial improvements to the WorldCover product were introduced between the first and second version). We instead chose the CGLS product as we at VITO are very familiar with this product and its limitations. Given these known limitations over e.g. Africa, we were very conservative in removing tiles from being processed by the WorldCereal system, as described in Section 4.

**Line 374: Since the authors were following the best practices in Stehman & Foody 2019 for accuracy assessment, why did they use a different technique than in that paper to compute the confidence intervals?**

We used a different method to calculate confidence intervals than the one proposed by Stehman & Foody 2019 called bootstrapping, because it does not require the assumption of a normal distribution. The method has been applied recently in Szantoi et al. (2021) and is based on Schreuder et al. (2004), Gallaun et al. (2015) and originally developed by Efron, B., & Tibshirani, R. (1986). Bootstrapping is a clever and relatively simple technique taking full advantage of the computing power that we now have.

Szantoi, Z., Jaffrain, G., Gallaun, H., Bielski, C., Ruf, K., Lupi, A., Miletich, P., Giroux, A.-C., Carlan, I., Croi, W., Augu, H., Kowalewski, C., and Brink, A.: Quality assurance and assessment framework for land cover maps validation in the 770 Copernicus Hot Spot Monitoring activity, *Eur. J. Remote Sens.*, 54, 538–557, <https://doi.org/10.1080/22797254.2021.1978001>, 2021

Schreuder, H., Ernst, R., & Ramirez-Maldonado, H. (2004). Statistical techniques for sampling and monitoring natural resources (Gen. Tech. Rep. RMRS-GTR-126; p. 111). U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station.

Gallaun, H., Steinegger, M., Wack, R., Schardt, M., Kornberger, B., & Schmitt, U. (2015). Remote Sensing Based Two-Stage Sampling for Accuracy Assessment and Area Estimation of Land Cover Changes. *Remote Sensing*, 7(9), 11992–12008.

Efron, B., & Tibshirani, R. (1986). Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy. *Statistical Science*, 1(1), 54–75. <https://doi.org/10.1214/ss/1177013815>

We have added a small statement in section 3.7.1 to justify our choice and added a reference to Schreuder et al. (2004):

*To calculate 95% confidence intervals for each metric, we applied bootstrapping with replacement (Schreuder et al., 2004; Szantoi et al., 2021). Unlike the original method proposed by Stehman & Foody (2019), this bootstrapping approach does not require the*

*assumption of a normal distribution and takes full advantage of today's computational power for estimating confidence intervals.*

**Section 3.7.2: Can the authors add a figure or table that shows the geographic distribution of the reference data from the Street Imagery validation? The metrics based on this are described as “Global” but it is not clear how globally representative that sample is.**

In this section we mention that coverage and availability of crop type data is limited, and we also write at the end of the section that “We did not use the term accuracy since the sample design is not probabilistic”.

We believe the coverage to be global as 78 countries have at least 1 crop type location checked, and 34 countries have more than 10 locations each. It is definitively not a representative but rather an opportunistic sample, based on available street-level imagery, where crop type could be identified using the mentioned tool.

We have added Figure 5 in Section 3.7.2 and inserted a reference to the figure in Section 5 (product validation).

The new figure is attached as a supplement to this comment.

*Figure 5. Locations with identified crop type across the globe, using the Street Imagery validation tool. Base map source: GADM.org.*

**Lines 546-647: It's not clear which of the 3 plots in Figure 9 are being described here. I don't see any hotspots (red) in the south Sahel region. In addition, can Figure 9 and Figure 10 be made higher resolution?**

The figures 9 and 10 indeed seem to have a resolution issue, which will be solved in the final version of the paper. For figure 9 (now Figure 12), we also changed the background color of the maps to yellow (indicating no difference between any of the two datasets) and added zoomed in examples of four different regions to make the maps easier to interpret. With hotspots we mean large blue areas of commission which are visible in all three of the comparison maps, this is now changed in the text. We also added a sidenote that not all areas of commission are a result of overestimating the irrigated area, but that this could also be a result of an increase in irrigation activity over the last couple of years, which is not recorded in either of the three reference maps due to their older production date:

*However, we assume that the blue hotspots in Sudan, the USA, Russia and Brazil are most likely WorldCereal commission errors since they become apparent when comparing our data with either of the other three reference datasets. Other blue areas, like the*

hotspots in Canada, do not necessarily have to be commission errors due to the recent increase in irrigation which occurred later than the production of the three reference maps (Statistics Canada, 2021).

The new Figure has been added as a supplement to this comment.

*Figure 12: Differences in percentages whilst comparing the WorldCereal combined irrigation product and: (a) the FAO global area equipped for irrigation in 2005 map, (b) the LGRIP30 irrigated area map for 2015, and (c) the International Commission on Irrigation and Drainage (ICID) world irrigated area dataset. The WorldCereal products show more irrigation in blue areas and less in red areas, compared to the other datasets.*

**Figure 10 and associated discussion: Are the areas in the yellow bars based on pixel counts? It is not clear how the authors computed the area. This analysis would be much stronger if there were error bars that we could use to compare, because it's possible an area calculated from these maps could be within error of each other even for large differences. Also, can the authors comment on how confident they are that the CIA and ICID estimates match the [theoretical] "ground truth"? It seems like they are being treated as ground truth, rather than simply assessing agreement between products which all lack ground truth. Since those are statistical estimates, did they provide errors that should be plotted too?**

The values in the yellow bars for the LGRIP30 dataset and our own dataset were derived by us. For this analysis, we resampled the products from their native resolution to the resolution of the FAO AQUASTAT dataset (0.083 degrees/ 5 arc minutes) and computed the irrigated area fraction for this pixel size based on pixel count. Subsequently, we vectorized the FAO AQUASTAT grid and calculated the area of each vector/pixel in QGIS. The pixel area was then multiplied by the irrigated area fraction to calculate the area of irrigated pixels per FAO AQUASTAT pixel, which was then summed to get the global irrigated area values. We added this explanation to the paper:

*We calculated the total irrigated area for the LGRIP30 and WorldCereal irrigated area maps by downscaling both maps to the resolution (5 arc minutes) of the FAO AQUASTAT area equipped for irrigation dataset (Siebert et al., 2013). To compare the three WorldCereal seasons with the other (annual) datasets, the three irrigated area datasets were merged into a single irrigated area map for 2021 where irrigated pixels indicate that in at least one of the three seasons irrigation was detected. During the downscaling process, the number of irrigated pixels within an AQUASTAT pixel was counted and used to calculate an irrigation fraction. These fractions were then combined with each pixel's surface area to compute the total irrigated area. The total irrigated area statistics of the other products shown in Figure 13 were calculated by their respective authors.*

The other global irrigated area statistics were computed by the authors of the cited studies. Wu et al. (2022) relied on pixel counting for their irrigated area estimation. Meier et al. (2018) and Siebert et al. (2013) do not elaborate how they calculate the total irrigated area. None of the above studies show any uncertainty metrics for the total irrigated area, so visualizing the uncertainties in error bars is unfortunately not feasible. Same goes for the statistical datasets, which have a different quality per country due to the level of detail in which irrigation is being monitored by the country as well as due to differences in reference periods for which the data was collected. For some countries very recent statistics are available, whilst for other countries data from 20 years ago was used. We would argue that on a country level, statistical datasets are superior, so also when calculating the total global irrigated area, the statistical datasets are in our opinion the best benchmark we have. However, the quality of these datasets differs per country as mentioned above. The only quality assessment we could find showed the quality of the statistical dataset based on a simple scoring system that ranges from very poor to very good (<https://hess.copernicus.org/articles/9/535/2005/hess-9-535-2005.html>, Table A1). Although we would agree it would improve the figure, providing error bars for the statistical datasets would be virtually impossible.

While reviewing our text we indeed realized it appears we treat the statistical datasets as the “truth”. To tackle this, we have (1) added a statement concerning the quality of statistical datasets for irrigation assessment and (2) rephrased some sentences to treat both datasets more equally.

*... statistical datasets from CIA and ICID. While the latter exhibit significant variation in quality and currency across different countries, these datasets are still regarded as the most accurate benchmarks available at country level.*

*Both the maps from Meier et al. (2018) and Wu et al. (2023) result in a slightly larger estimate of global irrigated areas (up to 30 % higher), while the LGRIP30 product shows an increase of more than 100 %. Finally, the WorldCereal irrigation product provides a significantly lower figure for global irrigated area (roughly 35 % less than statistical datasets), ...*

### **Code: Why not publish the code in a Github repository, which would be more accessible for many users?**

The code has been published in a Github repository (<https://github.com/WorldCereal/worldcereal-classification/tree/v1.1.1>). On top of this, Zenodo was used in order to link a specific DOI to the version of the code that was used to generate the official 2021 global products. You will find the link to the github repository included in the Zenodo repository.



## Reviewer Comment 3

<https://doi.org/10.5194/essd-2023-184-AC3>

The manuscript presents a suite of global crop products for year 2021, including annual temporary crop extent, seasonal active cropland, crop type maps of first and second season maize, winter cereals and spring cereals, as well as irrigation maps. Sentinel 2, Sentinel 1, thermal bands of Landsat 8, DEM, a biome map and geographic coordinates are the main input data. Vegetation indices and temporal features are derived to feed a CatBoost model with training data collected from various sources. The products are validated to a varying degree (see more details in comments below). The global map products, the reference datasets and the classification code are also made publicly available.

The entire work is very substantial. The authors are congratulated on the achievement.

Our main comment concerns the validation of these map layers. Only the annual temporary crop extent map was validated following established statistical protocols, and thus, the results included a complete list of accuracy measures. The crop type maps, arguably the main component of WorldCereal, were not properly validated, although the authors made substantial efforts to collect data. Using Google Street views to extract crop type information is very interesting. The irrigation maps were compared to existing irrigation datasets from FAO, whereas the seasonal active cropland maps were not assessed at all. Strictly speaking, the accuracies of most maps in WorldCereal are unknown, although this does not mean the quality of the maps is necessarily low. This should be acknowledged in the paper. It is done in the conclusion section but should also be reported in the abstract.

Thank you for this valuable comment regarding product validation. We agree the term “fully validated” used in both the abstract and conclusion section was a bit too strong and have adjusted formulation at both locations in the manuscript (cf. comment reviewer 2).

Additionally, we have added a few sentences in the abstract to clarify this. The relevant section now reads:

*Validation of the products was done based on best available reference data per product. A global statistical validation for the temporal crop extent product resulted in a user's and producer's accuracy of 88.5 % and 92.1 % respectively. For crop type, a verification was performed against a newly collected street view dataset (overall agreement 82.5 %) and various publicly available in-situ datasets (minimal agreement of 80 %). Finally,*

*global irrigated area estimates were derived from available maps and statistical datasets, revealing the conservative nature of the WorldCereal irrigation product.*

**We provide a few suggestions to improve the paper and to alert potential data users. First, the presentation of the validation method details needs to be improved. Global maps of validation reference data are strongly suggested to be included in the paper. Similar to Figure 1, which shows the distribution of training data, knowing the spatial distribution of validation data would allow readers to know which parts of the globe were better assessed.**

We agree such a new figure on the distribution of crop type validation data would significantly improve assessment of our validation effort by the reader. This was also suggested by reviewer 2. Accordingly, we have added the figure 5 to the manuscript, which has been added as a supplement to this comment.

*Figure 5. Locations with identified crop type across the globe, using the Street Imagery validation tool. Base map source: GADM.org.*

**In addition, Line 399 states that 3500 reference data points were collected, but the total number of records in Table 6 is about 2600, suggesting that the text descriptions need to be cross checked.**

Out of the 3500 collected locations, some samples were discarded from further analysis due to two main reasons: either the location contained a perennial crop (not useful for temporary crop type validation) or the location fell outside the WorldCereal temporary crop extent product. A total of 2617 locations were retained for further analysis. We have added the following sentence in section 3.7.2 to clarify this:

*After discarding perennial crops and locations outside the WorldCereal temporary crop mask, 2617 samples remained for crop type validation.*

**Second, the global crop type maps can be compared with regional products of known quality and similar resolution. For instance, the United States and Canada maintain annual production of crop type maps at 30 m resolution. These maps are coarser than WorldCereal and they have errors themselves. However, a pixel-by-pixel comparison would still provide very useful information that is absent in the current manuscript.**

We agree with the reviewer that such comparisons would prove useful to further demonstrate the quality of our crop type products. In fact, such efforts have been undertaken and will be made publicly available in the official WorldCereal validation report(s). For convenience, we included some additional figures in the manuscript. In response to a comment made by reviewer 2, we had already included the results of crop type validation over Ukraine, which presented a nice use case

demonstrating temporal robustness of our crop type detectors. In addition, we now added the comparison with USDA crop data layer and the Canadian 2021 maps as well (Section 5):

*In addition to this global effort, a regional comparison with USDA Crop Data Layer 2021 resulted in an overall agreement of 82.9 %, with class-specific agreements of 80.2 % and 93.8 % for maize and 84.9 % and 66.5 % for cereals respectively. For Canada, we found an agreement of 96 % and 80 % for maize and cereals respectively and noted major confusion between winter and spring cereals (Annual Crop Inventory Ground Truth Data, Canada, 2021). In order to demonstrate the temporal robustness of the WorldCereal models an additional validation effort was done for Ukraine based on an independent dataset obtained from the National Technical University of Ukraine, Kyiv Polytechnic Institute (Kussul et al., in preparation). Country-wide crop type maps were generated for the period 2018-2021. Overall, user's and producer's accuracies were found to remain stable across the years (OA of 92.5%, 92.7%, 85% and 93.9% respectively), despite only limited training data (2018-2019) being available for this particular country.*

We also added a small statement in Section 3.7.2 to highlight the additional analysis performed:

*In addition to this global effort, comparisons were performed with publicly available regional in-situ reference datasets (Sect. 3.1) and randomly sampled locations from existing crop type maps (USDA Crop Data Layer 2021) to further demonstrate crop type product quality.*

**Lastly, on the global scale, similar to the method used to assess the irrigation maps, the FAO county-level crop area statistics and other coarse-resolution crop type masks can be used to evaluate the overall bias of the crop type maps, again, acknowledging the uncertainties in the FAO and existing data themselves.**

We fully agree with the reviewer that such a country-level statistical comparison would prove an interesting means to further validate our crop type products. Considering the large variability in quality and currency of country-level statistics, this would represent a huge undertaking and was therefore considered to be out-of-scope for the present manuscript. The reason why we did perform such an analysis for the irrigation product (at global scale) is because we lacked the necessary in-situ reference data for conducting a holistic validation of the irrigation product.

Nevertheless, FAO has been involved in the WorldCereal project as a core user since the beginning of the project and is now actually working on performing such in-depth statistical analysis. FAO is planning to publish these results in a separate article and is currently investigating how the WorldCereal products can be used to improve and update their database on national crop statistics. In the newly inserted

section on Data usage and future prospects (Section 6), we also make reference to the potential of the WorldCereal products in this respect:

*Additionally, the data can be downscaled and as such prove useful to enhance and complement subnational to national agricultural/water use statistics (e.g. FAOSTAT and AQUASTAT, both produced by FAO). Important to note here is that our low resolution fraction maps (Figures 7, 8 and 9) have been generated purely for visualization purposes and do not represent proper area statistics. We refer to Olofsson et al. (2014) for detailed guidelines on deriving regional statistics from pixel-based classification maps.*

#### **Other minor comments:**

**Line 29-30. See comments on validation above. “Fully validated” sounds like an overstatement.**

This comment was already addressed in response to reviewer 2 (see also first comment above).

**Line 71. The Introduction section summarized a few studies on global cropland extent mapping, but provided no context on crop type and irrigation mapping. Existing global irrigation data such as those used in the product comparison should be mentioned here. There are many recent studies on regional crop type mapping that should also be summarized.**

Important to note here is that we do not intend in any way to provide a complete review of global (and especially not regional) mapping products on agricultural land cover, crop type and irrigation. Our intent here was to mainly highlight the limitations of existing mapping products. Nevertheless, we do agree with the reviewer that at minimum some context should be provided in terms of (global) crop type and irrigation mapping in this introductory section. We therefore added two paragraphs in the introduction to address this comment:

*With regards to crop type specific data, Han et al. (2021) produced the first global and annual maps of rapeseed planting area for 2017-2019 at 10m resolution based on Sentinel-1 and 2 data. However, most high-resolution crop type products available to date are restricted in terms of spatial coverage, highlighting the complexity of global crop type mapping. d’Andrimont et al. (2021) produced the first 10 m resolution crop type map for the European Union, covering 19 different crop types, based on Sentinel-1 data. Li et al. (2023) developed 10 m resolution maps for maize and soybean over China for 2019 based on a combination of PlanetScope and Sentinel-2 data. ESA’s GeoRice project generated high-resolution rice maps for Southeast Asia for 2018-2020 based on Sentinel-1 data, while various regional crop type mapping projects in Africa are being set up under the Digital Earth Africa umbrella. Becker-Reshef et al. (2023) collected and harmonized various regional crop type products to generate global Best Available Crop*

*Specific Masks (BACS) for wheat, maize, rice and soybeans in the context of global food security monitoring.*

*The first global irrigation datasets have been typically derived from a combination of statistics and inventories, with a minimal role for earth observation data (e.g. FAO's area equipped for irrigation map; Siebert et al., 2013). This map was further improved by Meier et al. (2018) through combination with remote sensing based land cover maps (ESA CCI), land suitability maps and long time series of the Normalized Difference Vegetation Index (NDVI) in a multi-criteria decision framework. Salmon et al. (2015) relied on a combination of survey data, remote sensing time series and climate data to train a supervised classification model to distinguish rainfed, irrigated and paddy croplands. Detecting irrigation purely from satellite observations is a challenging effort that can be addressed in various ways, employing microwave-based soil moisture estimates, optical satellite observations and/or measurements of crop water stress through thermal satellite data (Massari et al., 2021). Thenkabail et al. (2009) created a 1 km resolution irrigated area map based on a combination of optical satellite data (SPOT VGT), thermal satellite data (AVHRR), a digital elevation model and climate data. In contrast, Wu et al. (2022) relied exclusively on long time series of NDVI data in a locally tuned thresholding system to detect irrigation activities in dry periods at global scale. Most recently, Teluguntla et al. (2023) combined the Global Cropland-Extent Product at 30-m Resolution (GCEP30; Thenkabail et al., 2021) with multiple spectral bands and indices of Landsat-8 from 2014-2017 in a supervised machine learning approach, resulting in a high-resolution (30 m) global irrigated area product (LGRIP30).*

*Despite the clear increase in global agricultural mapping products, existing initiatives ...*

**Line 175. Were data of the calendar year 2021 used for mapping? If so, would the Southern Hemisphere growing seasons be interrupted?**

The WorldCereal system is a growing season specific mapping system, meaning that it starts from the definition of the major growing seasons for cereals and maize in each individual agro-ecological zone in order to determine the proper time range needed for generating each individual product. This means that we do not strictly follow calendar years, but mapping periods might cross calendar years in case the season of interest for instance starts in September and ends in March. During this global demonstration, products were generated for the year 2021, meaning that each growing season which ends in calendar year 2021 is covered by our suite of products. In order to clarify this link between crop calendars, agro-ecological zones and the range of time for which a certain product is valid, we have:

- Added a generic figure (Figure 1) describing the workflow of product generation, in which the link between growing season and processing period is explicitly made. Alongside Figure 1, we have added a generic introduction to the Materials and methods section.

- Added a specific figure (Figure 6) to provide two examples on how the different products are defined in terms of timing for two significantly different AEZ. Especially this figure should make it more clear to the reader that calendar years may be crossed depending on the definition of the crop seasonality in the AEZ under consideration.

The two new figures have been added as a supplement to this comment.

*Figure 1: WorldCereal production flowchart detailing the steps required to generate the WorldCereal products for a user-defined area and year of interest. This schematic does not include training of the classification model (detector).*

*Figure 6: Demonstration of WorldCereal 2021 product generation and timing for two distinctive AEZ's. (a) AEZ located in central U.S., where only one maize season occurs and spring cereals are mapped jointly with maize and (b) AEZ located in Somalia, where two maize seasons occur and no spring cereals are mapped. Product abbreviations are explained in Table 4, whereas locations of AEZ are highlighted in Figure 3. The "x" on each line defines the earliest moment in time the associated products can be generated.*

**Line 225. Can you include crop calendar maps for maize and cereals in the paper, in addition to Figures 2 and 3?**

In the context of this review we have already added 3 new figures to this manuscript (Figures 1, 5 and 6). The global crop calendars for maize and cereals are the explicit subject of another publication, hence we would prefer not to repeat this information in the current manuscript. Instead, we provide examples of the maize and cereal seasonality for two distinctive AEZ's in the new Figure 6.

**Line 258. Can you provide justification for the 2.5 and 10 degrees perturbations?**

This comment was already addressed in response to reviewer 2:

The random perturbation degree of lat/lon coordinates was empirically determined in a sensitivity analysis where a trade-off was found between (a) a model that is too sensitive to knowing exact locations (limited to no perturbation), which resulted in artefacts, especially in regions with considerable lack of training data, and (b) a model that no longer uses the localization information (too much perturbation). We agree with the reviewer that the resulting degree of perturbation is somewhat large. This could be diminished in future versions in case more training data would become available, especially over data-sparse regions.

In the manuscript, we added these ranges were "determined empirically".

**Line 364. Turning on the active cropland label when a pixel is classified as a crop type implicitly assumes that the crop type maps are more accurate than**

**the active cropland map. Is this true? If the active cropland map was more accurate, it could be used to remove some commissions errors of the crop type maps. Otherwise, this would introduce commission errors to the active cropland map.**

The reviewer is absolutely correct. The crop type detectors have been specifically trained and validated for their purpose and comes with a model-based confidence score indicating expected product quality. The active cropland marker represents a generic indication of crop growth, is known to be sensitive to cloud coverage (see one of the following comments), potentially suffers from commission problems (e.g. in case of recent fallow) and could not be properly validated due to a lack of reference data. For these reasons, we consider product quality of the crop type products to be superior to the active cropland marker and opted to use the former to correct the latter and not vice-versa.

**Line 385. Change Modis to MODIS.**

Change implemented, thank you for noticing.

**Line 420. Visual assessment of the temporal completeness of VI curves would allow you to evaluate the impacts of clouds and missing data on the active crop maps.**

The reviewer is correct that there is a clear relation between the temporal completeness of the VI curve (EVI) used as input for growing season delineation on the one hand and the resulting quality of the active cropland marker on the other hand. For each product we generate in the WorldCereal system, we also keep track of the number of valid optical (and SAR) observations used during product generation. During visual assessment of the active cropland marker, we could see that in regions with large data gaps in the optical time series, the active cropland marker becomes more erratic. This behaviour is also expected for the other products, although the impact is probably lower as the models have been specifically trained to deal with a low number of observations in the regions where this might present a potential problem. The minimum number of observations required to get an accurate indication of active cropland is however hard to determine, as it is additionally affected by the timing (and spread) of these observations and local agricultural practices (one large season or several short growing seasons). Any user who would run the system in the future would get access to this information for the generated products, so he/she can take this into account when using the active cropland marker (or any other product).

We have added a sentence to Section 3.7.4 to warn the user on this potential restriction of the product:

*Note that the quality of this product is expected to be lower in regions with few valid (non-cloudy) optical observations, as this is the only input used for delineating the growing season and determining active crop growth (Sect. 2.5).*

**Line 530. See comments above. The total amount of points in the table is less than 3500.**

We refer to our answer to an earlier comment for this. Clarification has been added to the manuscript:

*After discarding perennial crops and locations outside the WorldCereal temporary crop mask, 2617 samples remained for crop type validation.*

**Line 576-595. It is strange to see the map products, reference datasets and code are published on multiple websites.**

We have deliberately chosen to publish the different aspects of the WorldCereal system on the platform most suited to support its purpose and further use. The WorldCereal products were published on Zenodo (to ensure existence of a proper DOI) and made available through Google Earth Engine to ensure a large uptake by the user community. The reference data was published in a Zenodo repository specifically designed to have easy (API) access to the reference data, to ensure users being able to easily download individual datasets for their own use. The project code was published on Github as it is known as one of the most convenient platforms to use for hosting and collaborating on code. Additionally, it was also published as a Zenodo repository to ensure assignment of a unique DOI to the version used to generate the 2021 global WorldCereal demonstration products (cf. last comment by reviewer 2). Ultimately, all aspects were published on Zenodo to ensure proper reference (through DOI) could be made to the individual aspects of the project.