Food and Agriculture Organization of the United Nations

FAO Statistics Working Paper Series

Issue 25-46

COMPARISON OF WORLDCEREAL WITH FAOSTAT DATA: AN EXPLORATORY ANALYSIS

FAO Statistics Working Paper Series / 25-46

COMPARISON OF WORLDCEREAL WITH FAOSTAT DATA: AN EXPLORATORY ANALYSIS

Francesco N. Tubiello, Carola Fabi, Giulia Conchedda, Leon Casse and Gianfausto Bottini

Food and Agriculture Organization of the United Nations
Rome, 2025

Required citation:
Tubiello, F.N., Fabi, C., Conchedda, G., Casse, L. & Bottini, G. 2025. *Comparison of WorldCereal with FAOSTAT data: an exploratory analysis*. FAO Statistics Working Paper Series, No. 25-46. Rome, FAO.
https://doi.org/10.4060/cd4154en

ISBN 978-92-5-139585-1

Abstract

WorldCereal, a platform developed under the leadership of VITO Remote Sensing and the European Space Agency, provides high-resolution global cropland and crop type maps. This study evaluates the accuracy and reliability of these products using *in situ* data from FAOSTAT. We harmonized diverse datasets to enable fair assessments and employed robust statistical metrics to quantify the performance of maps detailing temporary crops, crop types and irrigation status.

The correlation between FAOSTAT data and WorldCereal data is excellent (R^2 around 0.9) for cereals and irrigated area, and strong for temporary crops (R^2 = 0.75) and maize (R^2 = 0.48). All the assessments showed a higher correlation and lower errors for data considered official than for imputed data. Most cases revealed strong correlations for Europe and South America, and poor ones for Oceania and Central America and the Caribbean that are due to a mismatch in data for Small Island Developing States.

Our findings offer valuable insights into the platform's utility for global agricultural monitoring and highlight its strengths relative to existing technologies. An accompanying interactive dashboard further enriches this work by presenting comprehensive maps and analyses.

Contents

Acknowledgements

We would like to express our gratitude to the reviewers, Zoltan Szantoi from the European Space Agency and Valérie Bizier from the Food and Agriculture Organization of the United Nations (FAO), for their insightful feedback and constructive suggestions, which significantly improved the quality of this manuscript. We also extend our appreciation to Christian Mongeau and Craig Matadeen from the FAO Data Lab, for their technical expertise and invaluable assistance in ensuring the effective deployment of the web application used in this study.

1 Introduction

Geospatial data on crops and land cover/land use enable us to expand our understanding of the spatial distribution of agricultural commodities. This improves our ability to analyse crop productivity and sustainability at multiple scales. At the national and subnational levels, geospatial data are an interesting source of data for regional and country statistics. Current FAO statistics on land use (FAO, 2024a) and crops and livestock products (FAO, 2024b) are data collected through country questionnaires. The completeness of this dataset therefore depends on the reporting of the countries to FAO. Resulting data gaps are currently filled using a variety of methods but are subject to continuously studying alternative methods.

For geospatial data to serve this purpose, a strong agreement between datasets and definitions is required, with data availability at an annual timescale. Previously, six land cover datasets were compared for their agreement on cropland extent for 2020 (Tubiello *et al.*, 2023a) and differences in definitions were found as an important source of disagreement between these datasets. A new country-level cropland area database was constructed from these datasets, showing a global cropland area of 1 500 ± 400 million hectares (Mha), consistent with the 1 560 Mha reported in FAOSTAT (Tubiello *et al.*, 2023b). However, the uncertainty of 400 Mha shows the complexity of mapping cropland globally and the impact of poorly aligned definitions. Three datasets were found to align closely to the FAOSTAT land use item Temporary crops, a subcomponent of cropland (FAO, 2024a), namely ESRI LC, GLAD and WorldCover (Karra *et al.*, 2021; Potapov *et al.*, 2022; Zanaga *et al.*, 2021). WorldCereal is the fourth, recently published dataset containing this land use class, which also maps the spatial distribution of several crops (maize and other cereals) as well as irrigated areas (van Tricht *et al.*, 2023).

The WorldCereal dataset currently consists of 11 products, with 2021 as the reference date. With plans for yearly releases and a dynamic cloud-based system (See *et al.*, 2023), it has the potential to be a valuable source of regional statistics for FAO, especially since the definitions used are similar to FAOSTAT definitions (Table 1). Validation showed high overall accuracy for temporary crops (97.8 percent) and reasonably good overall accuracy for the crop types (82.5 percent), although only limited validation was performed for the individual crops (van Tricht *et al.*, 2023).

Table 1. Overview of WorldCereal products and corresponding FAOSTAT domains and items. Items in a single cell are summed for our comparison

WorldCereal		FAOSTAT	
Item	**Definition**	**Domain**	**Item**
Annual temporary crops	Land used for crops with a less-than-one-year growing cycle which must be newly sown or planted for further production after harvest. Sugar cane, asparagus and cassava are also considered as temporary crops, even though they remain in the field for more than one year. The WorldCereal temporary crop maps exclude perennial crops as well as (temporary) pastures.	Land use	Temporary crops
Winter cereals – Irrigation, Maize main – Irrigation, Maize second – Irrigation	Irrigated areas which are actively cultivated. Areas are defined as irrigated agriculture only if, due to extensive irrigation over a prolonged period, a 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 focus of the irrigated-area mask.	Land use	Cropland area actually irrigated
Maize main, Maize second	Maize by main and second growing season. Maize second is only available for some agroecological zones.	Crops and livestock products	Maize (corn), Green corn (maize), Maize for forage and silage
Winter cereals, Spring cereals	Cereals are the combination of three separate cereal crops all belonging to the Triticeae tribe: wheat, barley and rye. Results are merged and disseminated for the single class 'cereals' by winter and spring seasons.	Crops and livestock products	Wheat, Barley, Rye
Maize main – Active cropland, Maize second – Active cropland, Winter cereals – Active cropland	The area cultivated during the specific season of reference (e.g. all the temporary crops that were detected during the main growing season of maize). The "annual temporary crops" combines all the areas that were detected as active over the entire year of reference, and areas that were detected as active over more than one season are counted only once.	Not compared	

Source: Authors' own elaboration.

2 Methods for the comparison

We compared combinations of items reported in Table 1 as the WorldCereal definitions could be matched to the FAO land use classification and the Central Product Classification (CPC). VITO Remote Sensing, the leading institution of the WorldCereal consortium, shared area statistics by country for each of the 11 products. Variable-specific methods for comparing the four variables are discussed below.

As part of this comparison, we also looked at whether FAOSTAT data were officially reported or estimated, to understand better their correspondence with WorldCereal and the potential of WorldCereal to gap fill the estimated data points. Country statistics were considered "official" if they have in FAOSTAT either the flags A (for official data) and X (figure from international organizations). They were considered "imputed" if they are flagged as E (estimated) or I (imputed) in FAOSTAT. Flags for the land use items land under temporary crops and cropland actually irrigated come from the corresponding 2021 data points in FAOSTAT. Flags for maize are those of the maize (corn) data points, and if no data are available for this item, the flags are taken from the other maize item (green corn or maize for silage and forage) with the largest share. Finally, the flags for the winter and spring cereals aggregate are the flags of the crop with the largest share among wheat, barley and rye (in most cases, wheat).

Correlations in this analysis are shown for the data normalized to the land area to account for the role of large countries, together with the R^2 and the normalized root mean square error (NRMSE), which is calculated as the root mean square error divided by the range of values in the regression.

3 Results

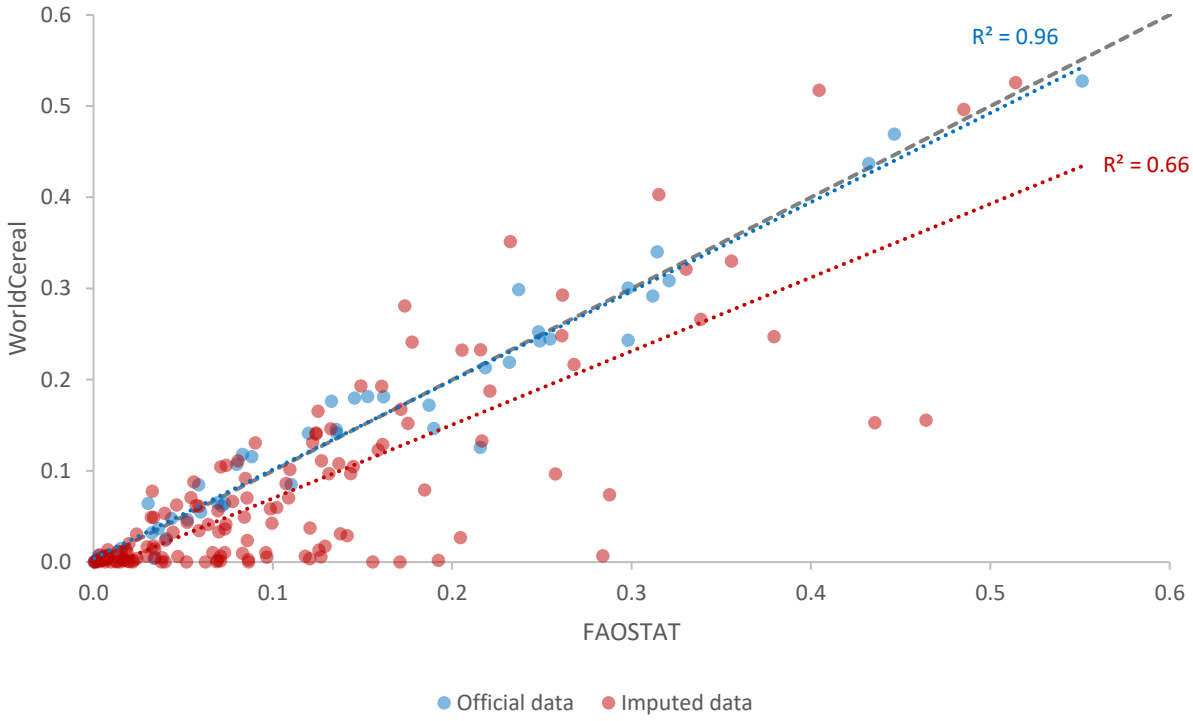
3.1 Temporary crops

FAO defines temporary crops as “Land used for crops with a less-than-one-year growing cycle, which must be newly sown or planted for further production after the harvest. Multiple-cropped areas are counted only once.” While “Some crops that remain in the field for more than one year may also be considered as temporary crops e.g., asparagus, strawberries, pineapples, bananas and sugar cane”, this should be regarded as an exception to adopt in country-specific cases rather than in general. To differentiate between temporary and permanent crops, FAO suggests using the Indicative Crop Classification of the World Census of Agriculture (FAO, 2017).

FAOSTAT land use data covered 221 countries and territories in 2021, compared to 202 for WorldCereal, with 192 common records between the two. At the time of the analysis, 29 countries and territories, including American Samoa, Bahamas, Bermuda, the Cayman Islands, the Cook Islands, French Polynesia, Guam, Kiribati, Liberia and Western Sahara were missing in WorldCereal. The comparison for 2021 is limited to these 192 countries and territories in common between the two sources, split between 51 official and 141 imputed datapoints.

The world total area of temporary crops in 2021 was 1 088 Mha according to FAOSTAT, which is consistent with the 1 129 Mha computed from WorldCereal. A good correlation is determined between FAOSTAT and WorldCereal data overall (R^2 = 0.75; NRMSE 11.7 percent). The results are significantly different between official data (R^2 = 0.96; NRMSE = 4.5 percent) and imputed data (R^2 = 0.66; NRMSE = 14.3 percent) (Figure 1 and Table 2).

Figure 1. Comparison of WorldCereal and FAOSTAT temporary crops data by data flag



Note: Values are normalized by land area.
Source: Authors' own elaboration.

Table 2. Global temporary crops comparison summary statistics

Flag	Number of observations	R^2	NRMSE (percent)
Official	51	0.96	4.5
Imputed	141	0.66	14.3
Total	192	0.75	11.7

Source: Authors' own elaboration.

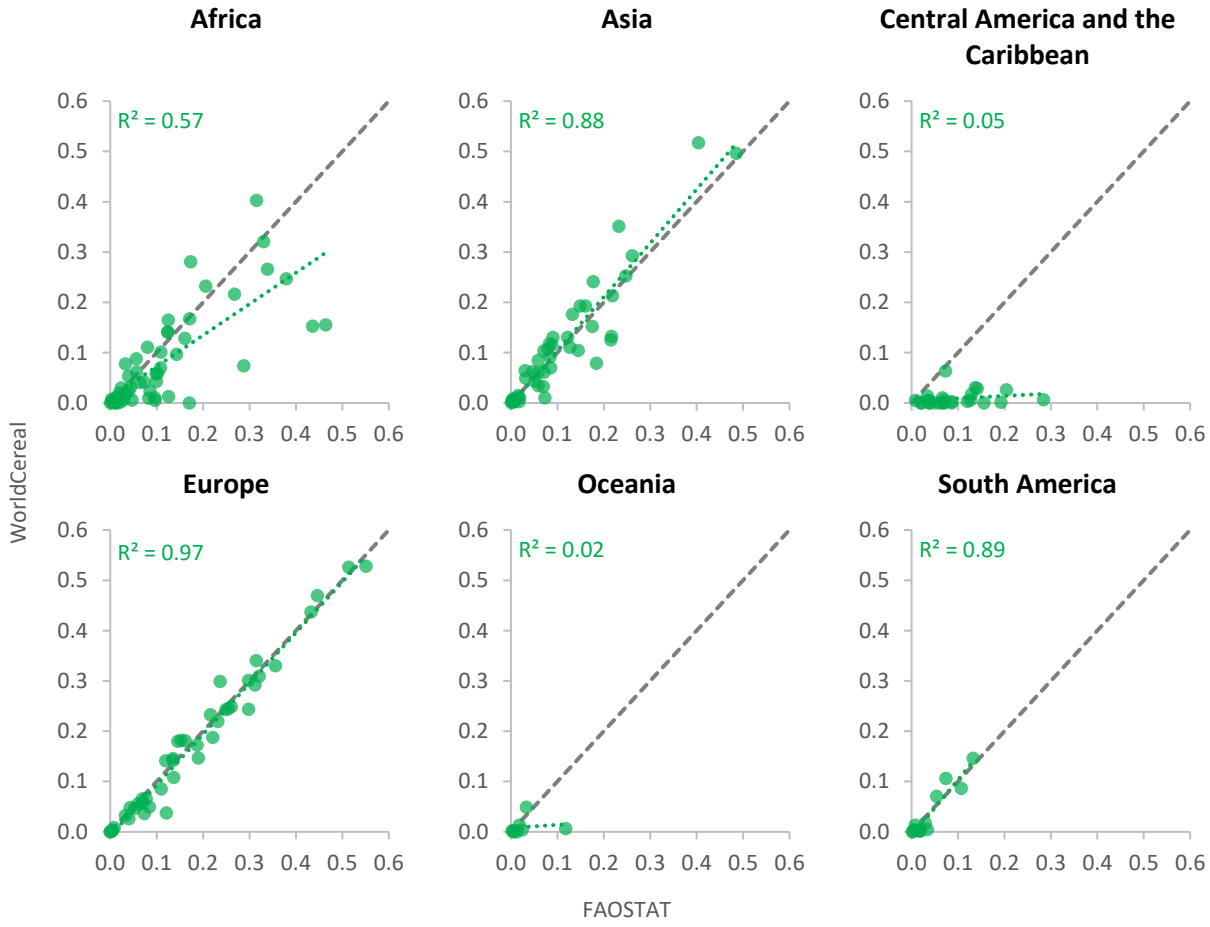
At the regional level[1] (Figure 2 and Table 3), regressions are very close and there is limited variance in Europe[2] (R^2 = 0.97; NRMSE = 4.6 percent) and in Asia (R^2 = 0.88; NRMSE = 8.6 percent). The positive relationship is also strong in South America but with much larger errors (R^2 = 0.89; NRMSE = 13.1 percent). The correspondence remains good in Africa (R^2 = 0.57; NRMSE = 18 percent). In this region, WorldCereal values are overall smaller than in FAOSTAT. The two datasets significantly differ in Central America and in the Caribbean, likely due to the predominance of small islands and small countries, for which the discrepancies between the two sources are overall wider than in larger countries (see below). Data for Oceania are shown but should be taken with caution since they are based on too few country values.[3] The

[1] Only regional-level regressions for the total number of observations are shown in this paper, as there are too few data points for official or imputed data in most regions.
[2] The FAOSTAT aggregate for Europe includes the Russian Federation.
[3] Values for Northern America are not shown since the FAOSTAT aggregate only includes Bermuda, Canada, Greenland, Saint Pierre and Miquelon, and the United States of America.

average share of temporary crops over the total land area ranged between 11 percent in Asia to 3 percent in Oceania.

For Africa, Oceania, South America, and Central America and the Caribbean, the share of official country statistics on temporary crops is rather low (less than 16 percent of data points). In contrast, Europe has the largest share of official statistics (68 percent) and Asia the second largest (Table 3); both regions also have the highest R^2 and lowest NRMSE.

Figure 2. Comparison of WorldCereal and FAOSTAT temporary crops data by region



Note: Values are normalized by land area.
Source: Authors' own elaboration.

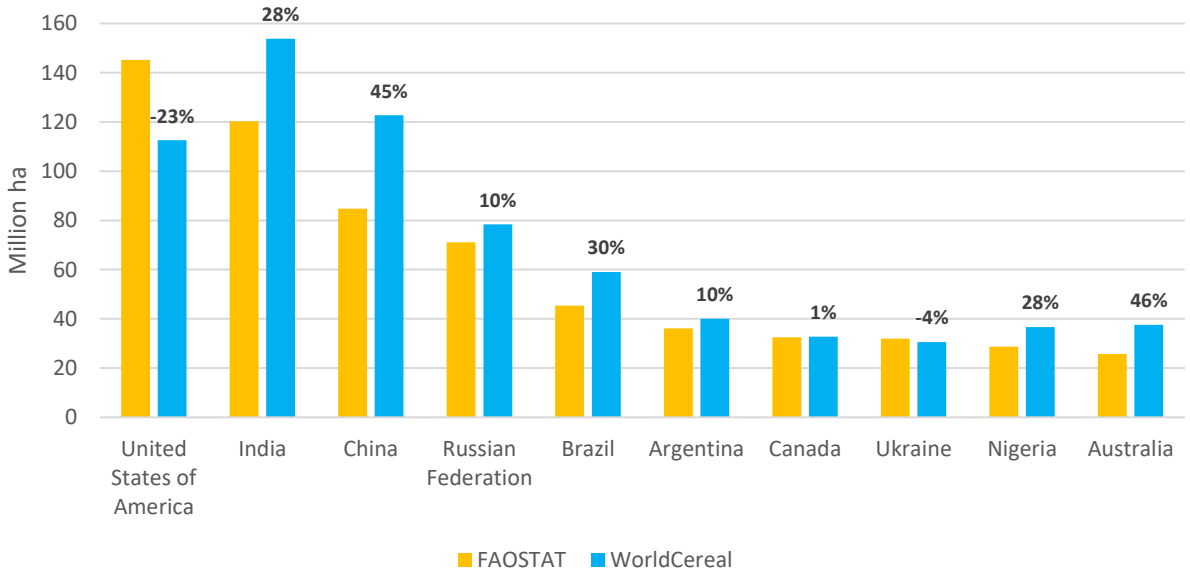
Table 3. Temporary crops comparison summary statistics by region

Region	Number of observations	R^2	NRMSE (percent)	Share of official data (percent)
Africa	51	0.57	18.0	2
Asia	46	0.88	8.6	31.9
Central America and the Caribbean	27	0.05	37.1	3.7
Europe	43	0.97	4.7	68.2
Oceania	8	0.02	35.3	12.5
South America	13	0.89	13.1	15.4

Source: Authors' own elaboration.

Figure 3 shows that WorldCereal and FAOSTAT values for temporary crops are very similar in Ukraine and Canada with a percentage difference less than 5 percent in absolute value, while larger differences are found for Australia (38 Mha for WorldCereal vs 26 Mha for FAOSTAT, a 46 percent difference) and in the United States of America (113 Mha for WorldCereal vs 145 Mha for FAOSTAT, a 23 percent difference in absolute value). Except for the United States of America and Ukraine (although to a lesser extent), WorldCereal produced higher estimates than FAOSTAT in the countries with the largest area of temporary crops. The relative difference is highest in Australia, China and Brazil, and lowest in the United States of America.

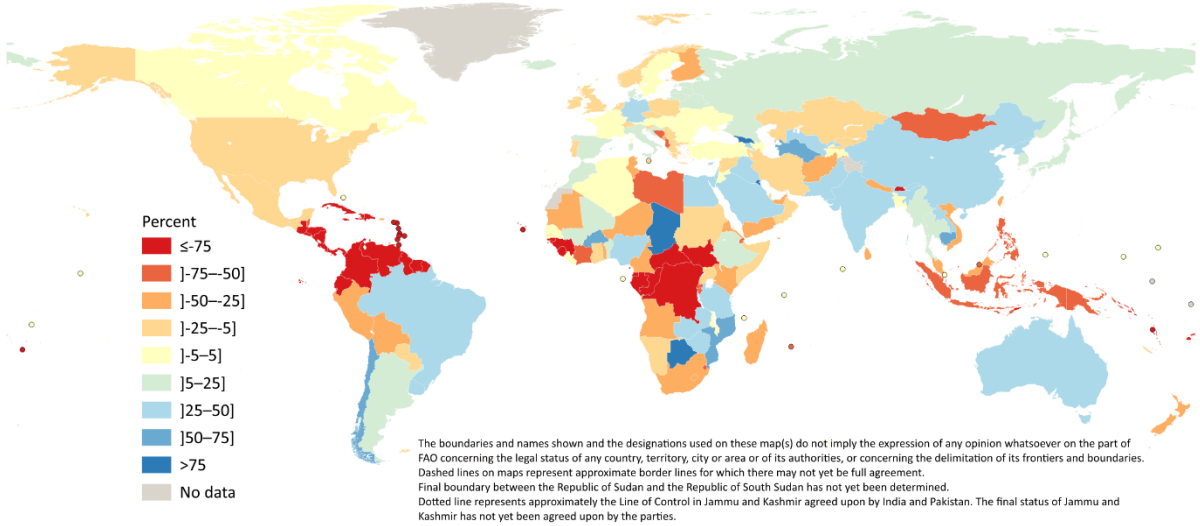
Figure 3. Comparison of WorldCereal and FAOSTAT temporary crops data for the countries with the largest area of temporary crops



Source: Authors' own elaboration.

Figure 4 provides a global overview of the differences between WorldCereal and FAOSTAT data for temporary crops. The coincidence is nearly perfect for Singapore, Bulgaria and Canada. Other countries with discrepancies within ±5 percent are France, Ukraine, Slovakia, Romania, Jordan, Malawi, Belgium, Türkiye, Senegal, Sweden, Hungary, Azerbaijan, the Republic of Moldova, Bangladesh, Algeria, Tajikistan and Austria. There is overall a reasonably good agreement between the two data sources.

Figure 4. Difference between WorldCereal and FAOSTAT temporary crops data



Note: Negative percentages indicate cases where FAOSTAT data are larger than WorldCereal data and positive percentages indicate cases where FAOSTAT data are smaller.
Source: Authors' own elaboration.

3.2 Maize

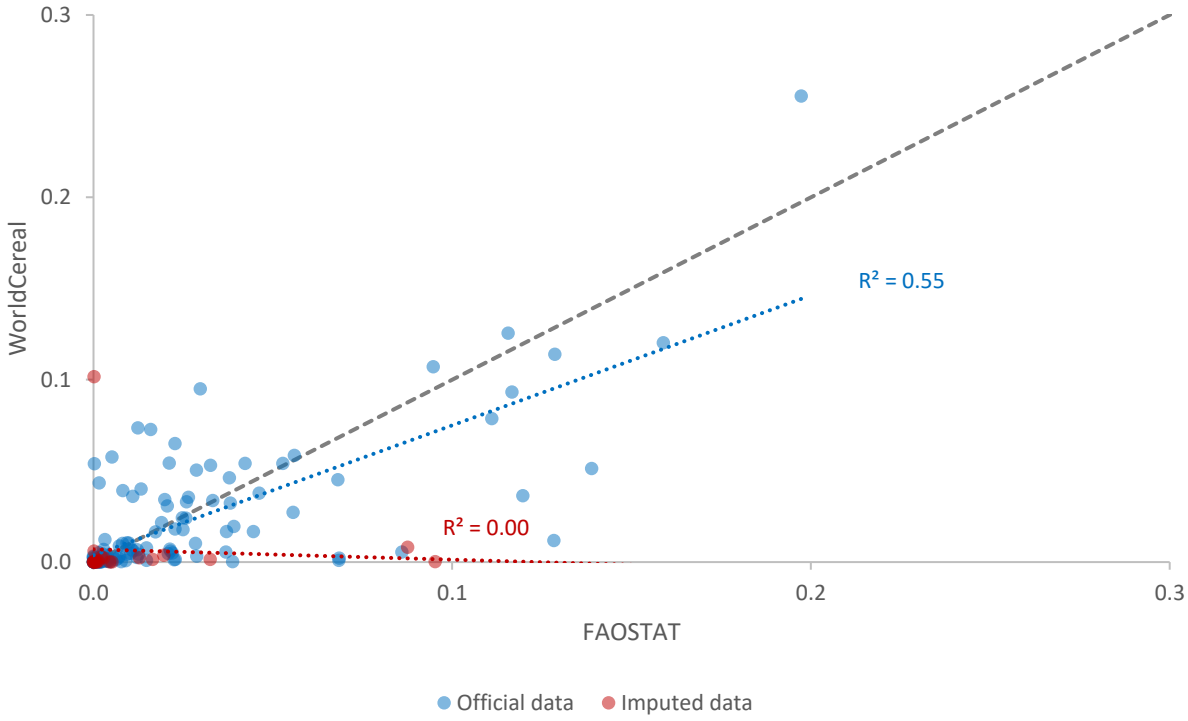
As seen in Table 1, we sum the WorldCereal area statistics of main maize and maize second and compare this to the sum of the FAOSTAT data on harvested areas for maize (corn), green corn, and maize cultivated for forage and silage.[4]

WorldCereal covers 182 countries and territories compared with 160 for FAOSTAT. The two datasets have 149 records in common. The areas covered by WorldCereal and missing in FAOSTAT data include the Plurinational State of Bolivia, Equatorial Guinea, Mongolia and Tunisia. Conversely, FAOSTAT data cover 11 areas that are missing in WorldCereal (including Liberia and several islands). Our comparison between WorldCereal and FAOSTAT data contained 149 records, the majority of which coming from official data.

The global area harvested with maize in 2021 was 219 Mha in FAOSTAT and 213 Mha according to WorldCereal. The area of maize for forage and silage was 12 Mha in 2021 (5.5 percent of the total). A positive linear relationship exists between the two sources. The R^2 for all data flags is close to 0.5 (Figure 5 and Table 4) with limited proportional errors (NRMSE = 13.9 percent). The correlation is higher (R^2 = 0.55; NRMSE = 13.1 percent) for FAOSTAT official data but extremely weak for the imputed and estimated data (R^2 = 0.00; NRMSE = 37.8 percent). WorldCereal values for the imputed and estimated data are usually less than FAOSTAT estimates.

[4] Statistics on maize cultivated for forage and silage are reported by countries to FAO or compiled from other official sources but are not disseminated due to the limited data availability.

Figure 5. Comparison of WorldCereal and FAOSTAT maize data by data flag



Note: Values are normalized by land area.
Source: Authors' own elaboration.

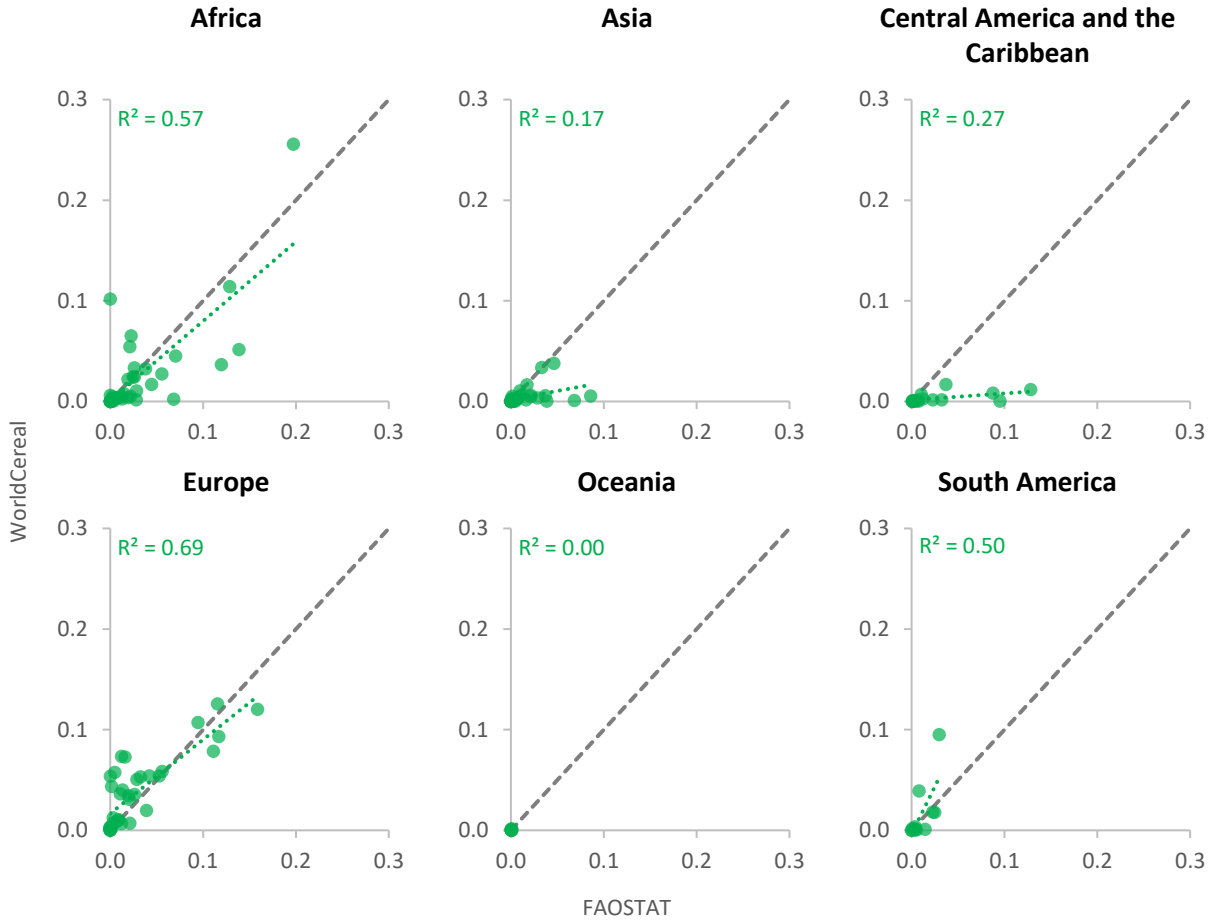
Table 4. Global maize comparison summary statistics

Flag	Number of observations	R^2	NRMSE (percent)
Official	128	0.55	13.1
Imputed	21	0.00	37.8
Total	149	0.48	13.9

Source: Authors' own elaboration.

At the regional level, Europe has the strongest regression and lowest errors variance (R^2 = 0.69; NRMSE = 15.3 percent). The R^2 is also good for Africa (0.57) and for South America (0.5) though the latter shows a much larger error (NRMSE = 72.1 percent). The correspondence is very weak in Asia (R^2 = 0.17) and in Central America and the Caribbean (R^2 = 0.27). FAOSTAT data on maize in Europe are derived exclusively from official sources and show the strongest agreement with WorldCereal, while in Asia the R^2 is much lower even though the region has the second highest share of FAOSTAT official statistics (85.4 percent) (Table 5).

Figure 6. Comparison of WorldCereal and FAOSTAT maize data by region



Note: Values are normalized by land area.
Source: Authors' own elaboration.

Table 5. Maize comparison summary statistics by region

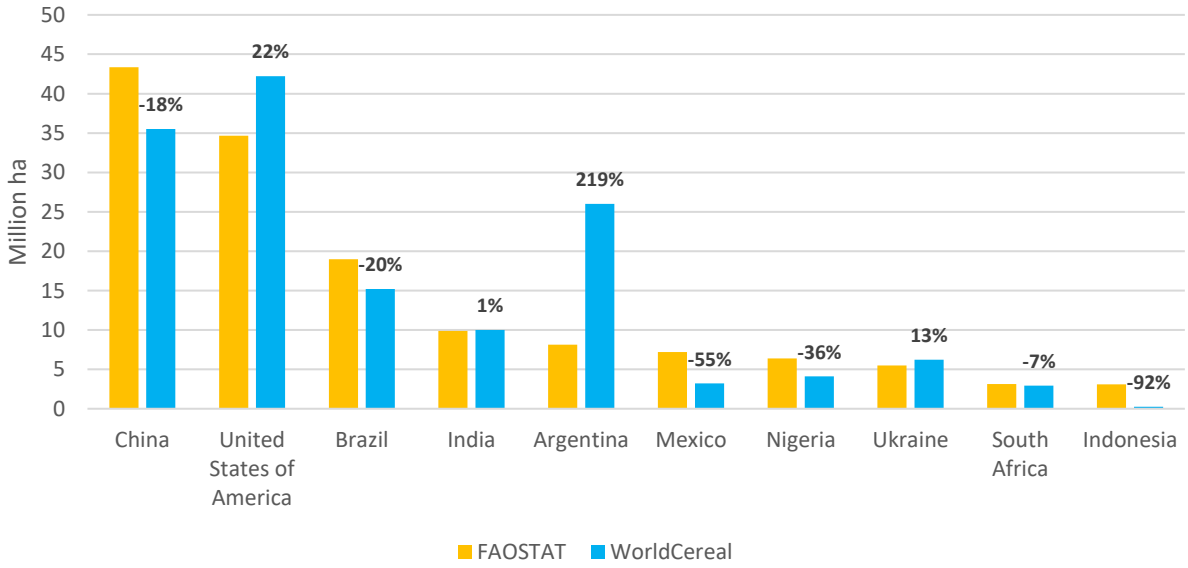
Region	Number of observations	R^2	NRMSE (percent)	Share of official data (percent)
Africa	38	0.57	16.4	85.1
Asia	39	0.17	22.9	85.4
Central America and the Caribbean	15	0.27	35.6	66.7
Europe	36	0.69	15.3	100.0
Oceania	6	0.00	56.7	50.0
South America	12	0.50	72.1	83.3

Source: Authors' own elaboration.

WorldCereal and FAOSTAT data provide different rankings for countries with the largest area of maize. The top thirteen countries according to both sources have ten countries in common (the United States of America, China, Argentina, Brazil, India, Ukraine, the Russian Federation, Nigeria, Mexico and South Africa); the other three are Ethiopia, Mozambique and the United Republic of Tanzania according to WorldCereal, and Angola, the Democratic Republic of the Congo and India according to FAOSTAT. The

values coincided in India (about 10 Mha in each source) and in Namibia (3.4 Mha). Ukraine,[5] Brazil, the United States of America, South Africa and China had relative differences ranging 10–25 percent. The discrepancies are much larger in Argentina, where the area of maize from WorldCereal is three times larger than in FAOSTAT data. Since both WorldCereal and FAOSTAT count these areas multiple times (as both sources include double-cropping and multicropping following similar definitions), the discrepancy is likely due to inaccuracies in one or both sources. FAOSTAT does not provide recent information on maize for forage and silage for Argentina and therefore it is missing from the data; however, data for 2018 suggest that this is only a small fraction of the total maize area (around 10 percent). India, Nigeria and Mexico also exhibit large differences, although in this case the FAOSTAT values are larger than the WorldCereal ones.

Figure 7. Comparison of WorldCereal and FAOSTAT maize data for the countries with the largest area of maize

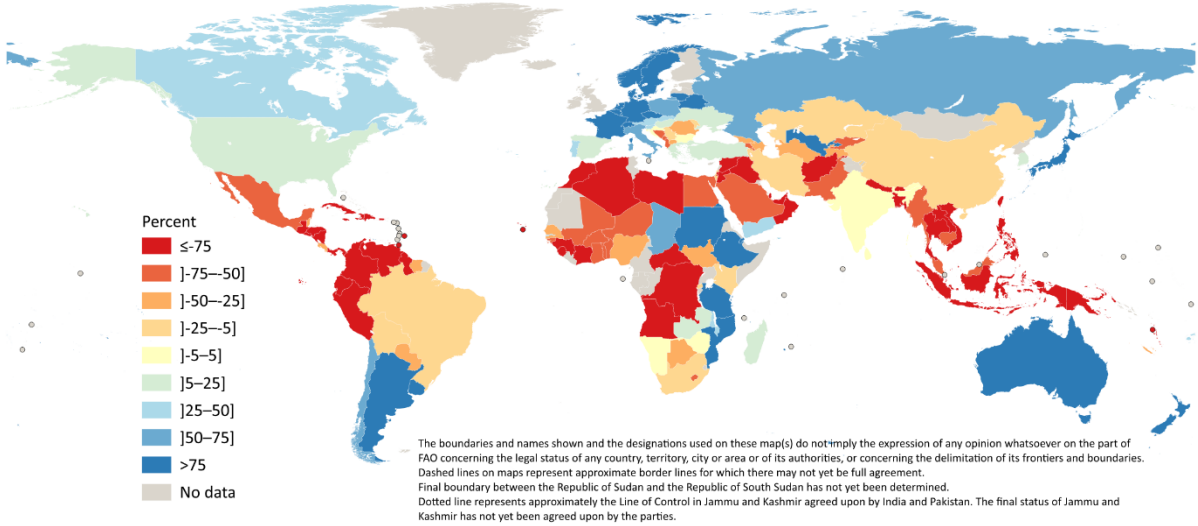


Source: Authors' own elaboration.

As observed earlier, WorldCereal and FAOSTAT maize estimates coincide for India and Namibia. Other countries with discrepancies within ±5 percent are Sri Lanka, Zimbabwe, Bulgaria and Croatia (Figure 8). Overall, WorldCereal maps smaller areas of maize compared to FAOSTAT, especially in Western Africa, Southern Asia and South-eastern Asia, which are likely related to the operational crop calendars and agroecological zones applied in the processing chain of WorldCereal. The spatial distribution and availability of reference in-situ data may also contribute to the observed discrepancies.

[5] Comparisons of WorldCereal data with FAOSTAT statistics for the Russian Federation and Ukraine may be affected by inconsistencies in reporting of data in occupied territories.

Figure 8. Difference between WorldCereal and FAOSTAT maize data



Note: Negative percentages indicate cases where FAOSTAT data are larger than WorldCereal data and positive percentages indicate cases where FAOSTAT data are smaller.
Source: Authors' own elaboration.

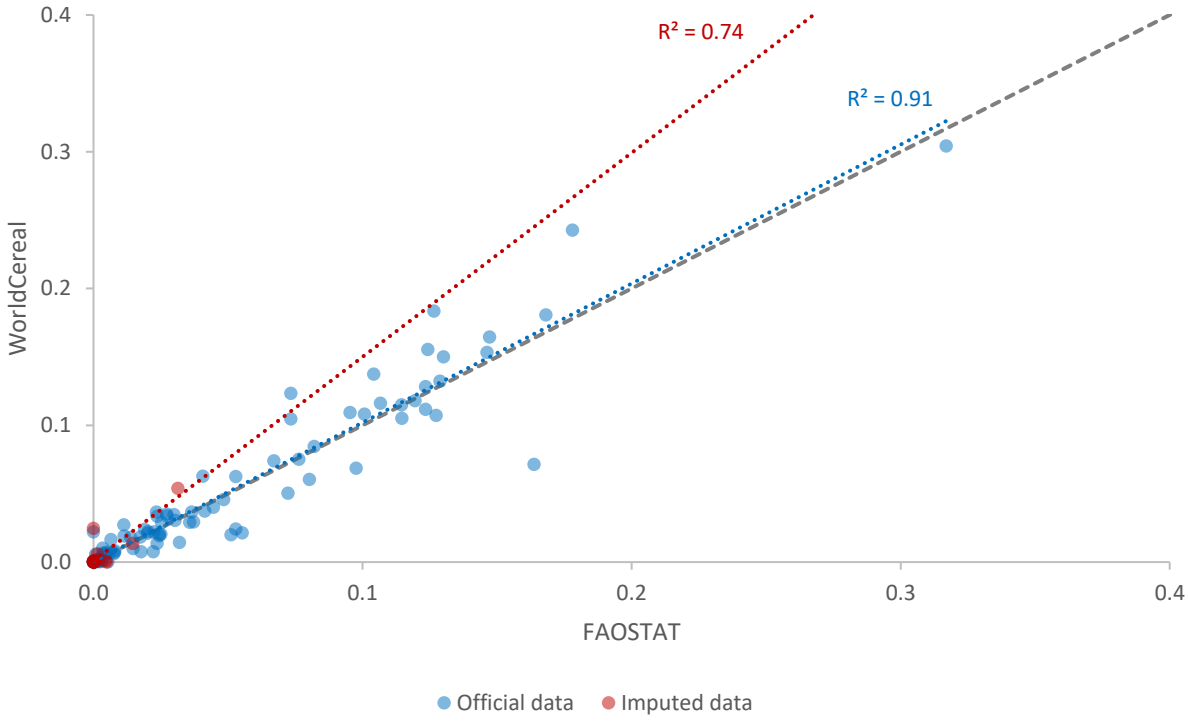
3.3 Cereals

WorldCereal country data for winter cereals and spring cereals are added and compared to the sum of the harvested areas of wheat, barley and rye in FAOSTAT in the absence of the corresponding item aggregate.

FAOSTAT land use data covered 114 countries and territories in 2021 compared to 181 in WorldCereal. All the countries covered by FAOSTAT data are also available in WorldCereal; the 67 countries and territories that do not report to FAO data on the harvested area for any of the three crops include Benin, Burkina Faso, Cambodia, the Central African Republic, Cuba, Ghana, the Lao People's Democratic Republic, Senegal, South Sudan, Togo and Viet Nam. In 2021, the majority of the data points came from official sources (including international organizations).

The global area of the three combined cereals was 278 Mha in WorldCereal and virtually identical in FAOSTAT (274 Mha), indicating that the missing countries mentioned above likely play a minor role. Globally, the area harvested with wheat accounted for 80 percent of the total in 2021, barley for 18 percent and rye for the remaining 2 percent. A very strong positive linear relationship exists between the two sources, as the R^2 for all data flags is 0.91, which is slightly higher than with official data (Figure 9 and Table 6). The correlation is very high, with a small error for FAOSTAT official data (R^2 = 0.91; NRMSE = 5.7 percent); it is strong (R^2 = 0.74) but with a larger error variance (NRMSE = 26.4 percent) for the imputed data.

Figure 9. Comparison of WorldCereal and FAOSTAT cereals data by data flag



Note: Values are normalized by land area.
Source: Authors' own elaboration.

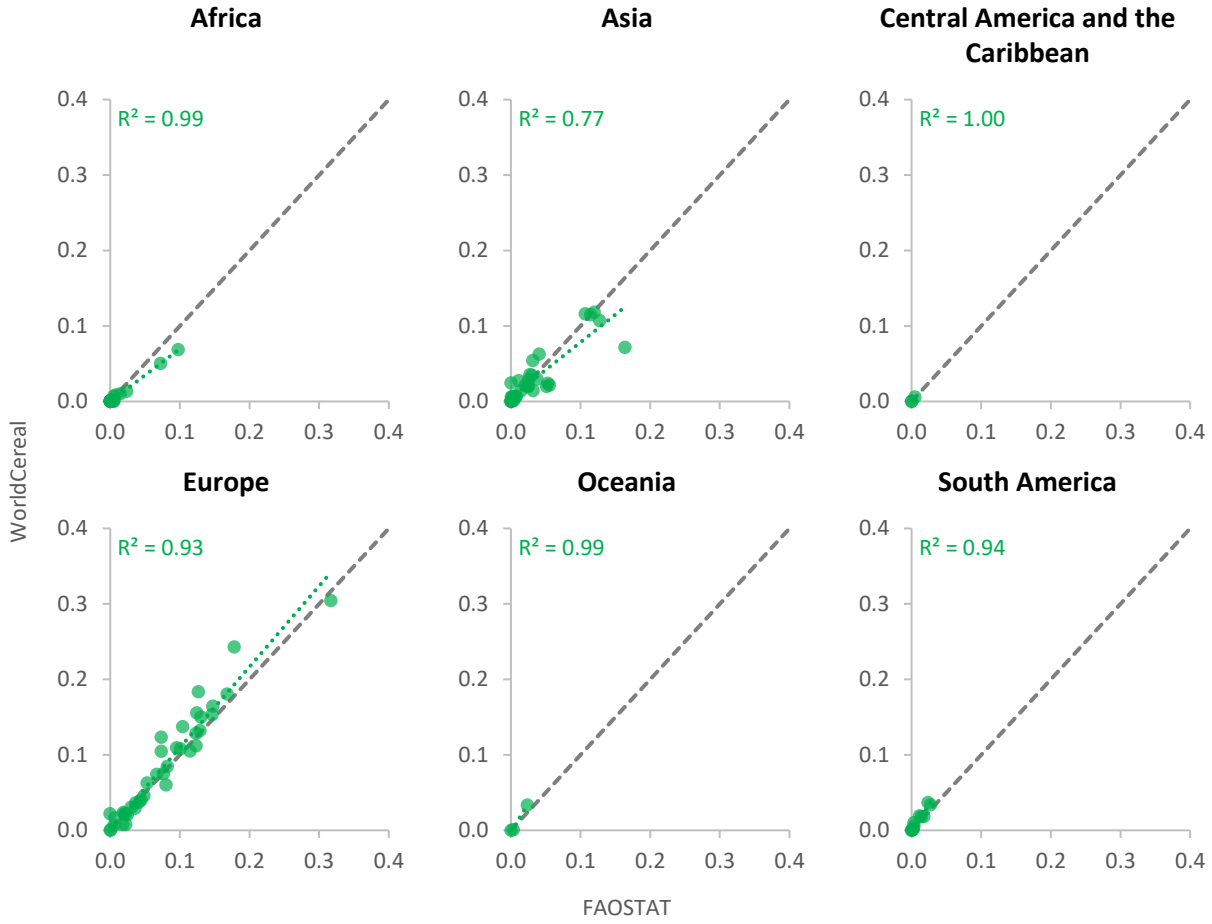
Table 6. Global cereals comparison summary statistics

Flag	Number of observations	R^2	NRMSE (percent)
Official	97	0.91	5.7
Imputed	17	0.74	26.4
All	114	0.91	5.3

Source: Authors' own elaboration.

The two sources show very high correlations and small relative errors in Africa (R^2 = 0.99; NRMSE = 8.3 percent) and Europe (R^2 = 0.93; NRMSE = 6.4 percent) (Figure 10 and Table 7). Results are also good in Asia but with larger errors (R^2 = 0.77; NRMSE = 12.6 percent). The comparisons for Central America and the Caribbean, and Oceania are based on too few data points to yield meaningful results. Europe and Asia had on average the largest proportional areas of cereals over land (4 percent in both regions) whereas Africa had on average a 1 percent share. Most FAOSTAT country statistics on cereals, like maize, come from official sources. Europe had the largest share of official data (100 percent), while South America, Asia and Africa had official data for more than 75 percent of the countries. Only Central America and the Caribbean had less than 50 percent of official data points (Table 7).

Figure 10. Comparison of WorldCereal and FAOSTAT cereals data by region



Note: Values are normalized by land area.
Source: Authors' own elaboration.

Table 7. Cereals comparison summary statistics by region

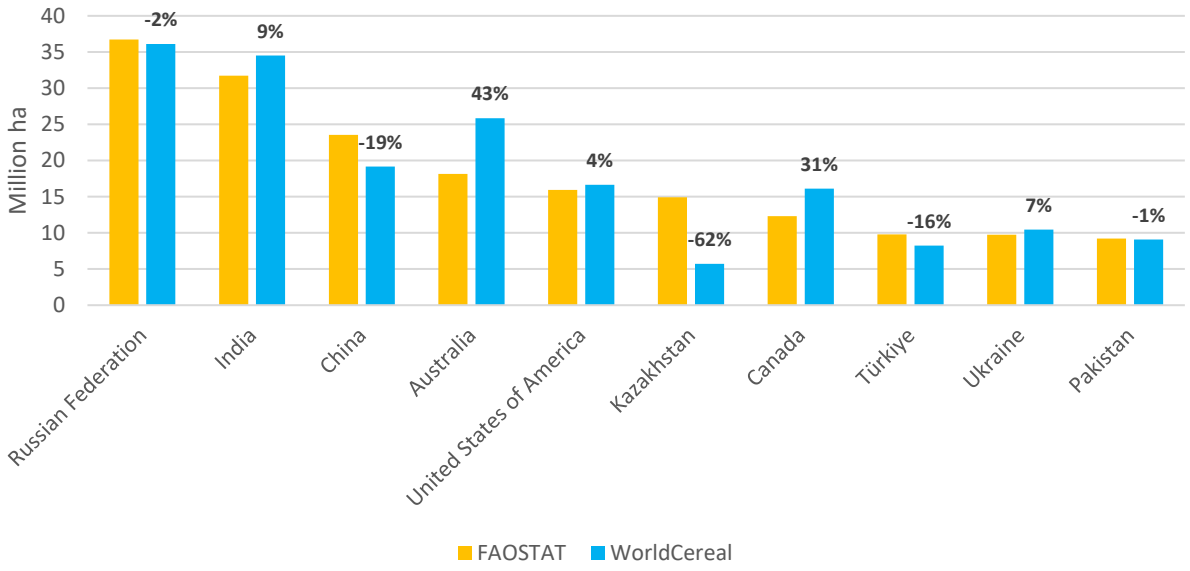
Region	Number of observations	R^2	NRMSE (percent)	Share of official data (percent)
Africa	23	0.99	8.3	75.0
Asia	34	0.77	12.6	86.1
Central America and the Caribbean	3	1.00	16.2	33.3
Europe	39	0.93	6.4	100.0
Oceania	3	0.99	25.3	66.7
South America	12	0.94	19.5	90.0

Source: Authors' own elaboration.

Figure 11 shows the WorldCereal and FAOSTAT values for the ten countries with the largest area of cereals, based on the FAOSTAT ranking. The data are very similar for Pakistan (9.1 Mha and 9.2 Mha for WorldCereal and FAOSTAT, respectively), the Russian Federation (36.1 Mha for WorldCereal and 36.8 for FAOSTAT), the United States of America (16.6 Mha for WorldCereal and 16 Mha for FAOSTAT) and India (34.5 Mha for WorldCereal and 31.7 Mha for FAOSTAT). The relative differences are larger for Kazakhstan and in Australia. WorldCereal estimates are much lower than FAOSTAT data for Kazakhstan (−62 percent)

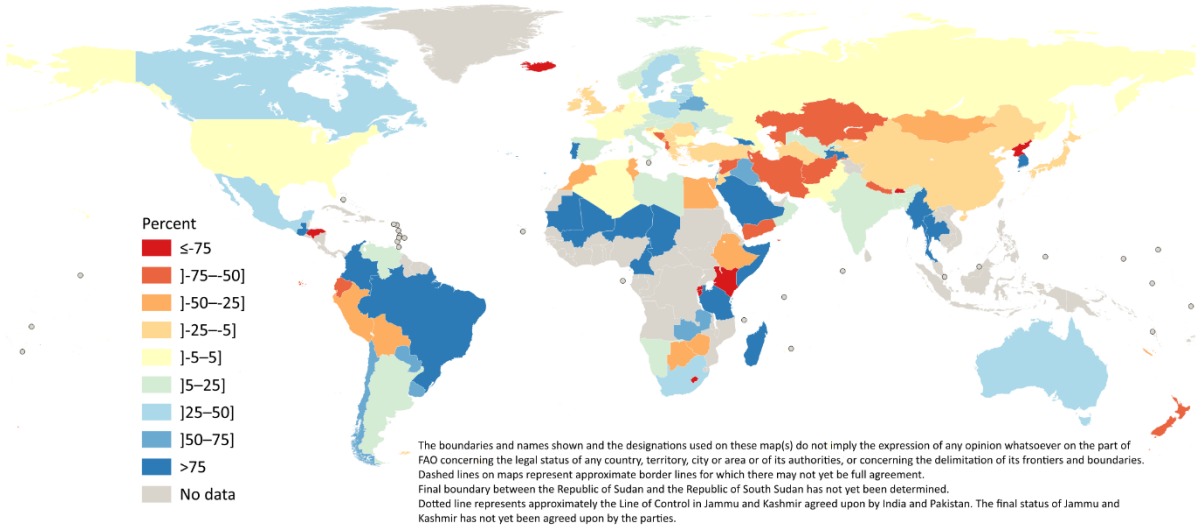
whereas they are 43 percent higher than FAOSTAT data for Australia. Global results and differences by countries are shown in Figure 12. WorldCereal has larger estimates than FAOSTAT in most of Latin America and in Sahel countries but lower in Central Asia, Eastern Africa and Middle Africa.

Figure 11. Comparison of WorldCereal and FAOSTAT cereals data for the countries with the largest area of cereals



Source: Authors' own elaboration.

Figure 12. Difference between WorldCereal and FAOSTAT cereals data



Note: Negative percentages indicate cases where FAOSTAT data are larger than WorldCereal data and positive percentages indicate cases where FAOSTAT data are smaller.
Source: Authors' own elaboration.

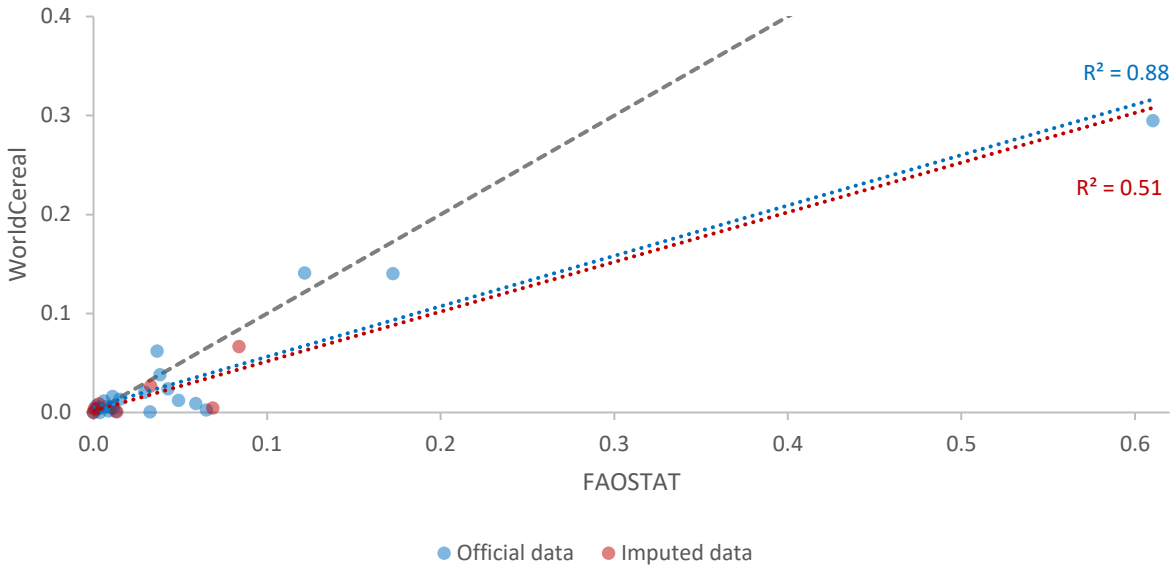
3.4 Irrigation

We sum for this comparison the seasonal country statistics of irrigated areas from WorldCereal (winter cereals – irrigation, maize main – irrigation and maize second – irrigation). This implies that if the same pixel is irrigated over multiple seasons its area is counted more than once. Conversely, the reporting on the variable cropland area actually irrigated should count the area irrigated only once (FAO, n.d.). Hence, WorldCereal values in this comparison are potentially larger than the corresponding FAOSTAT values for cropland area actually irrigated. On the other hand, the FAOSTAT variable also includes irrigated areas of permanent crops and temporary meadows and pastures, which are not mapped by WorldCereal. While WorldCereal only includes areas irrigated for extensive periods to reach significant crop yield (van Tricht *et al.*, 2023), FAOSTAT includes any type of irrigation. Furthermore, due to the incomplete data coverage, global and regional aggregates for the variable cropland area actually irrigated are not available in FAOSTAT.

Globally, WorldCereal maps 231 Mha of temporary crops that were actively irrigated at least one season in 2021 in 182 countries. There are 32 countries and territories with common records in the two datasets. For these areas both datasets showed a similar total (FAOSTAT: 23 Mha; WorldCereal: 27 Mha), despite the limitations in the comparability of data as explained above.

For this limited set of data, the overall correlation between these two sources is particularly good and with limited proportional error (R^2 = 0.87; NRMSE = 9.9 percent) (Figure 13 and Table 8).

Figure 13. Comparison of WorldCereal and FAOSTAT irrigated area data by data flag



Note: Values are normalized by land area.
Source: Authors' own elaboration.

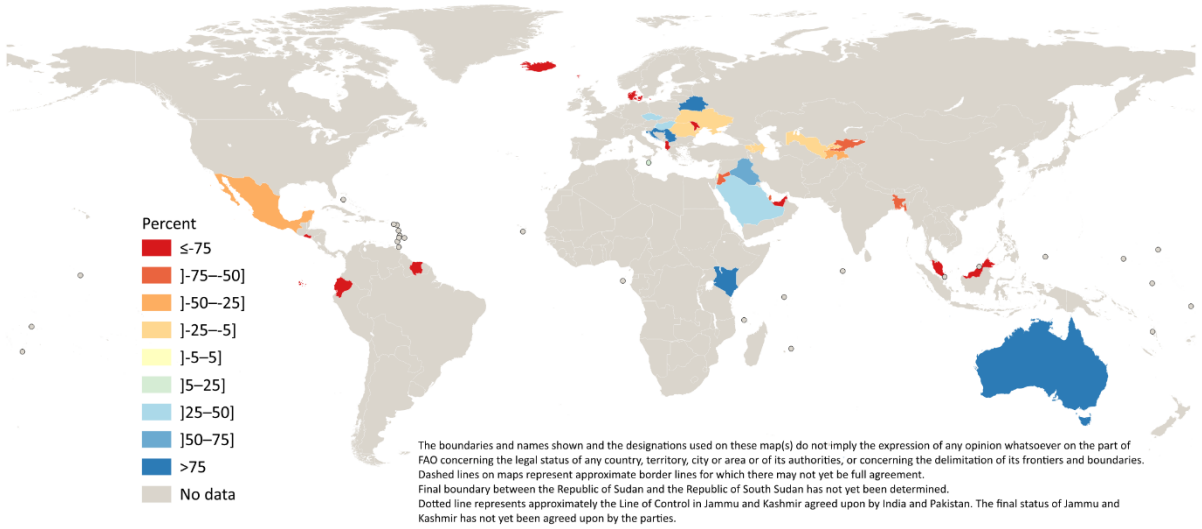
Table 8. Global irrigated area comparison summary statistics

Flag	Number of observations	R^2	NRMSE (percent)
Official	25	0.88	10.9
Imputed	7	0.51	30.7
Total	32	0.87	9.9

Source: Authors' own elaboration.

Figure 14 shows that the relative differences between the two sources are very large. However, conclusions on this variable cannot be definitive since the sum of WorldCereal country data may count the same area several times. A pixel-level analysis should first be done to extract annual country statistics that match more closely the FAOSTAT data for cropland area actually irrigated.

Figure 14. Difference between WorldCereal and FAOSTAT irrigated area data



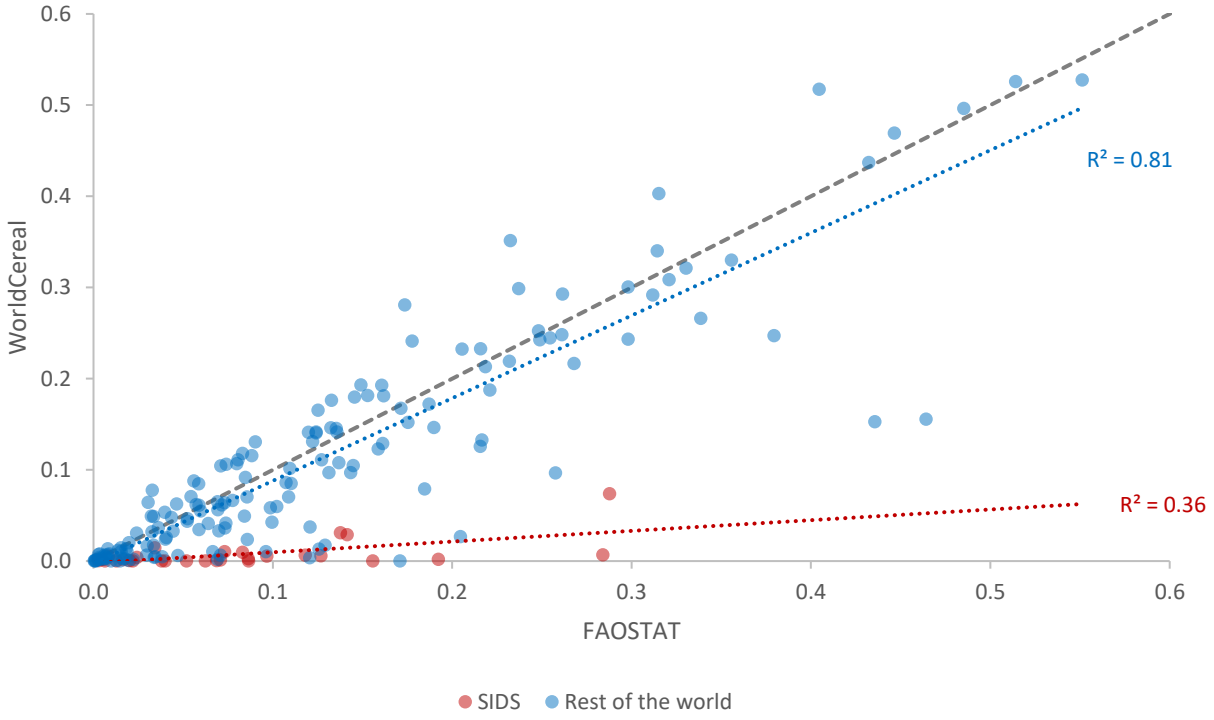
Note: Negative percentages indicate cases where FAOSTAT data are larger than WorldCereal data and positive percentages indicate cases where FAOSTAT data are smaller.
Source: Authors' own elaboration.

4 Correlation for Small Island Developing States

The analysis showed in most cases poor correlations between FAOSTAT and WorldCereal data in Oceania and Central America and the Caribbean. Many Small Island Developing States (SIDS) are in these regions, potentially contributing to this disagreement. To illustrate this, Figure 15 shows the correlations between WorldCereal and FAOSTAT temporary crops data for SIDS and for the rest of the world: the correlation is indeed much weaker and the error variance higher for SIDS (R^2 = 0.36; NRMSE = 34.3 percent) compared to the rest of the world (R^2 = 0.81; NRMSE = 10.1 percent) (Table 9). WorldCereal appears to considerably underestimate the temporary crop area compared to the areas reported in FAOSTAT for SIDS. Non-SIDS countries are close to the optimal 1:1 line, indicating better agreement between the two sources.

Table 9 shows that both areas have a reasonably low share of official FAOSTAT statistics (<33 percent), however the contribution of official data to SIDS is considerably lower, potentially contributing to the poorer agreement with WorldCereal.

Figure 15. Comparison of WorldCereal and FAOSTAT temporary crops data for SIDS and non-SIDS



Note: Values are normalized by land area.
Source: Authors' own elaboration.

Table 9. Temporary crops area comparison summary statistics for SIDS and non-SIDS

Region	Number of observations	R^2	NRMSE (percent)	Share of official data (percent)
Small Island Developing States	30	0.36	34.3	31.4
Rest of the world	162	0.81	10.1	3.0

Source: Authors' own elaboration.

5 Conclusion

The comparison of WorldCereal data with FAOSTAT data show very promising results for temporary crops and cereals, which indicates that WorldCereal, together with other remote-sensing derived products (Tubiello *et al.*, 2023b), could be included in the imputation and estimation processes for these variables. WorldCereal could be especially valuable since it has a nearly complete global coverage and reporting for temporary crops is limited. However, caution is needed for SIDS.

Our analysis shows a very good correlation between WorldCereal and FAOSTAT official data, and a poor correlation between WorldCereal and FAOSTAT estimated data, suggesting that more investigation may be needed on internal estimation methods. Clear spatial and regional patterns in the differences between the two sources may be worth investigating with the data providers to better understand the reasons for these discrepancies. Case-by-case analysis may provide additional insights worth exploring.

While this analysis uses tabular data, more conclusive results on the comparison of WorldCereal and FAOSTAT data for the areas irrigated would require prior pixel-level analysis, which could be done at a later stage of the research.

6 References

FAO. n.d. *Questionnaire on land use, irrigation and agricultural practices*. Rome. https://www.fao.org/fileadmin/user_upload/faoweb/statistics/questionnaires/FAO_LAND_QUEST_EN.xlsx

FAO. 2017. *World Programme for the Census of Agriculture 2020. Volume I – Programme, concepts and definitions*. Rome. https://openknowledge.fao.org/handle/20.500.14283/i4913e

FAO. 2024a. FAOSTAT: Land use. [Accessed December 2024]. https://www.fao.org/faostat/en/#data/RL. Licence: CC-BY-4.0.

FAO. 2024b. FAOSTAT: Crops and livestock products. [Accessed December 2024]. https://www.fao.org/faostat/en/#data/QCL. Licence: CC-BY-4.0.

Karra, K., Kontgis, C., Statman-Weil, Z., Mazzariello, J.C., Mathis, M. & Brumby, S.P. 2021. Global land use/land cover with Sentinel 2 and deep learning. *IEEE international geoscience and remote sensing symposium IGARSS*, 4704–4707, https://doi.org/10.1109/IGARSS47720.2021.9553499

Potapov, P., Turubanova, S., Hansen, M.C., Tyukavina, A., Zalles, V., Khan, A., Song, X.-P., Pickens, A., Shen, Q. & Cortez, J. 2022. Global maps of cropland extent and change show accelerated cropland expansion in the twenty-first century. *Nature Food*, 3, 19–28. https://doi.org/10.1038/s43016-021-00429-z

See, L., Gilliams, S., Conchedda, G., Degerickx, J., van Tricht, K., Fritz, S., Lesiv, M., Laso Bayas, J. C., Rosero, J., Tubiello, F.N. & Szantoi, Z. 2023. Dynamic global-scale crop and irrigation monitoring. *Nature Food* 4, 736–737. https://doi.org/10.1038/s43016-023-00841-7

Tubiello, F.N., Conchedda, G., Casse, L., Pengyu, J., Zhongxin, C., De Santis, G., Fritz, S. & Muchoney, D. 2023a. Measuring the world's cropland area. *Nature Food*, 4, 30–32. https://doi.org/10.1038/s43016-022-00667-9

Tubiello, F.N., Conchedda, G., Casse, L., Hao, P., De Santis, G., & Chen, Z. 2023b. A new cropland area database by country circa 2020. *Earth System Science Data*, 15, 4997–5015. https://doi.org/10.5194/essd-15-4997-2023

Van Tricht, K., Degerickx, J., Gilliams, S., Zanaga, D., Battude, M., Grosu, A., Brombacher, J., Lesiv, M., Bayas, J.C.L., Karanam, S., Fritz, S., Becker-Reshef, I., Franch, B., Mollà-Bononad, B., Boogaard, H., Pratihast, A.K., Koetz, B. & Szantoi, Z. 2023. WorldCereal: a dynamic open-source system for global-scale, seasonal, and reproducible crop and irrigation mapping. *Earth System Science Data*, 15, 5491–5515. https://doi.org/10.5194/essd-15-5491-2023

Zanaga, D., Van De Kerchove, R., De Keersmaecker, W., Souverijns, N., Brockmann, C., Quast, R., Wevers, J., Grosu, A., Paccini, A., Vergnaud, S., Cartus, O., Santoro, M., Fritz, S., Georgieva, I., Lesiv, M., Carter, S., Herold, M., Li, L., Tsendbazar, N.E. & Arino, O. 2021. ESA WorldCover 10 m 2020 v100 (Version v100). In: *Zenodo*. [Cited December 2024]. https://doi.org/10.5281/zenodo.5571936

Contact:

Statistics Division – Economic and Social Development
FAO-statistics@fao.org
www.fao.org/statistics/resources/en

Food and Agriculture Organization of the United Nations
Rome, Italy

ISBN 978-92-5-139585-1

CD4154EN/1/02.25