



OPEN

DATA DESCRIPTOR

County-level intensity of carbon emissions from crop farming in China during 2000–2019

Cheng Li^{1,5}, Junwen Jia^{2,3,5}, Fang Wu², Lijun Zuo⁴ & Xuefeng Cui²✉

Agriculture is an important contributor to global carbon emissions. With the implementation of the Sustainable Development Goals of the United Nations and China's carbon neutral strategy, accurate estimation of carbon emissions from crop farming is essential to reduce agricultural carbon emissions and promote sustainable food production systems in China. However, previous long-term time series estimates in China have mainly focused on the national and provincial levels, which are insufficient to characterize regional heterogeneity. Here, we selected the county-level administrative district as the basic geographical unit and then generated a county-level dataset on the intensity of carbon emissions from crop farming in China during 2000–2019, using random forest regression with multi-source data. This dataset can be used to delineate spatio-temporal changes in carbon emissions from crop farming in China, providing an important basis for decision makers and researchers to design agricultural carbon reduction strategies in China.

Background & Summary

The rapid increase in greenhouse gas emissions such as CO₂, CH₄, and N₂O has further exacerbated the global warming process, seriously threatening the sustainable development of human society^{1–3}. As a basic industry of the national economy, agriculture is easily affected by global warming, which can have a significant impact on crop farming and contribute further to the instability of crop production and supply^{4,5}. Meanwhile, crop farming is also one of the major sources of global greenhouse gas emissions of CO₂ and non-CO₂, and therefore acts as a crucial role in global warming^{6–8}. China is the world's leading agricultural producer⁹, which feeds 20% of the global population with only 7% of global croplands by securing arable land and increasing fertilizer and pesticide inputs over the past decades¹⁰. However, the fact that these main measures can further increase carbon emissions has been a great concern^{11–13}. With the implementation of the Sustainable Development Goals of the United Nations and China's carbon neutral strategy^{14,15}, a detailed understanding of the temporal and spatial changes in China's carbon emissions from crop farming is essential to reduce agricultural carbon emissions and promote sustainable food production systems in China.

There are several carbon emissions datasets related to crop farming at different spatial levels around the world. The Food and Agriculture Organization of the United Nations (FAO) provided a complete and coherent time series of emission statistics at national scale since 1961¹⁶. Furthermore, Carlson *et al.*¹⁷ developed the 5-arc-minute crop-specific circa 2000 estimates of carbon emissions, reporting the effects of rice paddy management, peatland draining, and nitrogen fertilizer on CH₄, CO₂ and N₂O. In China, Zuo *et al.*¹⁸ calculated carbon emissions from agricultural management (e.g., fertilizer application, rice cultivation, peatland draining) and cropland changes for 1987, 2000, and 2010. Zhang *et al.*¹⁹ calculated provincial-level carbon emissions of grain production from 1997 to 2020, including emissions from material inputs, straw burning, fertilizer application, and rice cultivation. In comparison, Liang *et al.*²⁰ reconsidered primary and secondary emissions from various agricultural activities involved in crop farming, in which crop residue open burning, rice cultivation, cropland change, cropland emissions, machinery use, nitrogen fertilizer production, and pesticide production were

¹Department of Ecology, School of Plant Protection / Joint International Research Laboratory of Agriculture and Agri-Product Safety of the Ministry of Education, Yangzhou University, Yangzhou, 225009, China. ²School of Systems Science, Beijing Normal University, Beijing, 100875, China. ³School of Earth and Environmental Sciences, Cardiff University, Cardiff, CF10 3AT, United Kingdom. ⁴Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, 100093, China. ⁵These authors contributed equally: Cheng Li, Junwen Jia. ✉e-mail: xuefeng.cui@bnu.edu.cn

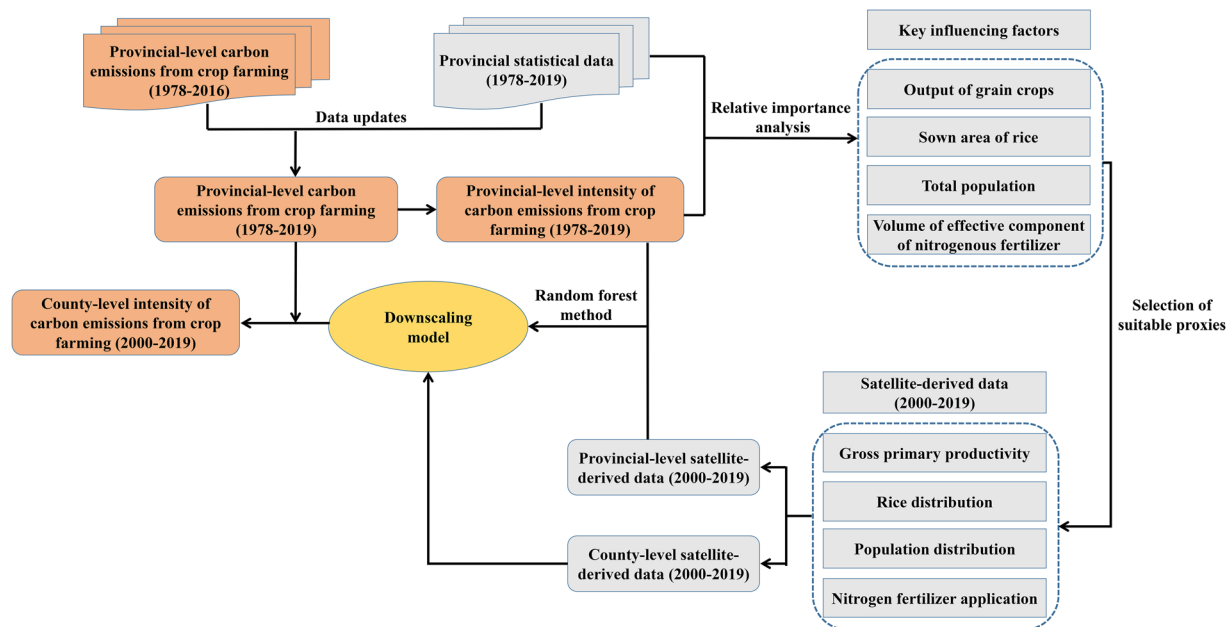


Fig. 1 Workflow of the study.

included, and then calculated provincial-level carbon emissions from 1978 to 2016 in China. These existed studies have mainly focused on carbon emissions at the national and provincial levels, and have taken into account different agricultural activities, parameters and study periods in their calculations, making comparisons difficult in China^{18–21}. A recent study indicated that the spatial heterogeneity of agricultural emissions highlighted necessity of tailored regional mitigation strategies in China²². Although counties are the most basic government units in China and their important role in the implementation of national policies²³, few studies have been conducted at the county level due to the low availability of data sources in China.

To fill those knowledge gaps, this study contributes the following work to this research area: (1) we updated provincial-level carbon emissions from crop farming during 1978–2019 in China based on Liang *et al.*²⁰, and calculated the corresponding intensity of carbon emissions; and (2) we used the random forest-based downscaling method to calculate the county-level intensity of carbon emissions from crop farming during 2000–2019 in China. The results of this study can help to delineate the spatio-temporal changes in carbon emissions from crop farming in China and provide an important basis to design China's agricultural carbon reduction strategies.

Methods

Calculation of the intensity of provincial-level carbon emissions from crop farming. Although Liang *et al.*²⁰ provided a comprehensive and long-term dataset on provincial-level carbon emissions from crop farming in China during 1978–2016 (Figure S1), there is a great uncertainty in the activity level data for cropland change^{24,25}, derived from satellite land use/cover data for 1980, 1990, 2000, 2010, and 2015 in their work²⁶ (Figure S1c). Moreover, there are also large differences in cropland changes compared to those from other commonly used land use/cover datasets (Figure S2). In order to reduce uncertainty, we consider emissions from the following six agricultural activities except for emissions from cropland change, including crop residue open burning, rice cultivation, cropland emissions, machinery use, nitrogen fertilizer production, and pesticide production.

According to data sources, emission factors, and calculation methods provided by Liang *et al.*²⁰, we first updated provincial-level carbon emissions from crop farming during 1978–2019 in China. We then further calculated the intensity of provincial-level carbon emissions from crop farming (Unit: ton CO₂-eq/km²) based on the land area of each province (Fig. 1).

Selection of important indicators affecting the intensity of carbon emissions from crop farming. Based on provincial statistics from the National Bureau of Statistics of China, we selected 12 indicators that could affect carbon emissions from crop farming (Table 1). Furthermore, the results per unit area were calculated for these indicators based on the land area of each province, which were set as independent variables. In this study, the intensity of provincial-level carbon emissions from crop farming was taken as the dependent variable, a random forest method was used to identify key factors that influence the intensity of provincial-level carbon emissions from crop farming^{27–29} (Fig. 1). As shown in Table 1, the output of grain crops, sown area of rice, total population, and volume of effective component of nitrogenous fertilizer had significant impacts on the intensity of provincial-level carbon emissions from crop farming, and their overall magnitudes of relative importance reached 95.76%. For these important influencing factors (Table 1), we selected suitable proxies based on available satellite-derived datasets in this work^{30–33}.

Indicators	Relative importance (%)	Suitable proxy indicators
Output of grain crops	2.40	Gross primary productivity ³⁰
Sown area of rice	36.44	Rice distribution ³¹
Total population	50.32	Population distribution ³²
Volume of effective component of nitrogenous fertilizer	6.60	Nitrogen fertilizer application ³³
Use of agricultural pesticide	1.13	/
Number of large animals	1.20	/
Number of sheep and goats	0.25	/
Total power of agricultural machinery	0.20	/
Number of large and medium-sized agricultural tractors	0.32	/
Number of small tractors	0.39	/
Gross output value of agriculture, forestry, animal husbandry and fishery	0.27	/
Gross output value of agriculture	0.48	/

Table 1. An overview of 12 indicators that could affect the intensity of provincial-level carbon emissions from crop farming.

Datasets	Evaluation indicators	Downscaling model
Training set	R^2	0.98
	RMSE	5.6×10^{-4}
Testing set	R^2	0.99
	RMSE	4.2×10^{-4}

Table 2. Evaluation indicators for the downscaling model on the intensity of provincial-level carbon emissions from crop farming.

Calculation of the intensity of county-level carbon emissions from crop farming. Based on the intensity of provincial-level carbon emissions from crop farming and satellite-derived data in China from 2000 to 2019, we constructed a downscaling model using the random forest method^{34,35} (Fig. 1). In order to evaluate the validation, we randomly divided 10% of the sample data into a testing set using for Leave-One-Out validation, and the remaining sample data was used as a training set^{36,37}. Among the training set, we chose 10% of the data randomly at one time for 10-fold cross-validation to avoid over-fitting^{38,39}. Two evaluation indicators were used to evaluate the performance of the simulation, including the determination coefficients (R^2) and the root mean square error (RMSE). As shown in Table 2, our constructed downscaling model provided the superior performance in both the training and testing sets, indicating that this model could reflect the intensity of carbon emissions from crop farming in China.

We then put the satellite-derived county-level data in China from 2000 to 2019 into the downscaling model, and used provincial-level carbon emissions of crop farming as a constraint to adjust the county-level carbon emissions for each province⁴⁰. At last, we generated a county-level dataset on the intensity of carbon emissions from crop farming in China during 2000–2019 (Fig. 1).

Data Records

The dataset available on the Figshare⁴¹ consists of two components. One is an Excel file describing the intensity of county-level carbon emissions from crop farming in China during 2000–2019. The other is the vector map, which can be used in the ArcGIS software to display all results by matching the Excel file. The temporal and spatial changes in the intensity of county-level carbon emissions from crop farming in China are shown in Fig. 2.

Technical Validation

Validity testing for time series of carbon emissions from crop farming in China. Based on the calculation method described above, we obtained provincial-level carbon emissions from crop farming during 1978–2019, and further calculated the corresponding intensity of carbon emissions from crop farming according to the land area of each province (Fig. 2). In order to compare our results with other studies, the Emissions Database for Global Atmospheric Research (EDGAR) dataset⁴² and previous accounting results^{18–20} were also collected and shown in Fig. 3.

At the national level, the results of our accounting showed an overall upward trend in carbon emissions from crop farming in China during 1978–2019, but there was a significant decline between 1997 and 2003 and between 2012 and 2020, respectively (Fig. 3a). Similar characteristics were also shown in the intensity of carbon emissions from crop farming in China (Fig. 3b), which were consistent with the EDGAR dataset⁴² and previous accounting results^{18–20}. However, direct numerical comparisons were difficult, considering that these studies selected different agricultural activities²⁰.

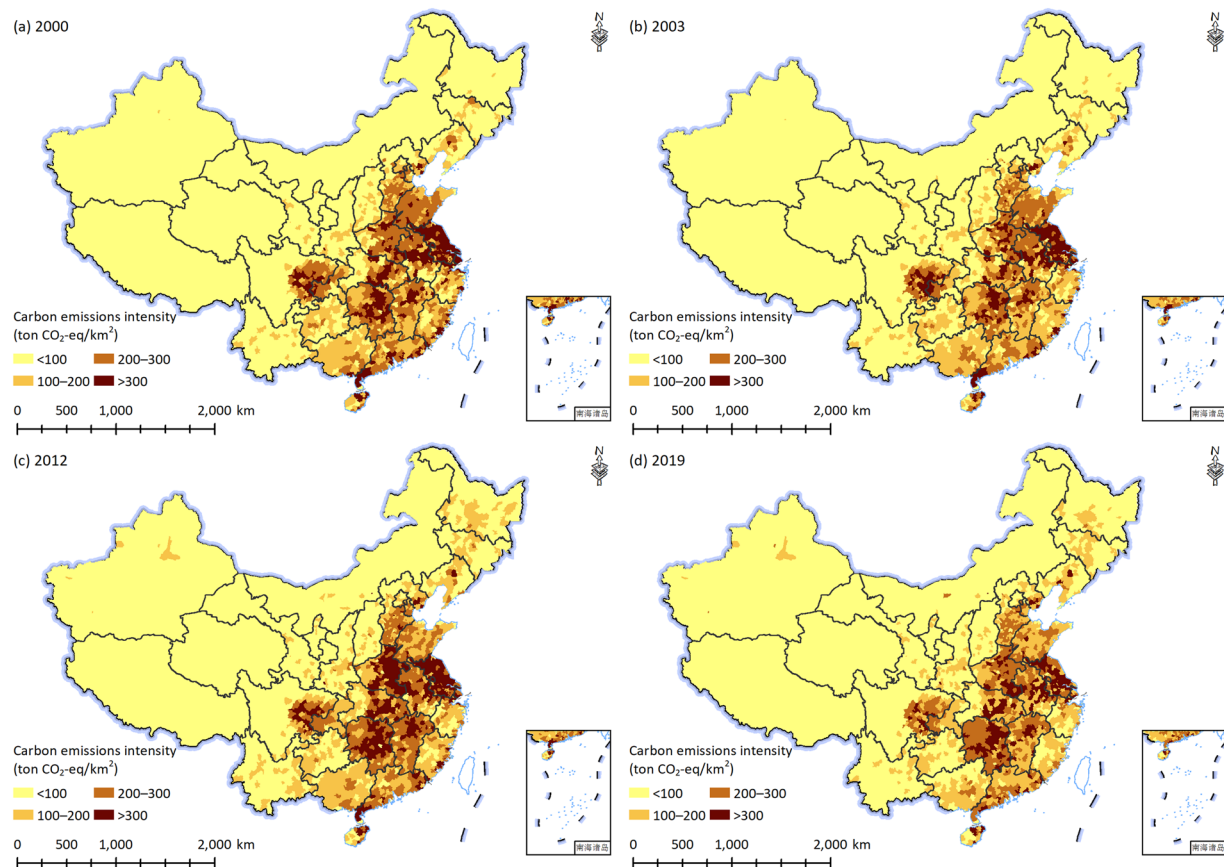


Fig. 2 Spatial patterns of county-level intensity of carbon emissions from crop farming in China during 2000–2019.

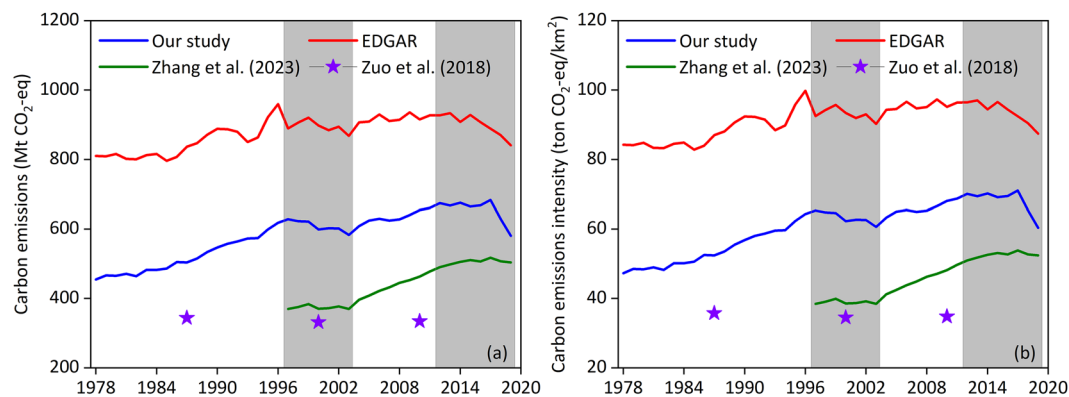


Fig. 3 Changes in carbon emissions from crop farming (a) and its intensity (b) in China.

Validity testing for the intensity of county-level carbon emissions from crop farming in China. At the spatial scale, we compared our findings with those of Zuo *et al.*¹⁸ and the EDGAR dataset for 2000 and 2010 in China (Fig. 4), which are raster data with a spatial resolution of 10 km. For comparison purposes, we extracted the results of Zuo *et al.*¹⁸ and the EDGAR dataset for each county in China using ArcGIS 10.2 software.

As shown in Fig. 4a and f, the results of our accounting showed an obvious spatial heterogeneity in the intensity of county-level carbon emissions from crop farming, with an overall distribution pattern of higher in the south and east and lower in the west and north. These spatial patterns were consistent with those of Zuo *et al.*¹⁸ (Fig. 4b and g) and the EDGAR dataset (Fig. 4c and h). Furthermore, there were significant positive relationships between our findings and the results of Zuo *et al.*¹⁸ for 2000 and 2010 (Fig. 4d and i). Moreover, significant positive relationships were also found between our findings and the results of the EDGAR dataset for 2000 and 2010 (Fig. 4e and j). The correlation coefficients between our findings and the results of Zuo *et al.*¹⁸ for 2000 and 2010 were 0.69 and 0.66, respectively, which were lower than those between our findings and the EDGAR dataset.

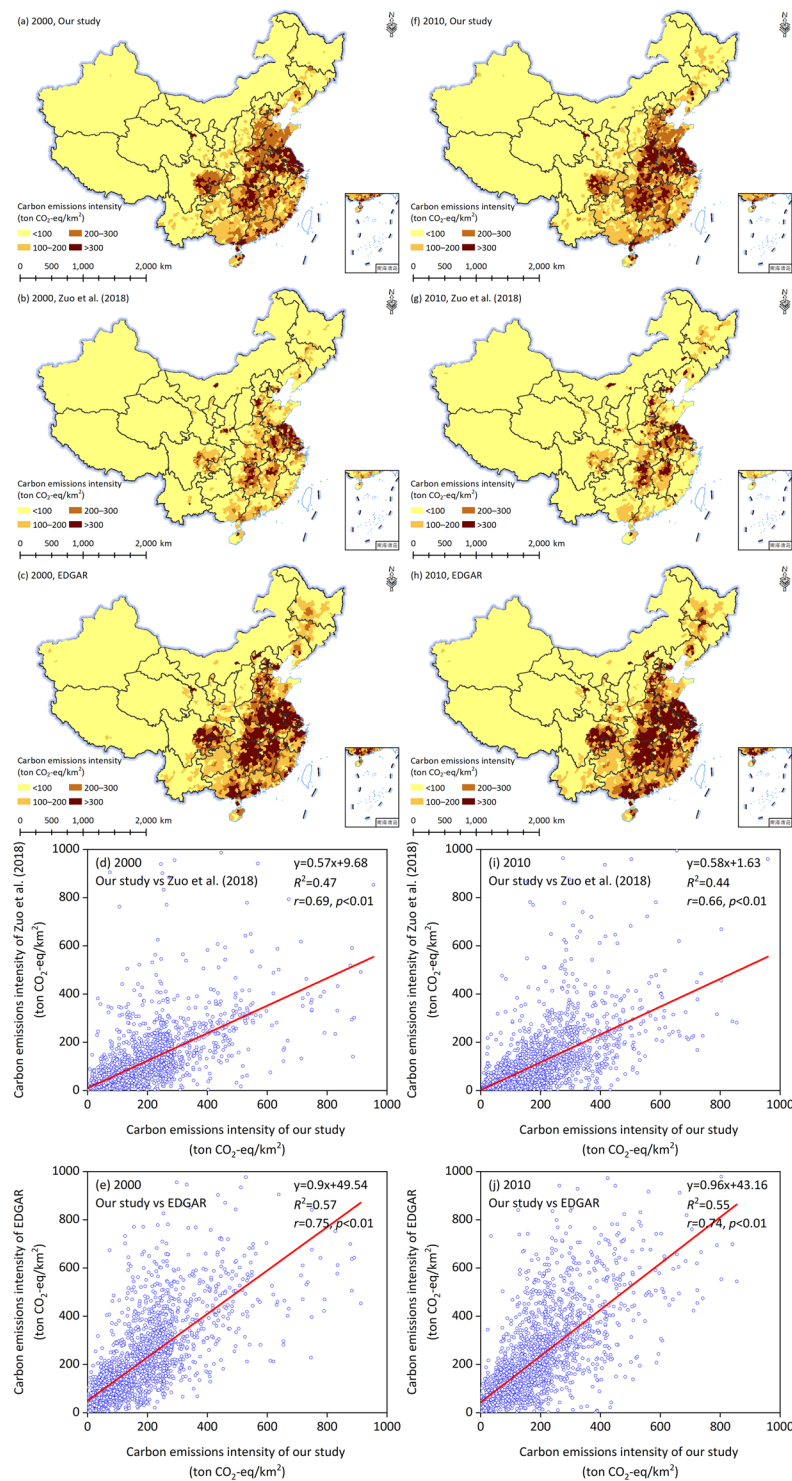


Fig. 4 Comparison of carbon emissions intensity estimated from Zuo *et al.*¹⁸ and the EDGAR dataset for 2000 (a–e) and 2010 (f–j).

Moreover, the corresponding R^2 values between our findings and the results of Zuo *et al.*¹⁸ for 2000 and 2010 were lower than those between our findings and the results of the EDGAR dataset. In addition to the spatial pattern, the results of our accounting were higher than those of Zuo *et al.*¹⁸ (Fig. 4b and g), but lower than those of the EDGAR dataset. An important reason was that different scholars considered carbon emissions from different agricultural activities and parameters in their work, which made direct comparisons difficult. For example, Zuo *et al.*¹⁸ considered carbon emissions from agricultural management (e.g., fertilizer application, rice cultivation, peatland draining) and cropland changes, whereas the EDGAR dataset mainly considered agricultural

carbon emissions. As previously described, we mainly focused on primary and secondary emissions from crop farming²⁰, and provided a long-term and spatially-precise profile of carbon emissions intensity from cropping farming in China.

Limitations and future work. The limitations of our dataset are based on three main aspects. First, our accounting system do not cover all agricultural activities involved in crop farming. For example, the contribution of CO₂ emissions resulting from land use transitions, such as the conversion of natural habitats for agricultural purposes or changes in agricultural land use, is not included in this dataset due to the high level of uncertainty in China's land use/land cover data^{24,43} (Figure S2). Additionally emissions from activities like forest product harvesting, peat drainage, and peat burning should be included as well^{17,18}, however, relevant long-term statistical data are difficult to obtain and are therefore not considered at the current stage. Second, there are many factors influencing the intensity of provincial-level carbon emissions from crop farming. However, it is difficult to find the corresponding satellite data or county-level statistical data for all these influencing factors when we constructed the downscaling model using the random forest method, which affects our calculating results. Third, this dataset is limited by different time spans between the provincial statistical data and satellite-derived data. Currently, this dataset is from 2000 to 2019 due to the availability of satellite-derived rice distribution data³¹. In the future, we will utilize multi-source data and data-model fusion method to further evaluate emissions from cropland change and other activities related to agriculture, and dynamically update our current dataset.

Usage Notes

In this study, we provide a county-level dataset on the intensity of carbon emissions from crop farming in China during 2000–2019, which can bridge existing gaps in the available datasets^{18–20}. On the one hand, this dataset can be used to analyse spatio-temporal changes in carbon emissions from crop farming and their driving factors in China from a county-scale perspective. On the other hand, this dataset also provides an important basis for decision makers and researchers to design agricultural carbon reduction strategies suitable for local contexts in China.

It is noted that the study area is mainland China (excluding Hong Kong, Macao and Taiwan). A total of 2369 data records are included in this dataset because we take the city-governed districts as a whole.

Code availability

Python and ArcGIS are the software used to generate all the results. The code used for the downscaling calculations is available on the Figshare⁴¹.

Received: 3 November 2023; Accepted: 23 April 2024;

Published online: 06 May 2024

References

1. Masson-Delmotte, V. *et al.* IPCC, 2021: Summary for Policymakers in Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change (Cambridge Univ. Press, 2021).
2. Zheng, X. *et al.* Drivers of change in China's energy-related CO₂ emissions. *Proc. Natl. Acad. Sci. USA* **117**, 29–36 (2020).
3. Duffy, K. A. *et al.* How close are we to the temperature tipping point of the terrestrial biosphere? *Sci. Adv.* **7**, eaay1052 (2021).
4. Ray, D. K., Gerber, J. S., MacDonald, G. K. & West, P. C. Climate variation explains a third of global crop yield variability. *Nat. Commun.* **6**, 5989 (2015).
5. Li, C. *et al.* Spatial heterogeneity in the response of winter wheat yield to meteorological dryness/wetness variations in Henan province, China. *Agronomy* **14**, 817 (2024).
6. Popp, A., Lotze-Campen, H. & Bodirsky, B. Food consumption, diet shifts and associated non-CO₂ greenhouse gases from agricultural production. *Glob. Environ. Change* **20**, 451–462 (2010).
7. Tian, H. *et al.* The terrestrial biosphere as a net source of greenhouse gases to the atmosphere. *Nature* **531**, 225–228 (2016).
8. Crippa, M. *et al.* Food systems are responsible for a third of global anthropogenic GHG emissions. *Nat. Food* **2**, 198–209 (2021).
9. Zhang, L., Pang, J., Chen, X. & Lu, Z. Carbon emissions, energy consumption and economic growth: Evidence from the agricultural sector of China's main grain-producing areas. *Sci. Total Environ.* **665**, 1017–1025 (2019).
10. Zeng, J. *et al.* Ecoefficiency of China's agricultural sector: What are the spatiotemporal characteristics and how are they determined? *J. Clean. Prod.* **325**, 129346 (2021).
11. Van Wessenbeeck, C. F. A., Keyzer, M. A., Van Veen, W. C. M. & Qiu, H. Can China's overuse of fertilizer be reduced without threatening food security and farm incomes? *Agr. Syst.* **190**, 103093 (2021).
12. Han, J. *et al.* A critical assessment of provincial-level variation in agricultural GHG emissions in China. *J. Environ. Manage.* **296**, 113190 (2021).
13. Hu, Y. *et al.* Food production in China requires intensified measures to be consistent with national and provincial environmental boundaries. *Nat. Food* **1**, 572–582 (2020).
14. Guo, H., Chen, F., Sun, Z., Liu, J. & Liang, D. Big Earth Data: A practice of sustainability science to achieve the Sustainable Development Goals. *Sci. Bull.* **66**, 1050–1053 (2021).
15. Zhao, X., Ma, X., Chen, B., Shang, Y. & Song, M. Challenges toward carbon neutrality in China: Strategies and countermeasures. *Resour. Conserv. Recycl.* **176**, 105959 (2022).
16. Tubiello, F. N. *et al.* The FAOSTAT database of greenhouse gas emissions from agriculture. *Environ. Res. Lett.* **8**, 015009 (2013).
17. Carlson, K. M. *et al.* Greenhouse gas emissions intensity of global croplands. *Nat. Clim. Chang.* **7**, 63–68 (2017).
18. Zuo, L. *et al.* Progress towards sustainable intensification in China challenged by land-use change. *Nat. Sustain.* **1**, 304–313 (2018).
19. Zhang, Q., Qu Y., Zhan, L., Su, D. & Wei, C. Dynamic evolution dataset of carbon emissions from grain production in China (1997–2020). Preprint at <https://doi.org/10.3974/geodb.2023.07.10.V1> (2023).
20. Liang, D. *et al.* China's greenhouse gas emissions for cropping systems from 1978–2016. *Sci. Data* **8**, 171 (2021).
21. Tubiello, F. N. *et al.* The contribution of agriculture, forestry and other land use activities to global warming, 1990–2012. *Glob. Change Biol.* **21**, 2655–2660 (2015).
22. Liu, G., Deng, X. & Zhang, F. The spatial and source heterogeneity of agricultural emissions highlight necessity of tailored regional mitigation strategies. *Sci. Total Environ.* **914**, 169917 (2024).
23. Chen, J. *et al.* County-level CO₂ emissions and sequestration in China during 1997–2017. *Sci. Data* **7**, 391 (2020).

24. Congalton, R. G., Gu, J., Yadav, K., Thenkabail, P. & Ozdogan, M. Global land cover mapping: A review and uncertainty analysis. *Remote Sens.* **6**, 12070–12093 (2014).
25. Yang, J. & Huang, X. The 30 m annual land cover dataset and its dynamics in China from 1990 to 2019. *Earth Syst. Sci. Data* **13**, 3907–3925 (2021).
26. Liu, J. *et al.* Spatiotemporal characteristics, patterns and causes of land use changes in China since the late 1980s. *Acta Geogr. Sin.* **69**, 3–14 (2014).
27. Azodi, C. B., Tang, J. & Shiu, S.-H. Opening the black box: interpretable machine learning for geneticists. *Trends Genet.* **36**, 442–455 (2020).
28. Guidotti, R. *et al.* A survey of methods for explaining black box models. *ACM Comput. Surv.* **51**, 93 (2019).
29. Luo, N. *et al.* China can be self-sufficient in maize production by 2030 with optimal crop management. *Nat. Commun.* **14**, 2637 (2023).
30. Li, X. & Xiao, J. Mapping photosynthesis solely from solar-induced chlorophyll fluorescence: A global, fine-resolution dataset of gross primary production derived from OCO-2. *Remote Sens.* **11**, 2563 (2019).
31. Luo, Y., Zhang, Z., Chen, Y., Li, Z. & Tao, F. ChinaCropPhen1km: a high-resolution crop phenological dataset for three staple crops in China during 2000–2015 based on leaf area index (LAI) products. *Earth Syst. Sci. Data* **12**, 197–214 (2020).
32. Dobson, J. E., Bright, E. A., Coleman, P. R., Durfee, R. C. & Worley, B. A. LandScan: A global population database for estimating populations at risk. *Photogramm. Eng. Remote Sens.* **66**, 849–857 (2000).
33. Tian, H. *et al.* History of anthropogenic Nitrogen inputs (HaNi) to the terrestrial biosphere: a 5 arcmin resolution annual dataset from 1860 to 2019. *Earth Syst. Sci. Data* **14**, 4551–4568 (2022).
34. Zhang, Y., Liu, X., Lei, L. & Liu, L. Estimating global anthropogenic CO₂ gridded emissions using a data-driven stacked random forest regression model. *Remote Sens.* **14**, 3899 (2022).
35. Jyolsna, P. J., Kambhammettu, B. & Gorugantula, S. Application of random forest and multi-linear regression methods in downscaling GRACE derived groundwater storage changes. *Hydrol. Sci. J.* **66**, 874–887 (2021).
36. Lipton, Z. C. The mythos of model interpretability. *Commun. ACM* **61**, 36–43 (2018).
37. Zhao, S. *et al.* Interpretable machine learning for predicting and evaluating hydrogen production via supercritical water gasification of biomass. *J. Clean. Prod.* **316**, 128244 (2021).
38. Fushiki, T. Estimation of prediction error by using K-fold cross-validation. *Stat. Comput.* **21**, 137–146 (2011).
39. Zhou, W. *et al.* Integrating climate and satellite remote sensing data for predicting county-level wheat yield in China using machine learning methods. *Int. J. Appl. Earth Obs. Geoinf.* **111**, 102861 (2022).
40. Halpern, B. S. *et al.* The environmental footprint of global food production. *Nat. Sustain.* **5**, 1027–1039 (2022).
41. Li, C., Jia, J., Wu, F., Zuo, L., & Cui, X. County-level intensity of carbon emissions from crop farming in China during 2000–2019, *figshare*, <https://doi.org/10.6084/m9.figshare.24421261.v1> (2023).
42. Janssens-Maenhout, G. *et al.* EDGAR v4.3.2 Global Atlas of the three major greenhouse gas emissions for the period 1970–2012. *Earth Syst. Sci. Data* **11**, 959–1002 (2019).
43. McGlynn, E., Li, S., Berger, M. F., Amend, M. & Harper, L. K. Addressing uncertainty and bias in land use, land use change, and forestry greenhouse gas inventories. *Clim. Change* **170**, 5 (2022).

Acknowledgements

This work was financially supported by the National Key Research and Development Program of China (2023YFF1303702), the International Research Centre of Big Data for Sustainable Development Goals (CBAS2023SDG001), the National Natural Science Foundation of China (41801013), the Strategic Priority Research Program of the Chinese Academy of Sciences (XDA19090119), the Young Scientific and Technological Talents Support Project of Jiangsu Association for Science and Technology (TJ-2023-032), the Humanities and Social Sciences Foundation of Yangzhou University (xjj2023-05) and the Qinglan Project of Yangzhou University.

Author contributions

C.L. collected the data and wrote the manuscript. J.J. designed the algorithm and performed the downscaling calculations. F.W., L.Z. and X.C. revised the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41597-024-03296-y>.

Correspondence and requests for materials should be addressed to X.C.

Reprints and permissions information is available at www.nature.com/reprints.

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



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

© The Author(s) 2024