Earth System Science Data

Paper # essd-2023-432

May. 25, 2024

Dear editor and reviewers:

We are very grateful to you for your constructive comments and suggested amendments on our manuscript entitled: "**High-resolution mapping of global winter-triticeae crops using a sample-free identification method**" (essd-2023-432). The comments have improved the paper quite tremendously. We have carefully studied the comments and revised our manuscript accordingly.

Here are our detailed responses to your comments. Please note that the comments from you are in **bold font** followed by our responses in regular font, changes/additions to the manuscript are <u>underlined</u>.

Sincerely yours,

Wenping Yuan on behalf of all co-authors

Corresponding author: Wenping Yuan, Ph.D., Professor

School of Atmospheric Sciences,

Sun Yat-sen University

135 West Xingang Road, Guangzhou 510275, China

E-mail address: yuanwp3@mail.sysu.edu.cn

Detailed responses to reviewers' comments

Reviewer #1

1). There is no methodology mentioned in abstract. "First global 30 m resolution distribution maps was produced from what and how? After this, add one to two sentences about how the study was carried out?

Response: Thanks for your comments. In fact, we have described the method used to produce the distribution map of winter -triticeae crops before this sentence. Here, we have revised and refined this sentence:

"In this study, we propose a new method based on the Winter-Triticeae Crops Index (WTCI) for global winter-triticeae crops mapping. This is a new sample-free method for identifying winter-triticeae crops based on differences in their normalized difference vegetation index (NDVI) characteristics from the heading to the harvesting stages and those of other types of vegetation. <u>We considered state (or province) or country as an identification unit and employed WTCI to produce the first global 30 m resolution distribution maps of winter-triticeae crops from 2017 to 2022 using Landsat and Sentinel images.</u>"

2). Line 50: clarify what is 'long-term distribution maps"

Line 50: Therefore, it is necessary to produce long-term distribution maps of winter-triticeae crops with a high-spatial resolution for these countries.

Response: We have revised this sentence:

"<u>Therefore, it is necessary to produce distribution maps of winter-triticeae crops with high-spatial resolution and continuous years for these countries.</u>"

3). Line 64 -70: these are about method and better fits in section 2 Data and method not in section 1.

Line 64 -70: Here, based on Landsat 7, Landsat 8 and Sentinel-2 satellite data, this study developed the Winter-Triticeae Crops Index (WTCI), a sample-free method for identifying the global distribution of winter-triticeae crops. Specifically, we first designed the WTCI based on the NDVI differences between winter-triticeae crops and other vegetation types. Then, we applied this method to identify the winter-triticeae crops in 66 countries worldwide. Finally, we assessed the accuracy and spatiotemporal transferability of the WTCI method based on field survey samples, visual interpretation samples from high-resolution images on Google Earth, CDL dataset, the Land Parcel Identification System (LPIS) dataset and agricultural statistical data. Ultimately, we produced 30 m spatial resolution distribution maps of winter-triticeae crops from 2017 to 2022 in 66 countries (2020 for US, see 2.2.2 for details) worldwide to fill such product gaps, providing a data basis for yield estimation and crop management.

Response: Thanks for your suggestion. We have revised this paragraph:

"This study aims to develop a new sample-free method, i.e., Winter-Triticeae Crops Index (WTCI), to identify global winter-triticeae crops based on Landsat 7, Landsat 8, Sentinel-1 and Sentinel-2 satellite data. The main goals are to (1) assess the accuracy and spatiotemporal transferability of the new method using field survey samples, visual interpretation samples from high-resolution images on Google Earth, CDL dataset, the Land Parcel Identification System (LPIS) dataset and agricultural statistical data, (2) produce 30 m spatial resolution distribution maps of winter-triticeae crops in 66 countries worldwide from 2017 to 2022 to fill such product gaps, providing a data basis for yield estimation and crop management."

4). Methodology needs better explanation because results can be trusted based on methodology. Line95: how many data sets used of Landsat, Sentinel-1. Similarly, procedure is too short, for example what was the process of noise removal, radiometric calibration, terrain correction – there are no theorical foundation nor procedure explained.

Line 95: In this study, we used Landsat 7 collection 2 data and Landsat 8 collection 2 data, as well as Sentinel-2 data on the Google Earth Engine (GEE) platform to obtain NDVI from 2016 to 2022, all of which were surface reflectance (SR) products and have undergone atmospheric correction.

Response: Thank you for your suggestion. We used Landsat 7, Landsat 8, Sentinel-2 and Sentinel-1 datasets in our study. We obtained all available images during the study period from these datasets, and processed these data pixel by pixel (2.2.1 section in the original manuscript), therefore, we did not calculate how many scenes were used in this study. In addition, the sentinel-1 data for each scene provided on the GEE platform has been pre-processed with thermal noise removal, radiometric calibration, and terrain correction (https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S1_GRD). Each procedure of processing NDVI and VH data is described in section 2.2.1 of the original manuscript. Here, we revised some contents, and the details are as follows:

"In this study, we used all available Landsat 7 collection 2 data (USGS Landsat 7 Level 2, Collection 2, Tier 1) and Landsat 8 collection 2 data (USGS Landsat 8 Level 2, Collection 2, Tier 1), as well as Sentinel-2 data (Harmonized Sentinel-2 MSI: MultiSpectral Instrument, Level-2A) on the Google Earth Engine (GEE) platform to obtain NDVI from 2016 to 2022, all of which were surface reflectance (SR) products and have undergone atmospheric correction."

"The data provided on GEE platform has undergone thermal noise removal, radiometric calibration, and terrain correction. We applied a refined Lee filter (Abramov et al., 2017) to alleviate the impact of speckle noise caused by the interferences between adjacent backscatter returns, and finally obtained the monthly maximum composite values of VH from 2016 to 2022 and resampled them to 30 m using the nearest neighbour method to keep consistency with NDVI. These operations were also run on the GEE platform."

5). Line 171-172. "On the contrary, the NDVI of natural vegetation peaks". Incomplete sentence.

Line 171-172: On the contrary, the NDVI of natural vegetation peaks.

Response: Thank you for your reminder. We have modified this sentence:

"On the contrary, the NDVI of natural vegetation approaches its peak in a year."

6). Section 2.4 what was references used for accuracy assessment.

Response: We added some references for accuracy assessment, and the details are as follows:

"The producer's accuracy (PA), user's accuracy (UA), overall accuracy (OA) and F1 score (Congalton, 1991; Hripcsak and Rothschild, 2005; Lin et al., 2022) were employed to validate the identification accuracy at the pixel scale.”

“The correlation coefficient (R^2) and relative mean absolute error (RMAE) were used to examine the consistency between the identified area and the statistical area (Shen et al., 2023; Zheng et al., 2022).”

References:

Lin, C. X., Zhong, L. H., Song, X. P., Dong, J. W., Lobell, D. B., Jin, Z. N.: Early- and in-season crop type mapping without current-year ground truth: Generating labels from historical information via a topology-based approach, Remote Sens. Environ., 274, 112994, https://doi.org/10.1016/j.rse.2022.112994, 2022.

Shen, R. Q., Pan, B. H., Peng, Q. Y., Dong, J., Chen, X. B., Zhang, X., Ye, T., Huang, J. X., and Yuan, W. P.: High-resolution distribution maps of single-season rice in China from 2017 to 2022, Earth Syst. Sci. Data., 15, 3203–3222, https://doi.org/10.5194/essd-15-3203-2023, 2023.

Zheng, Y., dos Santos Luciano, A. C., Dong, J., Yuan, W. P.: High-resolution map of sugarcane cultivation in Brazil using a phenology-based method, Earth Syst. Sci. Data., 14, 2065–2080, https://doi.org/10.5194/essd-14-2065-2022, 2022.

7). Texts are often confusing and readability is hampered. Even in results section, the sentences are composed as if it is method section still describing the method, e.g. line 261-266, line 311, 321,…. Sentences could be better composed to reflect the contents of the respective section. Another example, when you say, in conclusion, the study proposed a new method, then follow-up with a sentence describing salient feature of that method WTCI, what and how to do. Also introduce the limitations, of the method if any, and the recommendation. The contents in conclusion section mostly reads as findings.

line 261-266: This study first identified the spatial distribution of winter-triticeae crops in 66 countries in 2020 based on the WTCI. Fig. 4 shows that winter-triticeae crops are mainly distributed in mid-latitude regions, including most countries in Europe (Fig. 4b), the plains of Asia (Fig. 4c), northern Africa (Fig. 4d), the southern edge of Australia (Fig. 4e), middle of US (Fig. 4f) and the southeast regions of South America (Fig. 4g). To display the detailed information of the winter-triticeae crops map produced by this study, we selected twelve typical areas in different countries to zoom in and compared them with high-resolution images from Google Earth (Fig. 5).

Line 311: To examine the temporal transferability of the WTCI method, this study applied the optimal percentile of the V and B lines in 2020 to other years.

Line 321: Based on CDL and LPIS datasets, we further validated the performance of the WTCI method in the US and Europe.

Response: Thank you for your suggestion. We have revised and refined some contents in results and conclusions sections, the details are as follows:

"The spatial distribution map of winter-triticeae crops in 66 countries in 2020 was first produced based on the WTCI method (Fig. 4), which effectively presented the distribution of winter-triticeae crops in the study area. Specifically, the winter-triticeae crops were mainly distributed in most European countries and Asian plains (Fig. 4b and 4c). To display the detailed information of the map of winter-triticeae crops, we selected twelve typical areas in different countries to zoom in and compared them with high-resolution images from Google Earth (Fig. 5). In general, despite some noise, the identification map clearly displays the fields planted with winter-triticeae crops and effectively distinguishes roads and rivers between the fields."

"The comparison between the identified and statistical areas of winter-triticeae crops indicates that the WTCI method can be effectively applied to other years."

"The distribution map of winter-triticeae crops exhibited high consistency with CDL and LPIS datasets."

"This study proposed a new sample-free method (WTCI) for mapping winter-triticeae crops and examined its performance in 66 countries worldwide. The new method exhibits high accuracy and strong spatiotemporal transferability by comparing the produced maps with field survey and Google Earth samples, the CDL and LPIS datasets, and agricultural statistical data. Overall, the OA and F1 score were more than 80% and 75% in most of identification units, respectively. The R^2 between identified and statistical areas in most of regions was greater than 0.6 in all years, and RMAE less than 30%. These satisfactory results indicate that the WTCI method can be used for long-term and large-scale crop mapping. At the same time, the first 30 m spatial resolution distribution maps of winter-triticeae crops from 2017 to 2022 produced by the WTCI method fills the current product gaps, which can be further served for the harvest area monitoring, yield estimation and agricultural management."

Reviewer #2

The authors responded to my comments and improved the study substantially. The main improvement includes the validation of obtained results with CDL and LPIS datasets. However, I still have some comments.

Response: Thank you for your comments and affirmation of our revised manuscript. Here, we have revised our manuscript based on your comments this time, and we also attached a point-by point letter to you. The detailed responses are listed below.

1). Thank you for addressing the issue of (lack of) harmonization of Sentinel-2 and Landsat data, as well as Landsat-7 failure-related striping issue. But I think it should be already mentioned in the methods/data description. And maybe it would be good to discuss more deeply the limitations and errors, such as maquis mentioned in my previous review. Maybe also adding one figure with examples of errors/uncertainties would be beneficial.

Response: Thank you for your deep thought and suggestion. We have added some contents in data and discussions sections, and the details are as follows:

"We choose Landsat 7 satellite to obtain more available data although a malfunction in its scan line corrector. To ensure the data quantity and quality, we first removed the pixels with clouds."

"Additionally, due to the scan line corrector failed of the Landsat 7 sensor, the striping issues and reduced data availability may also impact the accuracy of NDVI time series (Ju and Roy., 2008), resulting in the errors in identification results. In our study, there were some striping issues in the distribution map of winter-triticeae crops in a few regions (Fig. S1a), which may lead to errors in winter-triticeae crops identification and the differences in identification results between different years (Fig. S1)."

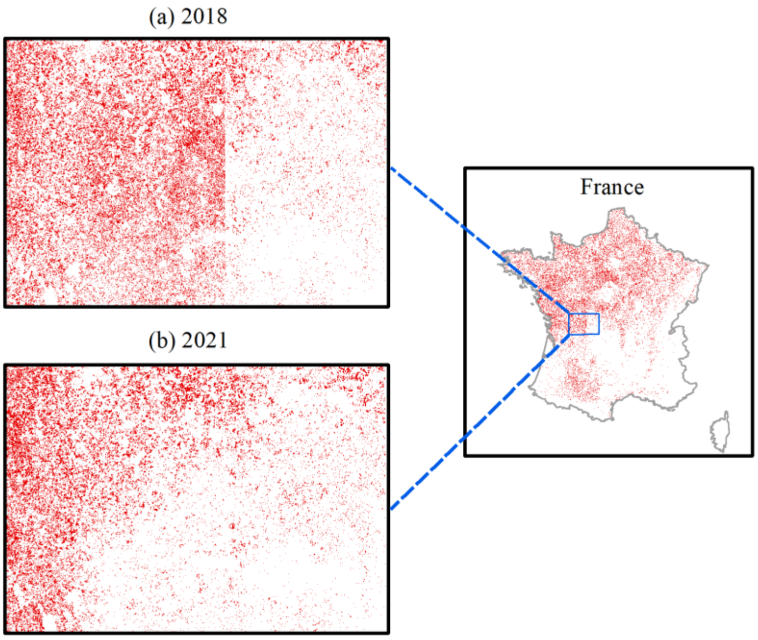

Figure S1: Comparison of distribution maps of winter-triticeae crops between different years. (a) and (b) show the zoomed-in maps of subregion in France in 2018 and 2021, respectively.

2). Furthermore, I still cannot fully understand how the authors selected samples based on Google Earth imagery and how they could say that these were winter triticeae and not other types of crops. I understand that they were later again checked with the NDVI curves. However, what you can see for example in Figure 5 – some of the identified winter triticeae crops are represented with vegetation and some with bare soils on the corresponding Google Earth high resolution images. What are the dates high resolution images from Google Earth were collected? I think this is important information and needs further clarification, because the types of crops may change from year to year.

Response: Thank you for your meticulous thinking and suggestion. We have added some contents and pictures to illustrate how we selected samples based on Google Earth image. The new pictures in Fig .2a are the Google Earth image corresponding to the filed survey. The details are as follows:

"We first chose regions with available images during the growing season of winter-triticeae crops (section 2.3.3), and selected samples from these regions based on the texture features and colors. Winter-triticeae crops have deeper color or stronger texture than winter rapeseed and grassland, and their roughness is lower than that of forest, which can be used to distinguish winter-triticeae crops from other land cover types (Fig .2a). Crops with different growing season (such as, maize, rice, and soybean) will not affect the visual interpretation. To ensure the accuracy of the samples, we then validated the selected samples on GEE platform by checking whether the NDVI temporal features of these samples matched the characteristics of winter-triticeae crops, and finally obtained 7,029 winter-triticeae crops samples and 8,897 non-winter-triticeae crops samples (Fig. 1)."

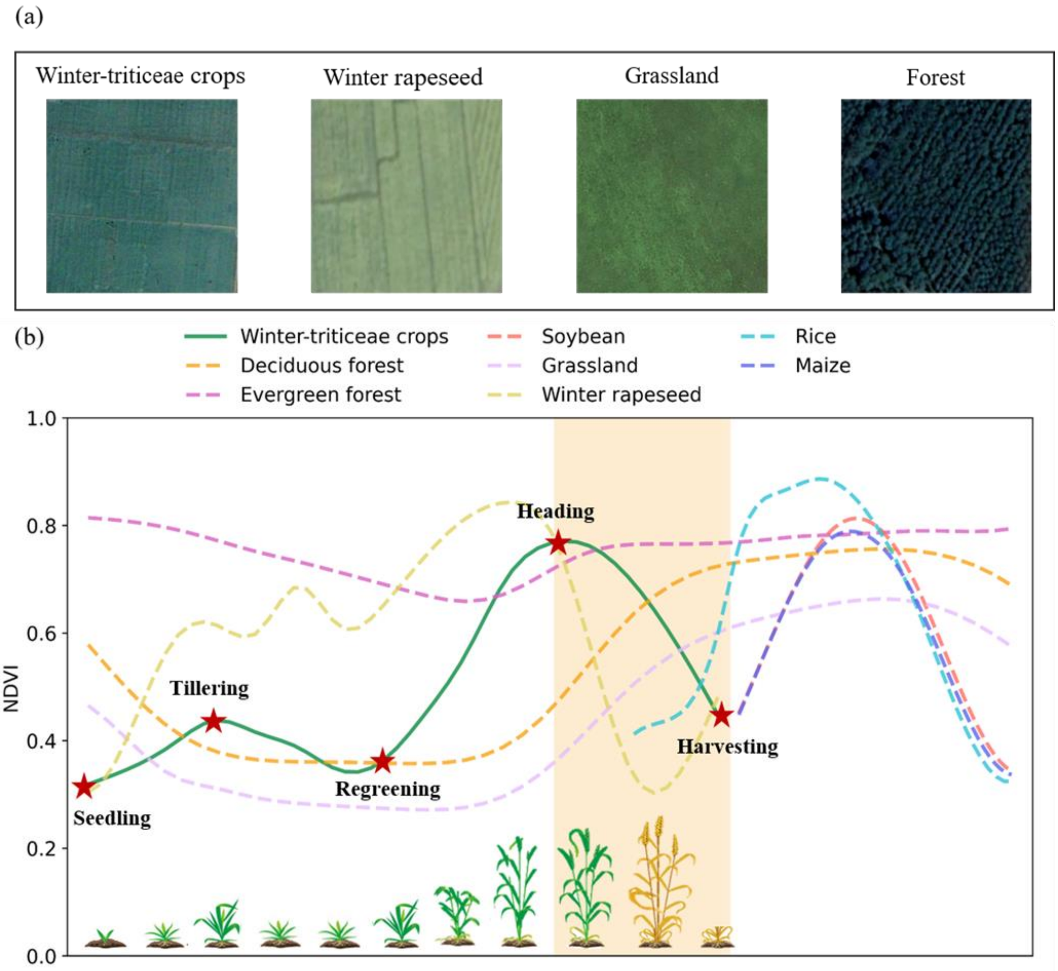

Figure 2: Example of the (a) textures and colors on the high-resolution images from © Google Earth and (b) NDVI time series characteristics of different land cover types. The red five-pointed stars represent the different phenological stages of winter-triticeae crops.

In addition, we have added dates for each image in Figure 5 and updated Fig. 5a and 5f to ensure that the displayed images are during the growing season of winter-triticeae crops. The identified winter-triticeae crops appears as vegetation or bare soil on the corresponding Google Earth high-resolution image may be related to the date of the image or the planting habits of farmers, such as the early or late planting or harvesting time. Besides, as described in the discussion section, the quality of satellite data can also lead to some identification errors. In general, the identification map clearly displays the winter-triticeae crops fields and effectively distinguishes other land covers.

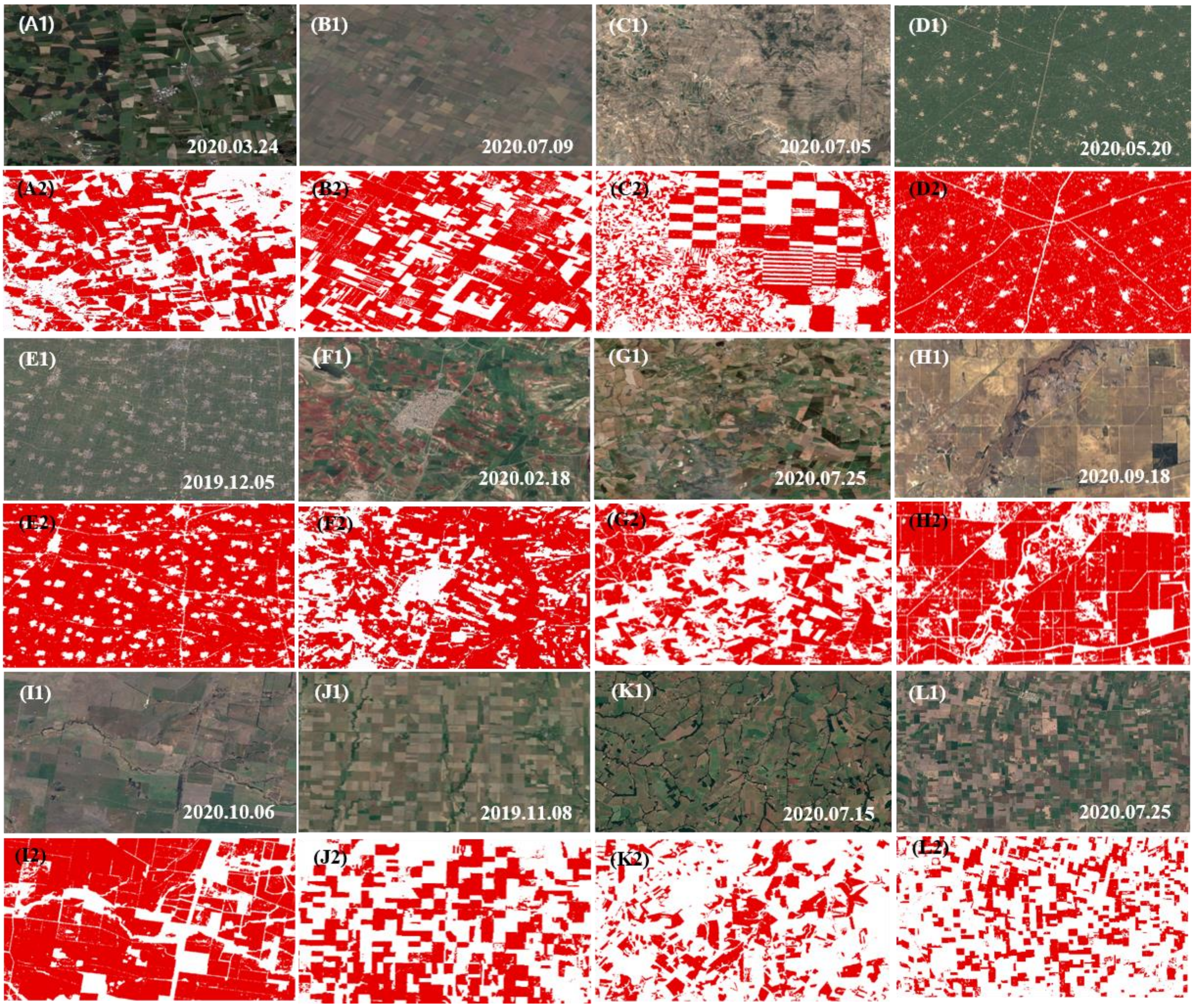

Figure 5: Comparison between the identification maps of winter-triticeae crops and high-resolution images from © Google Earth in the study area. (A1-L1) represent the high-resolution images from Google Earth of different regions; (A2-L2) represent the zoomed-in maps of area A-L in Figure 4.

Other comments:

3). Line 23 and others: F1 should be also reported in %, similarly as producer's and user's accuracy

Line 23： the overall accuracy and F1 score in most regions of the United States and Europe were more than 80% and 0.75.

Response: Thank you for your suggestion. We have revised F1 score in the manuscript to be expressed in %, and the details are as follows. Meanwhile, we have modified the corresponding figures (Figure 6, Figure 7 and Figure 10), and the details can be found in the revised manuscript.

"Moreover, compared with the Cropland Data Layer (CDL) and the Land Parcel Identification System (LPIS) datasets, the overall accuracy and F1 score in most regions of the United States and Europe were more than 80% and 75%."

"the overall accuracy (OA), producer's accuracy (PA), and user's accuracy (UA) of the winter-triticeae crops identification maps in 65 countries (except US) were 87.7%, 81.12% and 87.85%, respectively, and the F1 score was 84.04% (Fig. 6). PA and UA varied between 52% and 97.73%, 63.64% and 97.83% over the various countries, and OA and F1 ranged from 70.86% to 96.05% and 65.63% to 96.09%, respectively. At state (province) scale, the variation range of OA and F1 score in China were 77.68% to 95.9% and 71.79% to 94.47%, respectively (Fig. 7a). In Brazil, the OA and F1 score were in the range of 76.99%-94.74% and 78.26%-96.24% (Fig. 7b). The OA in India was between 67.53% and 92.07%, and the F1 score was between 65.24% and 92.05% (Fig. 7c). The OA and F1 score in Australia lied in the range of 79.21% to 91.67% and 69.23% to 91% (Fig. 7d). In general, the F1 score in most of the identification units was greater than 75%, indicating that the WTCI method shows satisfactory accuracy in identifying winter-triticeae crops. The regions with F1 scores less than 75% were mainly found in small winter-triticeae crops planting areas and complex winter crop types, such as Croatia (HRV), Albania (ALB), Sichuan (SC) province in China, and Bihar (BR) state in India."

"In 2020, the OA and F1 score in the US were 86.84% and 82.09%."

"For all states planting winter-triticeae crops, the OA varied from 70.42% to 94.24%, and the F1 score ranged from 66.67% to 91.01% (Fig. 10a-10c)."

"In major planting states, such as Kansas, Oklahoma and Texas, the planting area of winter-triticeae crops accounted for approximately 50% of the total area of winter-triticeae crops in the US, with OA and F1 score over 92% and 85%, respectively (Fig. 10)"

"Among the 10 European countries from LPIS datasets, the OA, F1 score, PA and UA ranged from 71.22% to 94.79%, 67.67% to 90.14%, 63.68% to 84.77% and 71.43% to 96.24%, with the mean value of 83.88%, 78.87%, 73.18% and 86% (Fig. 10d), respectively.”

“In general, the OA and F1 score in most of regions of US and Europe were higher than 80% and 75%, implying that the WTCI method exhibited satisfactory performance compared to the CDL and LPIS datasets.”

“Overall, the OA and F1 score were more than 80% and 75% in most of identification units, respectively.”

4). Line 55: ability to what?

Line 55: For example, Ge et al. (2021) combined Landsat images with the CDL production of Arkansas to train a classifier and then assessed the ability of the classifier in California, USA, and Liaoning, China.

Response: Thank you for your reminder. we have revised this sentence:

“For example, Ge et al. (2021) combined Landsat images with the CDL production of Arkansas to train a classifier and then assessed the spatial transferability of the classifier in California, USA, and Liaoning, China.”

5). Line 73: I think you should use words for cardinal numbers less than 10

Line 73: The study area covers 66 countries, including 36 European countries, 15 Asian countries, 8 African countries, 2 North American country, 4 South American countries, and 1 Oceania country (Fig.1).

Response: Thank you for your suggestion. we have modified the representation of numbers less than 10:

“The study area covers 66 countries, including 36 European countries, 15 Asian countries, eight African countries, two North American country, four South American countries, and one Oceania country (Fig.1).”

6). Figure 1: validation samples are not very well visible, maybe use more distinct colors

Response: Thank you for your suggestion. We have modified the Figure 1:

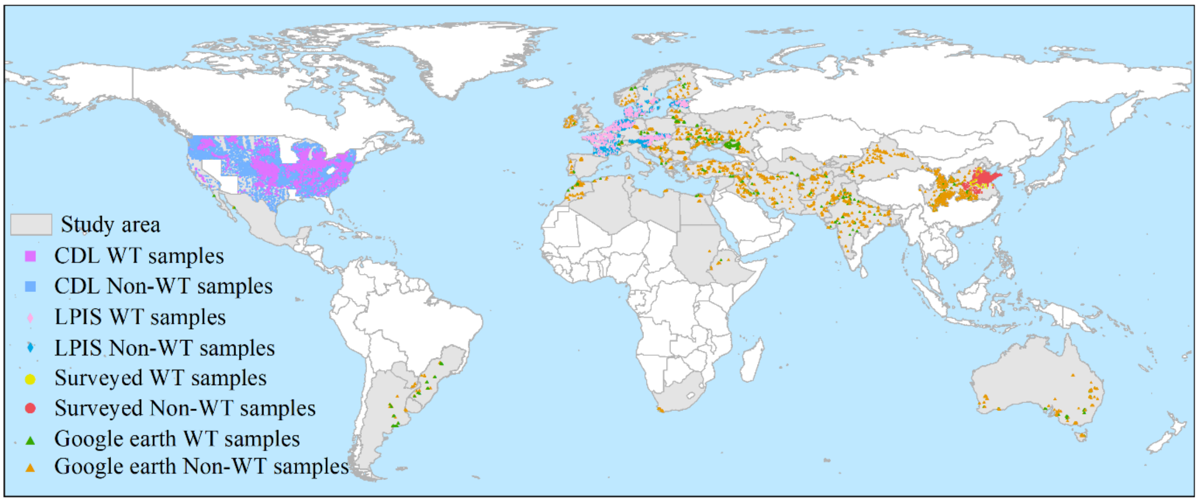

Figure 1: Distribution of the study area and validation samples. The study area is the region covered in grey; The legend indicates the winter-triticeae (WT) crops samples and non-winter-triticeae (Non-WT) crops samples from Cropland Data Layer (CDL) dataset of the United States, the Land Parcel Identification System (LPIS) dataset of Europe, and field survey in China, as well as visual interpretation base on Google Earth images, respectively.

7). Line 95: Please add the names of the used collections in GEE

Line 95: In this study, we used Landsat 7 collection 2 data and Landsat 8 collection 2 data, as well as Sentinel-2 data on the Google Earth Engine (GEE) platform to obtain NDVI from 2016 to 2022, all of which were surface reflectance (SR) products and have undergone atmospheric correction.

Response: We have added the names of the used collections in GEE, and the details are as follows:

"In this study, we used all available Landsat 7 collection 2 data (USGS Landsat 7 Level 2, Collection 2, Tier 1) and Landsat 8 collection 2 data (USGS Landsat 8 Level 2, Collection 2, Tier 1), as well as Sentinel-2 data (Harmonized Sentinel-2 MSI: MultiSpectral Instrument, Level-2A) on the Google Earth Engine (GEE) platform to obtain NDVI from 2016 to 2022, all of which were surface reflectance (SR) products and have undergone atmospheric correction."

8). Line 284: plating -> planting

Line 284: The regions with F1 scores less than 0.75 were mainly found in small winter-triticeae crops planting areas and complex winter crop types.

Line 286: On the contrary, the identification accuracy of regions with larger planting areas of winter-triticeae crops was significantly higher than that of regions with smaller plating areas.

Response: Thank you for your reminder. We speculate that you are referring to line 286, and we have revised this word:

"On the contrary, the identification accuracy of regions with larger planting areas of winter-triticeae crops was significantly higher than that of regions with smaller planting areas."

9). Figure 6: what a-e represent? Maybe you can sort countries by descending/ascending accuracy?

Response: Thank you for your suggestion. Figure 6a-6e represent the producer's accuracy (PA), user's accuracy (UA), overall accuracy (OA) and F1 score of the identification maps of winter-triticeae crops at each country in 2020. It is a good idea to sort countries by descending/ascending accuracy. But we think sorting them in ascending order of country abbreviations is more convenient for readers to match the identification results of these countries in the figure with the full names of the countries in the supplement, therefore, we modified Figure 6 according to this rule. Meanwhile, we have also made modifications to Figures 7-11 according to the same rules to ensure consistency throughout the entire manuscript. The details can be found in the revised manuscript.

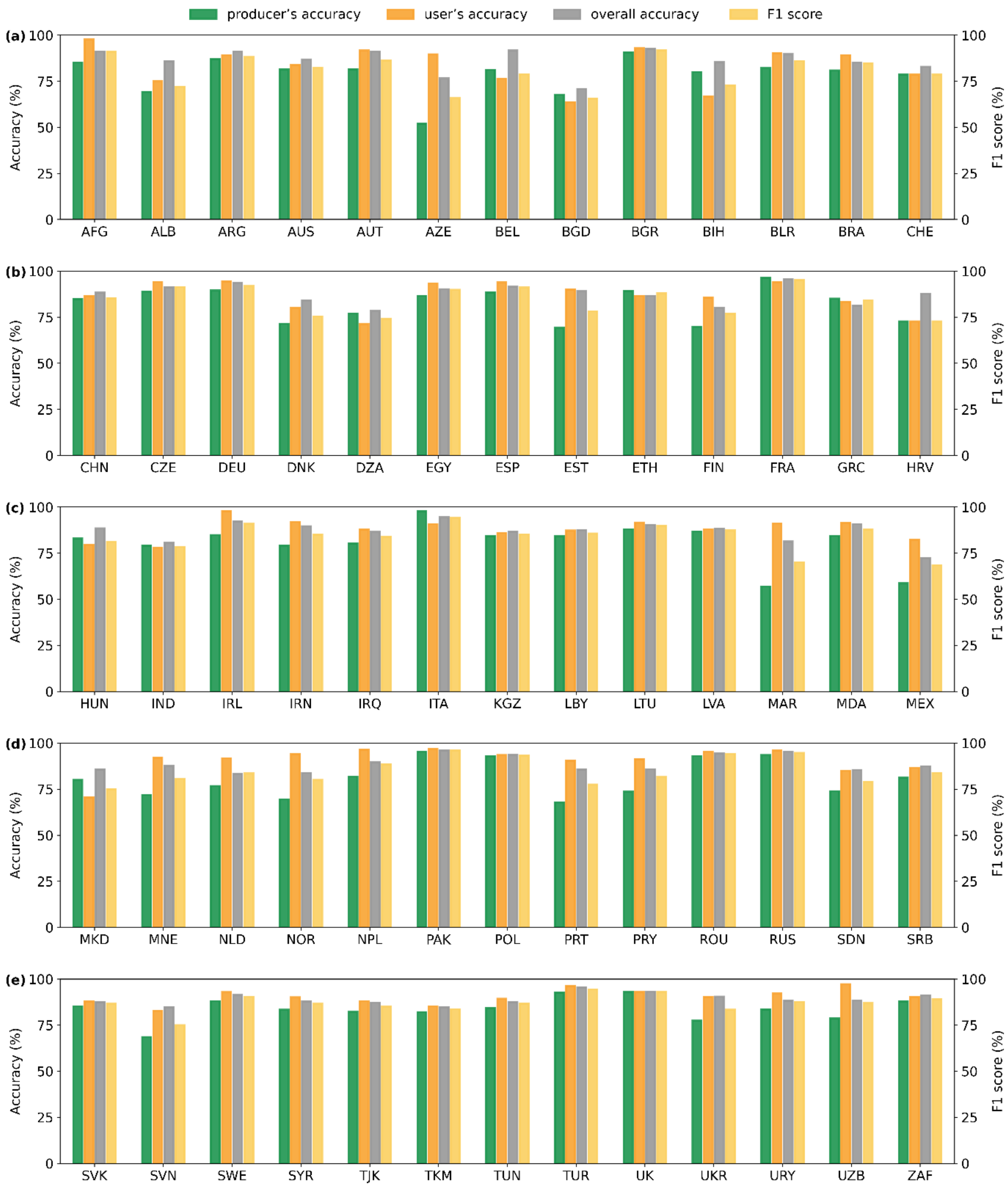

Figure 6: The producer's accuracy (PA), user's accuracy (UA), overall accuracy (OA) and F1 score of the identification maps of winter-triticeae crops at national scale in 2020. The abbreviations of countries are shown in Table S2 in the supplement.

10). Figures 8 and 9: I would use reversed color palette for RMAE (green – low error, red – high error).

Response: Thank you for your suggestion. We have revised Figure 8 and 9, and the details are as follows:

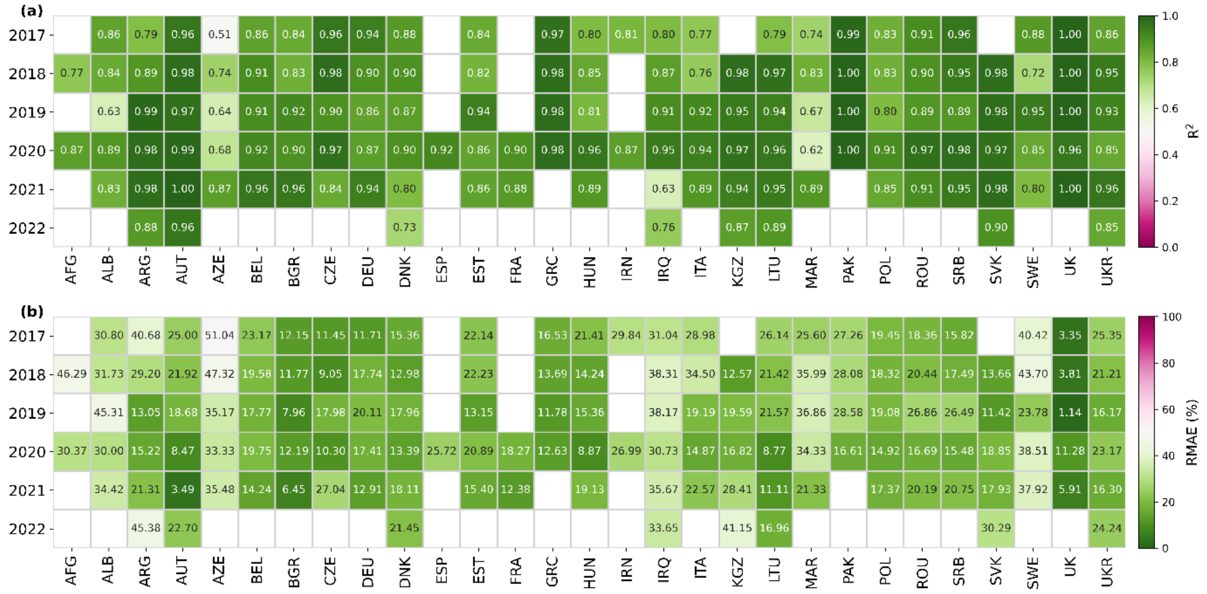

Figure 8: Comparison between identified and statistical areas of winter-triticeae crops at national scale from 2017 to 2022. (a) and (b) show the correlation coefficient and RMAE between identified and statistical areas, respectively.

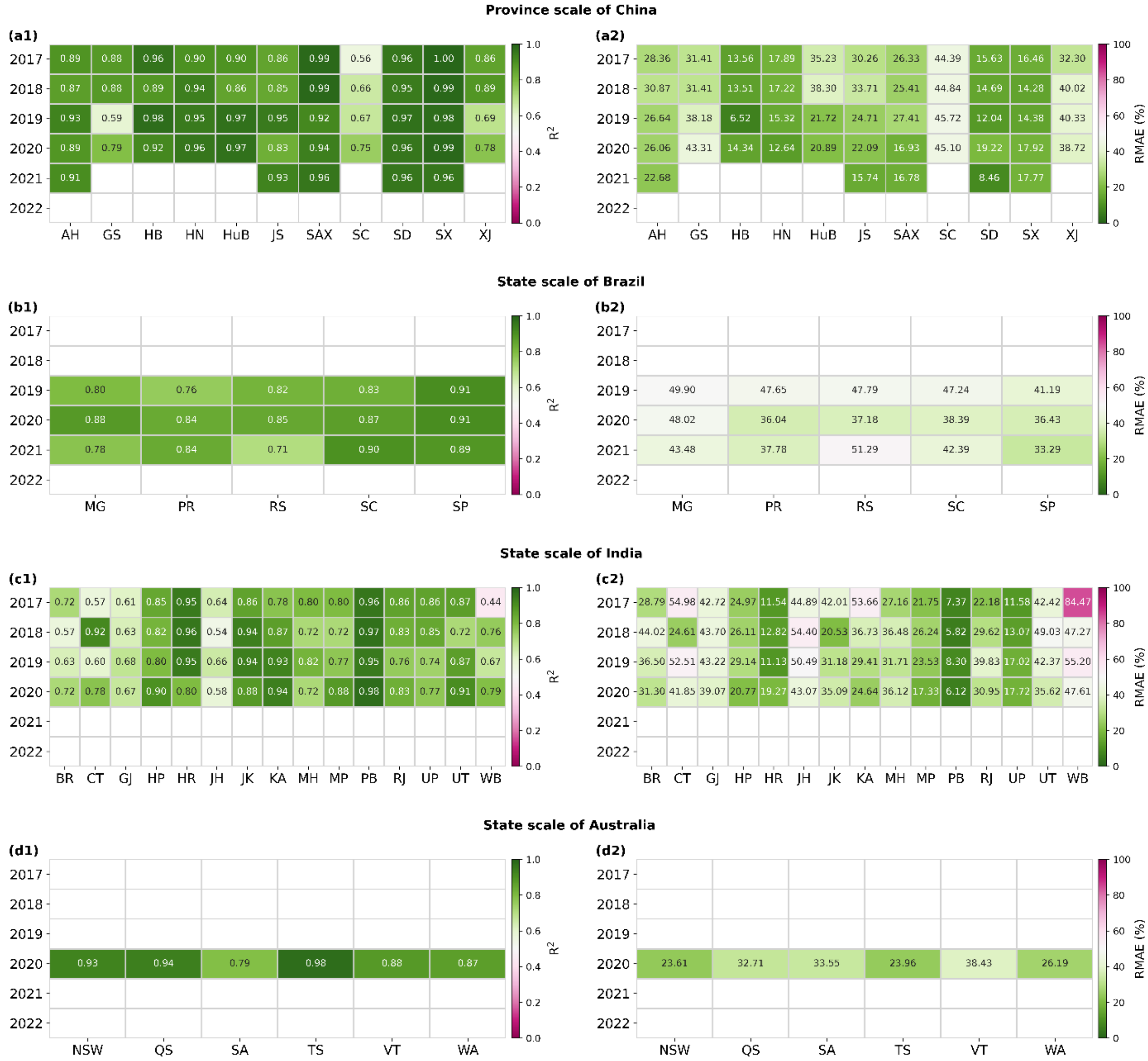

Figure 9: Comparison between identified and statistical areas of winter-triticeae crops at state (province) scale from 2017 to 2022. (a1-d1) represent the correlation coefficient at state (province) scale in China, Brazil, India, and Australia, respectively; (a2-d2) represent the RMAE at state (province) scale in China, Brazil, India, and Australia, respectively.

11). Figure 12: Why is LPIS shown as points and CDL as polygons?

Response: The CDL product is a crop classification map with a spatial resolution of 30 meters, so we would like to compare the spatial consistency between the identification map produced using the WTCI method and the CDL. The LPIS data we collected is the location information of land cover types, so we compared it with our identification map using coordinate points.

12). Figures 4 &13: are coordinates necessary here? You don't use them in the Figure 12.

Response: Thank you for your reminder. We have removed the coordinates to keep these figures consistent, and the details are as follows:

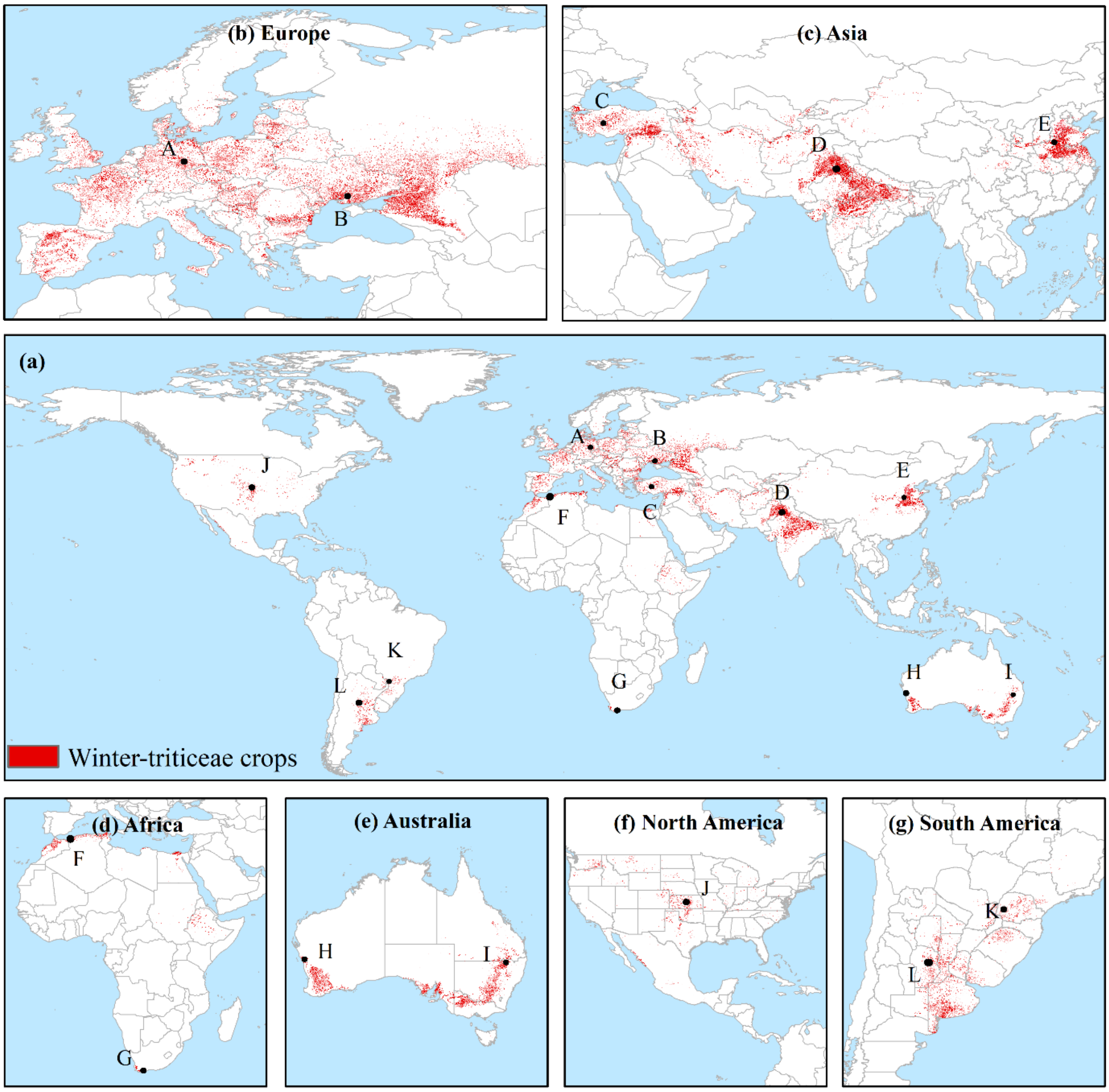

Figure 4: Spatial distribution of winter-triticeae crops in the study area in 2020. (a) shows the distribution of winter-triticeae crops in 66 countries; (b-g) show the zoomed-in maps of Europe, Asia, Africa, Australia, North America and South America, respectively.

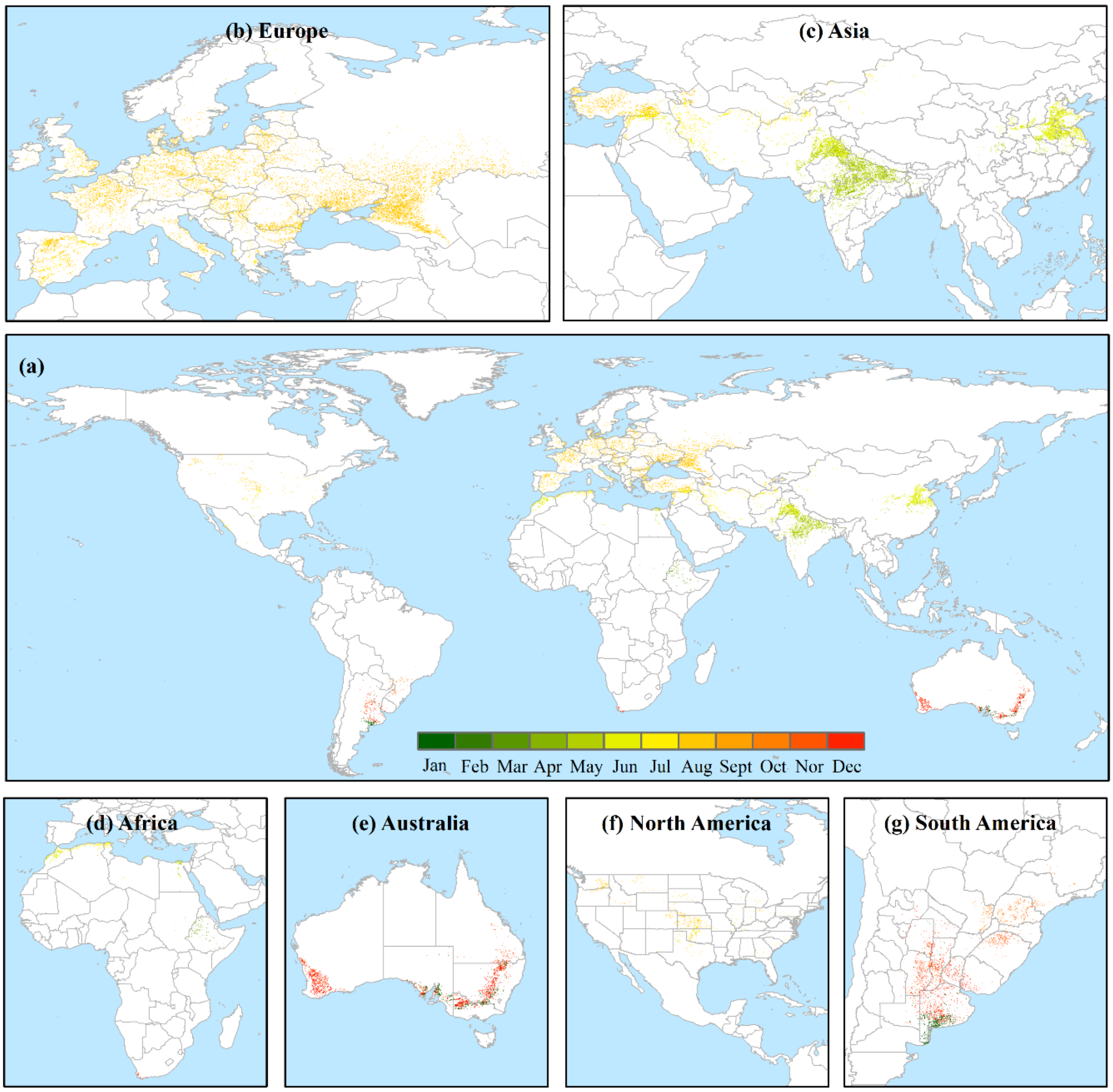

Figure 13: Harvest time of winter-triticeae crops in the study area in 2020.