Responses to the comments of Referee #1

Article ID: essd-2024-586
Title: NortheastChinaSoybeanYield20m: an annual soybean yield dataset at 20 m in Northeast China from 2019 to 2023
Authors: Jingyuan Xu, Xin Du, Taifeng Dong, Qiangzi Li, Yuan Zhang, Hongyan Wang, Jing Xiao, Jiashu Zhang, Yunqi Shen, Yong Dong

Dear Reviewer,

Thank you very much for your thorough review and valuable comments on our manuscript. Your insightful feedback has significantly contributed to improving the quality and clarity of our work. In response to your suggestions, we have rigorously revised the manuscript to address each comment and suggestion. The key modifications include:

(1) Data Description Revisions: the data sections (e.g., field measurements, remote sensing data, and statistical sources) have been thoroughly revised to eliminate ambiguity and ensure reproducibility. Clarifications have been added to prevent misinterpretations on the datasets, including explicit dentitions of sampling protocols and spatial-temporal resolution.
(2) Methodological Enhancements: Additional details on the hybrid approach, such as descriptions of scenario parameterization, temporal feature extraction and computational have been added to strengthen methodological transparency and rigor.
(3) Readability Improvements: Key paragraphs and sentences have been restructured for logical flow. Figures or tables have been refined to align with revised text, ensuring visual clarity and consistency with results.

The detailed point-to-point responses are as follows. Texts in black are the reviewer's comments; those in blue are our responses to the reviewer's comments; and those in *red and italics* are the revised texts appeared in the revised manuscript.

We will attach a clean version (Manuscript_Clean_Version.docx) as well as a tracking enabled version (Manuscript_Marked_Version.docx) with editing marks for your reference.

I am very familiar with the WOFOST model and the dataset used by the author. It is not a good simulation project, not only because the simulation accuracy did not meet industry standards, but also because the author withheld many critical details and settings of the WOFOST in the manuscript, which makes it difficult for me to assess the rationality and scientific validity of the simulation. *Earth System Science Data*, as the name suggests, focuses on the application of datasets, but the author's professionalism in describing and processing the dataset is not good. Moreover, the description of CRU is severely inadequate. After reading the entire manuscript, I still do not understand the role of the CRU used by the author in this study.

Reply: Thank you for your thorough review and valuable feedback.

We sincerely appreciate your expertise and the time invested in evaluating our work. Below, we address your concerns point by point:

(1) "The simulation accuracy did not meet industry standards, and critical details/settings of WOFOST were withheld, making it difficult to assess scientific validity."

We acknowledge your concern and regret any lack of clarity about the WOFOST model and its reparameterization in our original manuscript.

In the revised manuscript, key edits to improve transparency include:

(a) Additional Details Added: detailed description on the soil data (Section 2.2.3) and statistical data (Section 2.2.6) was added. We also have provided more detailed information to explicitly outline the WOFOST parameterization, including soil properties (Line 217 – 225), crop variety parameters (e.g., TSUM1, TSUM2, Table A1), and agro-management scenarios (e.g., planting date, Line 238 – 245).

(b) Model Calibration and Validation: we have revised the section 5.2 to include more detailed comparisons with existing studies at various spatial scales. For instance, our accuracy at the municipal-scale was RMSE of 272.36 kg ha^{-1} outperformed the 420 kg ha^{-1} reported by Von Bloh et al., (2023). Additionally, we have expanded the discussion on the comparison of our dataset with soybean yield datasets from other countries, demonstrating the higher precision of our estimates. The comparisons include a detailed analysis of RMSE, supporting the claim that our method provides reliable and accurate yield estimates (Line 469 – 478).

(c) Supplementary Material: A table of WOFOST input parameters is provide in Supplementary Table S1 to enhance reproducibility.

(2) "The description of CRU is severely inadequate, and the dataset's role in the study is unclear."

We apologize for the initial oversight.

In the revised manuscript, we have enhanced the description on the WOFOST- GRU integration. We have detailed and clarified the coupling mechanism between the GRU and WOFOST models including generating input features from time-series LAI data, and explicitly outlining the training workflow (Section 3.2, Line 264 – 279). Figure 2 has been revised to explicitly illustrate the flow of data between WOFOST simulations, remote sensing inputs, and GRU-based yield estimation.

(3) ***"The dataset description lacks professionalism, and the manuscript does not align with Earth System Science Data's focus."***

We have rigorously revised the manuscript to emphasize: (a) Dataset Documentation, the "data availability" has been revised to detail file formats, data spatial-temporal resolution. Metadata with these information has been added to the dataset hosted on Zenodo. (b) Application Focus, we have revised the conclusions to synthetically illustrate the innovation of the method for dataset construction and demonstrate the dataset's utility for agricultural management practise (e.g., optimizing sowing date for maxing yield), aligning with the journal emphasis on actionable Earth system data.

We are grateful for your critique, which has significantly strengthened our work. We hope these revisions address your concerns raised and significantly enhance the quality and clarity of the manuscript.

Below, we provide a detailed point-by-point response to address the specific concerns you raised:

1. The study spanned from 2019 to 2023, but the sampling data was only from 2022 and 2023 (Fig.1). The author should explain this issue in the text.

Reply: Thank you very much for the suggestion.

In the revised manuscript, we have clarified that although the study spanned from 2019 to 2023, field observations were not conducted from 2019 to 2021 due to resource limitations. As a result, the sampling data for this study were collected only in 2022 and 2023 (Line 106 - 108).

Due to resources and personnel constraints, in-situ yield measurements were unavailable for the initial years (2019 - 2021). Field-scale yield data was collected through field surveys in September 2022 and 2023, with 21 and 18 sample plots surveyed annually, respectively.

2. The soil data should be described in more detail, for example, which soil parameters were used in this study.

Reply: Thanks for your suggestion.

As suggested, we have provided a more detailed description of the soil data used in this study. For our study, we did not use the soil attribute data (such as chemical characteristics) in this analysis, focusing only on the spatial distribution of soil types to characterize the soil environment in the study region. This clarification has been added to the revised manuscript (Section 2.2.3).

Soil data was obtained from the 1:1000,000 Chinese Soil Database, accessible via the Geographic Data Sharing Infrastructure and the Global Resources Data Cloud platform (www.gis5g.com). The dataset adopted the traditional "Chinese soil classification system", with subcategories serving as the foundational mapping unit. It categorized soil hierarchically into 12 soil orders, 61 soil types, and 227 subcategories, providing comprehensive coverage of China's soil diversity. The dataset consisted of two components: spatial data for vector layers delineating soil type distributions at the national scale, and attribute data that include key chemical properties (e.g., pH, organic matter) and physical characteristics (e.g., texture, bulk density). In this study, the 1:1000,000 soil spatial data was obtained to map soil types for the study area, enabling integration with agro-ecological variables (e.g., crop suitability, irrigation requirements) in the hybrid yield estimation framework.

3. The author used statistical data from 1980 to 2022, but the study's time scale is from 2019 to 2023. This is confusing for the readers. Please provide an explanation.

Reply: Thanks for your suggestion.

We have clarified the use of statistical data in the revision. In this study, the statistical data served two main purposes. Firstly, to ensure that the multi-scenarios soybean growth dataset, which was constructed in this study, accurately represented a wide range of soybean production conditions, the statistical data from 1980 to 2022 were all used to validate the reasonableness of the model simulations. Secondly, statistical data from 2019 to 2022 were specifically applied to evaluate accuracy of soybean yield estimation at the regional scale.

This distinction has been clearly explained in the revised manuscript to avoid any confusion for the readers (Section 2.2.6).

The crop yield records (1980-2022) were obtained from the Statistical Yearbook published by provincial authorities: Heilongjiang (http://tjj.hlj.gov.cn), Jilin (http://tjj.jl.gov.cn), Liaoning (https://tjj.ln.gov.cn) , and Inner Mongolia Autonomous Region (IMAR , https://tj.nmg.gov.cn). Due to the incomplete publication of the 2022 Statistical Yearbook, municipal-level yield for that year were limited to a subset of cities. The statistical data were used for model simulation validation and regional-

scale yield accuracy evaluation. The multi-scenario soybean growth dataset, simulated using multi-years meteorological data, various soil types, multiple soybean varieties and different agro-managements, required validation against real production conditions. The historical statistics (1980 – 2022), supplemented by published yield datasets and field samples, were used to evaluate the plausibility of the simulated dataset validated the reasonableness of the model simulations. For validating yield estimates at regional scale, data from 2019 to 2022 employed to quantify estimation errors at regional scales.

4. The technology roadmap that needs improvement. 1) The author mentioned agro-management data in Figure 2, but it is not mentioned in Section 2.2 Data collections. 2) The sampling data mentioned in the Data collections section is not reflected in the figures, as well as meteorological data from National Meteorological Information Center. 3) The method of combining remote sensing data and model output through GRU is described too simplistically. 4) The author allocates a large proportion of the figures to how WOFOST conducts simulations, but this is not the focus of this study. The focus of this study should be on how to use models and remote sensing coupled for yield estimation, just as the author introduces in the research objective: "Designing a hybrid model coupling crop growth model and deep learning model for soybean yield estimation." The technology roadmap should more detailed display the research focus.

Reply: Thank you for pointing this out.

We have made substantial revisions to improve the technology roadmap, addressing the points raised:

(1) The study primarily simulated different soybean agro-management scenarios by setting different planting dates as literatures have highlight the importance of planting date on soybean yield. In addition, for large-scale estimation modeling, it's more practical to focus on broader factors like planting dates rather than detailed management measures. The agro-management data was collected alongside in-situ measurements. We have now included an explanation of the agro-management data collection process in Section 2.2.1 (Line 112 - 114) to ensure consistency between the figure and the text.

Due to resources and personnel constraints, in-situ yield measurements were unavailable for the initial years (2019 - 2021). Field-scale yield data was collected through field surveys in September 2022 and 2023, with 21 and 18 sample plots surveyed annually, respectively (Fig. 1). Each plot covered an area of approximately 100 m × 100 m, within which nine randomly distributed 1 m × 1 m quadrats were selected for destructive soybean yield sampling. The central geo-location of each quadrat was recorded using a GPS device with accuracy of 1 m. Harvested beans for each quadrat were then oven-dried 72 hours to determine moisture-free yield, which was processed at the Hailun Agricultural Ecology Experimental Station, Chinese

Academy of Sciences. Finally, the mean yield of the nine quadrats was calculated to represent the plot-level yield. ***Additionally, soybean planting dates across regions were collected through field surveys to support agro-management parameterization in the model.***

(2) The sampling data were used to validate soybean yield estimation accuracy at the field scale, while the meteorological data from the National Meteorological Information Center were used as input for the WOFOST model to provide essential weather parameters. Both of these datasets are represented in the revised figure (field measured samples and meteorological station data, respectively), and their roles have been explicitly described in the text.

(3) We have revised the flowchart to better represent the hybrid model's structure. The updated version now clearly distinguishes two main components: the first part describes the construction of the hybrid model, combining WOFOST and GRU for yield estimation, while the second part focuses on how remote sensing data are used in conjunction with this hybrid model to estimate soybean yield at the regional scale. We have provided more detail in the flowchart regarding how remote sensing data are integrated as inputs into the GRU model and how they contribute to spatial yield predictions.

(4) We agree that the focus of the study should be on the coupling of remote sensing data with crop growth models for yield estimation, as stated in the research objectives. Therefore, the updated technology roadmap places greater emphasis on this aspect, reducing the proportion dedicated to the WOFOST model and highlighting how the hybrid model (WOFOST + GRU) is used for yield estimation. This revision better aligns the roadmap with the research focus and research objectives, as well as the methodology described in the manuscript.

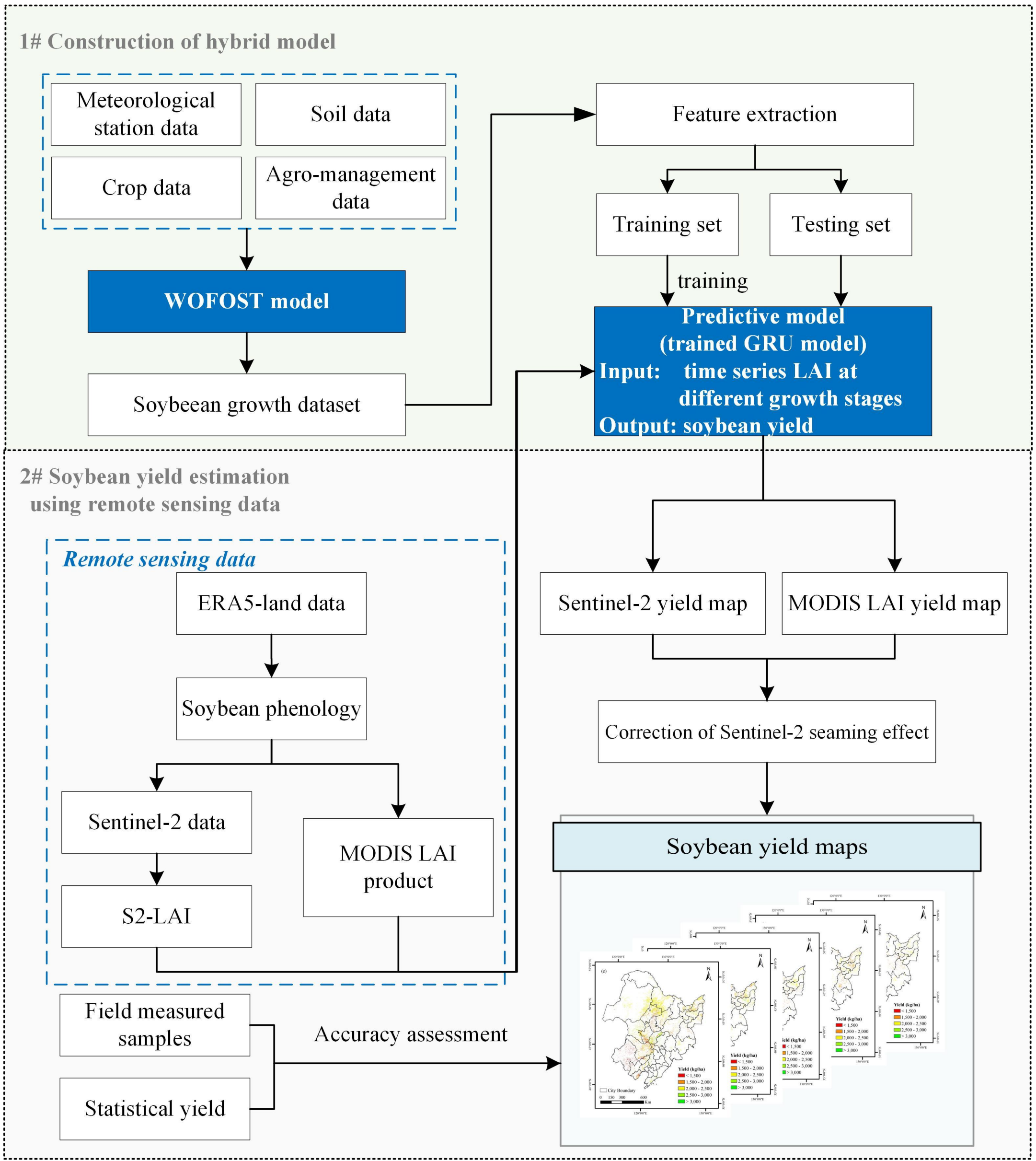

Figure 2: The flowchart of the overall yield estimation methodology in this study.

5. Which sub-model of PCSE did the author use? LINTUL3 or Wofost72_PP?

Reply: Thanks for your suggestion.

Our soybean simulations were conducted using the Wofost72_WLP_CWB (water-limited mode) within the PCSE (v5.5.) for soil-water dynamics and crop responses to water stress. This is because rainfed agricultural practices dominant in Northeast China, where water availability rather than nutrient constraints is the major limiting factor on crop growth.
Relevant edits have been revised in the revision (Line 197 – 199).

In this study, the World Food Studies Simulation Model (WOFOST) model (Diepen et al., 1989) was employed to generate knowledge-based soybean growth dataset.

Detailed information about the model could be found in Huang et al., (2015). ***Since rainfed agricultural practices dominant in Northeast China, the water-limited mode of the WOFOST (Wofost72_WLP_CWB) was employed for soybean simulation from the Python Crop Simulation Environment (PCSE) framework version 5.5.***

6. As far as I know, VAP is not included in the ERA5 dataset. How did the author obtain the VAP data?

Reply: Thank you for pointing out this.

The study only used the daily aggregated air temperature data at 2 meters from the ERA5 dataset to calculate the spatial distribution of soybean phenology, which was used to guide the acquisition time for remote sensing data. For the meteorological parameters required by the WOFOST model (including VAP), all data were sourced from meteorological station data rather than ERA5.

This clarification has been added to the manuscript in Line 132 – 137 and Line 211 – 212.

The climate reanalysis data was obtained from the ERA5-land Daily Aggregated ECMWF product. The data was only used to calculate soybean phenology for preparation of yield estimations. *It was a global climate reanalysis product that provided continuous climate data at a resolution of 0.1° × 0.1° (e.g., air temperature and atmospheric pressure) starting from 1950. The daily aggregated air temperature data at 2 m above the surface of land measured in kelvin (K)* ***during the soybean growth periods from 2019 to 2023 was collected in this study*** *from the Google Earth Engine (http://earthengine.google.com). The product was then resampled to 20 m using bilinear interpolation model.*

The meteorological parameters required in WOFOST is shown in Table 1. ***The data was only provided by meteorological stations.*** *To account for various meteorological conditions, the meteorological data collected from the selected 51 meteorological stations over a period of 42 years (1980-2021) was all utilized to provide values of these parameters. The meteorological inputs were then processed into the format recognized by the model.*

7. Line 209-215: The description of the calculation process for soil parameters is too simplistic; a detailed calculation process should be provided. For example, which parameters from the Chinese soil database were used in the study, and what theories/formulas were utilized to calculate the SMW, SMFCF, SM0, and K0 required by the WOFOST model? Is Table 2 a lookup table? Where did it come from?

Reply: Thanks for your suggestion.

In the revised version of the manuscript, we have provided more detailed information on the soil parameter settings (Line 217 – 225). Based on the soil spatial data, we found that the main soil types in the study area can be categorized as sandy loam, light loam, medium loam, and heavy loam. The parameter settings for these different soil types were primarily obtained from existing literature. Table 2 is not a lookup table, but a compilation of the parameters based on previous studies.

We have now indicated the sources of the soil parameters presented in Table 2 in the manuscript.

The soil parameters in the WOFOST mainly include soil moisture content at wilting point (SMW), field capacity (SMFCF) and saturation (SM0) as well as hydraulic conductivity of saturated soil (K0). According to the 1:1000,000 Chinese soil database, the soil texture in study area is predominantly loam soil, which can be further divided into sandy loam, light loam, medium loam and heavy loam. Sun et al., (2022) found that the properties of typical heavy loam suitable for soybean cultivation were similar to those provided in the model source file WOFOST Control Centre/SOILD/EC2.NEW and further updated the soil parameters of heavy loam based on this file by incorporating soil data observed at meteorological stations. The parameters of sandy loam, light loam and heavy loam were collected from Du et al., (2025). The value settings of soil parameters for different soil types in the study area was presented in Table 2 (Du et al., 2025; Sun et al., 2022). They were all used as model inputs to consider the impact of different soil types on soybean production.

8. Line 215: The description of Table 3 is redundant. It suffices to directly list the values and sources of the WOFOST crop parameters. Table 4 should list all crop parameters in WOFOST, not just the main crop parameters.

Reply: Thanks for your suggestion. We have removed Table 3 and its accompanying description. Instead, we have provided a complete list of the WOFOST crop parameters and their corresponding values as suggested. Due to the length of the table, it has been placed in the appendix (Table A1). This change ensures a more comprehensive and concise presentation of the crop parameters used in the study.

In this study, the soybeans were classified into five types including early maturity, medium-early maturity, intermediate maturity, medium-late maturity and late maturity according to Qu et al., (2023). In the WOFOST model, soybean growth stages are mainly determined by temperature-related parameters including the minimum and maximum threshold temperature for emergence (TBASEM, TEFFMX, respectively), accumulated temperature (Te) from sowing to emergence (TSUMEM), from emergence to anthesis (TSUM1) and from anthesis to maturity (TSUM2). The accumulated temperature for different growth stage is sensitive to crop varieties. They were set according to the historical meteorological data and observation data of *soybean and had been validated using actual measurements of soybean development periods in Qu et al., (2023). Other crop parameters were set used the default values of the WOFOST model or the optimal values from the study of Sun et al., (2022). Values of all crop parameters could be found in Table A1.*

Table A1 Values of crop parameters in WOFOST.

Parameter	Description	Units	Value	Source
Crop initial parameters				
TDWI	Initial total crop dry weight	kg ha^{-1}	120	Default value in WOFOST
RGRLAI	Maximum relative increase in LAI	ha ha^{-1} d^{-1}	0.01	Default value in WOFOST
Parameters for emergence				
TBASEM	Minimum threshold temperature for emergence	°C	8.0	Qu et al., (2023)
TEFFMX	Maximum threshold temperature for emergence	°C	22.0	Qu et al., (2023)
TSUMEM	Accumulated temperature from sowing to emergence	°C	70.0	Qu et al., (2023)
Phenological parameters				
DLO	Optimal daylength for development	h	-99	Default value in WOFOST
DLC	Critical daylength	h	-99	Default value in WOFOST
TSUM1	Cumulative temperature from emergence to anthesis	°C	450 (early maturity) 480 (medium-early maturity) 520 (intermediate maturity) 540 (medium-late maturity) 580 (late maturity)	Qu et al., (2023)
TSUM2	Cumulative temperature from anthesis to maturity	°C	660 (early maturity) 770 (medium-early maturity) 870 (intermediate maturity) 960 (medium-late maturity) 1000 (late maturity)	Qu et al., (2023)
Green area parameters				
TBASE	Lower threshold temperature for aging of leaves	°C	7.0	Default value in WOFOST
SPAN	Life span of leaves growing at 35 °C	d	23	Default value in WOFOST
SLATB00	Specific leaf area at DVS = 0.00	ha kg^{-1}	0.00140	Default value in WOFOST
SLATB045	Specific leaf area at DVS = 0.45	ha kg^{-1}	0.00250	Default value in WOFOST
SLATB090	Specific leaf area at DVS = 0.90	ha kg^{-1}	0.00250	Default value in WOFOST
SLATB200	Specific leaf area at DVS = 2.00	ha kg^{-1}	0.00070	Default value in WOFOST
Assimilation parameters				

KDIFTB00	Extinction coefficient for diffuse visible light (DVS = 0)	-	0.80	Default value in WOFOST
KDIFTB200	Extinction coefficient for diffuse visible light (DVS = 2)	-	0.80	Default value in WOFOST
EFFTB0	Light use efficiency of a single leaf (T = 0 °C)	kg ha^{-1} h^{-1} J^{-1} m^2 s^{-1}	0.40	Default value in WOFOST
EFTB40	Light use efficiency of a single leaf (T = 40 °C)	kg ha^{-1} h^{-1} J^{-1} m^2 s^{-1}	0.40	Default value in WOFOST
AMAXTB00	Maximum leaf CO_2 assimilation rate (DVS = 0)	kg ha^{-1} h^{-1}	29.00	Default value in WOFOST
AMAXTB170	Maximum leaf CO_2 assimilation rate (DVS = 1.7)	kg ha^{-1} h^{-1}	25.31	Sun et al., (2022)
AMAXTB200	Maximum leaf CO_2 assimilation rate (DVS = 2)	kg ha^{-1} h^{-1}	0.00	Default value in WOFOST
TMPFTB00	Reduction factor of AMAX (T = 0 °C)	-	0.00	Default value in WOFOST
TMPFTB10	Reduction factor of AMAX (T = 10 °C)	-	0.30	Default value in WOFOST
TMPFTB20	Reduction factor of AMAX (T = 20 °C)	-	0.60	Default value in WOFOST
TMPFTB25	Reduction factor of AMAX (T = 25 °C)	-	0.80	Default value in WOFOST
TMPFTB30	Reduction factor of AMAX (T = 30 °C)	-	1.00	Default value in WOFOST
TMPFTB35	Reduction factor of AMAX (T = 35 °C)	-	1.00	Default value in WOFOST
Conversion of assimilates into biomass				
CVL	Conversion efficiency of assimilates into leaf tissue	kg kg^{-1}	0.72	Default value in WOFOST
CVO	Conversion efficiency of assimilates into storage organs	kg kg^{-1}	0.48	Default value in WOFOST
CVR	Conversion efficiency of assimilates into root tissue	kg kg^{-1}	0.72	Default value in WOFOST
CVS	Conversion efficiency of assimilates into stem tissue	kg kg^{-1}	0.69	Default value in WOFOST
Maintenance respiration parameters				
Q10	Relative change in respiration rate per 10 °C temperature increase	-	2.0	Default value in WOFOST
RML	Ralative maintenance respiration rate of leaves	kg CH_2O kg^{-1} d^{-1}	0.03	Default value in WOFOST
RMO	Ralative maintenance respiration rate of storage organs	kg CH_2O kg^{-1} d^{-1}	0.017	Default value in WOFOST
RMR	Ralative maintenance respiration rate of toots	kg CH_2O kg^{-1} d^{-1}	0.01	Default value in WOFOST
RMS	Ralative maintenance respiration rate of stems	kg CH_2O kg^{-1} d^{-1}	0.015	Default value in WOFOST
Partitioning parameters				
FRTB00	Fraction of total dry matter to roots at DVS = 0	kg kg^{-1}	0.62	Sun et al., (2022)
FRTB075	Fraction of total dry matter to roots at DVS = 0.75	kg kg^{-1}	0.35	Default value in WOFOST
FRTB100	Fraction of total dry matter	kg kg^{-1}	0.15	Default value in

	to roots at DVS = 1			WOFOST
FRTB150	Fraction of total dry matter to roots at DVS = 1.5	kg kg^{-1}	0.00	Default value in WOFOST
FRTB200	Fraction of total dry matter to roots at DVS = 2.0	kg kg^{-1}	0.00	Default value in WOFOST
FLTB00	Fraction of total dry matter to leaves at DVS = 0	kg kg^{-1}	0.70	Default value in WOFOST
FLTB100	Fraction of total dry matter to leaves at DVS = 1.0	kg kg^{-1}	0.70	Default value in WOFOST
FLTB115	Fraction of total dry matter to leaves at DVS = 1.15	kg kg^{-1}	0.60	Default value in WOFOST
FLTB130	Fraction of total dry matter to leaves at DVS = 1.3	kg kg^{-1}	0.43	Default value in WOFOST
FLTB150	Fraction of total dry matter to leaves at DVS = 1.5	kg kg^{-1}	0.15	Default value in WOFOST
FLTB200	Fraction of total dry matter to leaves at DVS = 2.0	kg kg^{-1}	0.00	Default value in WOFOST
FSTB00	Fraction of total dry matter to stems at DVS = 0	kg kg^{-1}	0.30	Default value in WOFOST
FSTB100	Fraction of total dry matter to stems at DVS = 1.0	kg kg^{-1}	0.30	Default value in WOFOST
FSTB115	Fraction of total dry matter to stems at DVS = 1.15	kg kg^{-1}	0.25	Default value in WOFOST
FSTB130	Fraction of total dry matter to stems at DVS = 1.3	kg kg^{-1}	0.10	Default value in WOFOST
FSTB150	Fraction of total dry matter to stems at DVS = 1.5	kg kg^{-1}	0.10	Default value in WOFOST
FSTB200	Fraction of total dry matter to stems at DVS = 2.0	kg kg^{-1}	0.00	Default value in WOFOST
FOTB00	Fraction of total dry matter to storage organs at DVS = 0	kg kg^{-1}	0.00	Default value in WOFOST
FOTB100	Fraction of total dry matter to storage organs at DVS = 1.0	kg kg^{-1}	0.00	Default value in WOFOST
FOTB115	Fraction of total dry matter to storage organs at DVS = 1.15	kg kg^{-1}	0.15	Default value in WOFOST
FOTB130	Fraction of total dry matter to storage organs at DVS = 1.3	kg kg^{-1}	0.47	Default value in WOFOST
FOTB150	Fraction of total dry matter to storage organs at DVS = 1.5	kg kg^{-1}	0.75	Default value in WOFOST
FOTB200	Fraction of total dry matter to storage organs at DVS = 2.0	kg kg^{-1}	1.00	Default value in WOFOST
Death rate parameters				
PERDL	Maximum relative death rate of leaves due to water stress	kg kg^{-1} d^{-1}	0.03	Default value in WOFOST
RDRRTB00	Relative death rate of roots at DVS = 0	kg kg^{-1} d^{-1}	0.00	Default value in WOFOST
RDRRTB150	Relative death rate of roots at DVS = 1.5	kg kg^{-1} d^{-1}	0.00	Default value in WOFOST
RDRRTB151	Relative death rate of roots at DVS = 1.51	kg kg^{-1} d^{-1}	0.02	Default value in WOFOST
RDRRTB200	Relative death rate of roots at DVS = 2.0	kg kg^{-1} d^{-1}	0.02	Default value in WOFOST
RDRSTB00	Relative death rate of stems	kg kg^{-1} d^{-1}	0.00	Default value in

	at DVS = 0			WOFOST
RDRSTB150	Relative death rate of stems at DVS = 1.5	kg kg^{-1} d^{-1}	0.00	Default value in WOFOST
RDRSTB151	Relative death rate of stems at DVS = 1.51	kg kg^{-1} d^{-1}	0.02	Default value in WOFOST
RDRSTB200	Relative death rate of stems at DVS = 2.0	kg kg^{-1} d^{-1}	0.02	Default value in WOFOST
Water use parameters				
CFET	Correction factor transpiration rate	-	1.0	Default value in WOFOST
DEPNR	Crop group number for soil water depletion	-	5.0	Default value in WOFOST
IAIRDU	Air ducts in roots present (=1) or not (=0)	-	0	Default value in WOFOST
IOX	Oxygen stress effect enabled (=1) or not (=0)	-	0	Default value in WOFOST
Rooting parameters				
RDI	Initial rooting depth	cm	10	Default value in WOFOST
RRI	Maximum daily increase in rooting depth	cm d^{-1}	1.2	Default value in WOFOST
RDMCR	Maximum rooting depth	cm	120	Default value in WOFOST

9. Line 235: What's the setting of the fertilizer application rate and timing in the WOFOST?

Reply: Thanks for your suggestion.

We would like to clarify that in this study, no fertilizer applications were considered. For agricultural production management, literatures have highlight that the planting date has a significant impact on soybean yield. The study simulated different agricultural management practices by varying the planting dates. As soybeans in study area are seldom subjected to nutrient stress according to the field surveys, the WOFOST model was not configured with specific fertilizer application rates or timings for this study. As when it comes to large-scale estimation modeling, focusing too much on detailed management measures might not be practical. A more generalized approach, taking into account broader factors like planting dates, can provide a better balance between accuracy and feasibility for large-area predictions.

This has been explained in the revised manuscript (Line 238 – 245).

Research demonstrates that planting date significantly influences soybean yield components, including pods per plant, seed count, and plant biomass, while also affecting oil and protein content (Urda et al., 2024). When conducting large-scale estimation modelling, prioritizing overly detailed management practices may lack practicality. Furthermore, field surveys indicate that soybeans in study area rarely experience nutrient stress. Given these findings, the simulations did not account for fertilizer applications. To balance realism and scalability, agro-management variability was represented solely through planting date adjustments. Since soybeans *in the study area are typically sown between late April and late May, so simulations incorporated four planting scenarios: 20 April, 30 April, 10 May, and 20 May. Overall, the simulations capture phenological variability while avoiding unnecessary complexity in regional scale.*

10. Line 244: After reading Section 3.2 Development of the Grated Recurrent Unit model (GRU), I am still unclear about the role of GRU in this study. The author's explanation of the principles of GRU is unclear. It does not directly describe how GRU combines the output of the WOFOST model with remote sensing data, as shown in the technical roadmap. Figure 3 lacks self-explanatory power, leaving it unclear what exactly the inputs and outputs of the GRU are.

Reply: Thanks for your suggestion. In the revised version, we further clarified the role of the GRU model in this study. The GRU was used for large-scale soybean yield estimation in this study. Since the internal structure of the GRU model was not adjusted in this study, we reduced the description of the GRU model's principles and instead focused more on its application in yield estimation. We removed the description of the GRU cell structure in Figure 3 and replaced it with a more detailed explanation of how the GRU model was trained using the soybean growth dataset simulated by the WOFOST model, enabling it to quickly estimate soybean yield at the regional scale. Additionally, we specified the inputs and outputs of the GRU model. In this study, the GRU model uses the average LAI values at different soybean growth stages as input features and soybean yield as the output. The trained GRU model was then integrated with remote sensing data for large-scale soybean yield mapping using the features derived from the remote sensing data as inputs (Section 3.2).

After construction of soybean growth dataset, the GRU model was used for large-scale soybean yield estimation. The GRU is a type of RNN model similar to LSTM. *It controls the flow of information through two gates, update and reset gates (Cho et al., 2014). The update gate aims to control how much of the past information that are retrained and will be used in the future calculation. The reset gate aims to evaluate whether the remained previous information can be ignored in the new candidate hidden state. The use of two gates maintains the balance between retaining the hold hidden state and incorporating new information (Peng and Yili, 2022; Zhang et al., 2022). This improves the training speed of the model and helps mitigate the vanishing gradient problem during training. Since GRU can effectively capture long-term dependencies in time series data, it has achieved good performance in applications of crop yield estimation (Gopi and Karthikeyan, 2023; Ren et al., 2023b).* ***The model was constructed based on TensorFlow 2.6 in this study. More details about GRU could be found in Cho et al., (2014).***

The GRU model was trained based on the multi-scenario soybean growth dataset generated by the WOFOST model. We firstly determined the features sensitive to soybean yield, and then extracted the corresponding features and soybean yield from soybean growth dataset, which were used as the GRU's input

and output to train the GRU model. Through this approach, the connection between the WOFOST model and the GRU model was effectively established. *As a crucial state variable in WOFOST, LAI signifies the photosynthetic capability of crops and can effectively characterize the potential yield (Huang et al., 2015). Accounting for the computational efficiency of the model in large areas, the average value of LAI was used as the input feature of GRU. To better capture the growth dynamics of soybean, the mean LAI at vegetative (from emergence to flowering) and reproductive (from flowering to maturity) growth period in the soybean growth dataset were calculated, represented as* LAI_{mean1} *and* LAI_{mean2}*, respectively. The two stage LAI values derived from the simulations,* ***serving as temporal input features,*** *were then combined with model simulated soybean yield in knowledge base,* ***serving as model output,*** *to train the GRU model. The simulated soybean growth dataset was splits into training and testing datasets using 10-fold cross-validation* ***for model training.*** *The hyperparameters of GRU was optimized using a grid search method (Açikkar, 2024). The root mean squared error (RMSE, Eq. (5)) was applied to assess the predictive performance of different set of hyperparameters. After optimization of each fold, the hyperparameters that yielded the smallest predictive error were selected as the optimal ones.* ***The trained GRU model was then coupled with remote sensing data for large-scale soybean yield mapping by inputting the feature variables inversed with remote sensing data.***

11. Why is MODIS data mentioned again in Line 315? MODIS data was not mentioned in the data collection section.

Reply: Thank you for pointing out this. This is a writing error, and we apologize for the confusion. The term "MODIS data" in Line 315 actually refers to the MODIS LAI product (MCD15A3H), which was used in the study. We have corrected this in the revised manuscript to avoid any misunderstanding (Line 317 - 322).

In this framework, the MODIS LAI data, MCD15A3H, was used for intercalibration. Due to its higher temporal resolution and broader image coverage, MCD15A3H generally provided more continuous estimation results. For correction, yield maps were also generated using MODIS LAI products. We utilized the estimation results from MODIS LAI to calibrate the mean yield for each Sentinel-2 tile. The difference between the mean value of the yield derived from MODIS LAI for the region cover the tile and the initial Sentinel-2 estimations was then added to the yield values from Sentinel-2.

12. As shown in Figures 6, 7, and A2, the model simulation accuracy is below industry standards.

Reply: Thank you for pointing this out.

We have made clarifications in the revised manuscript. While the accuracy of our model may appear to be below industry standards in Figures 6, 7, and A2, we have provided a more comprehensive evaluation of its performance in comparison to other studies. The comparison between the performance of our study with that of other studies at multi-scales (field and municipal scale) showed that our method outperformed existing approaches in terms of accuracy. Moreover, we compared our results with soybean yield datasets from other countries with similar resolution. Results showed that our dataset demonstrated superior accuracy. We acknowledge that some studies based on UAV and RGB data have reported higher accuracy for soybean yield estimation. However, these methods are limited by challenges related to data acquisition and high costs, making them suitable only for individual plant or field scale analysis. This limits their applicability for large-scale studies.

The primary goal of this study is to provide a hybrid modeling approach that enables rapid and large-scale soybean yield estimation. The method we produced balances computational efficiency, accuracy, and high resolution, making it suitable for regional-scale and field scale applications. This approach represents a practical solution for large-scale yield estimation despite the lower accuracy compared to some high-cost methods.

Details of revision could be found in Line 469 – 478.

Accurate monitoring of soybean yield is crucial for food policy decision-making and security assessment. Previous studies have primarily focused on the impact of various factors (e.g., climate) on soybean yield (Guo et al., 2022; Zhao et al., 2023a). To our knowledge, high-resolution soybean yield dataset is currently unavailable in the main production regions of China. The study combined crop growth model with deep learning to construct a hybrid model driven by data and knowledge simultaneously for soybean yield estimation. The model retained its data mining capabilities while incorporating mechanistic constraints, thereby enhancing the model's interpretability and transferability. Accuracy verification based on in-situ and statistical data showed that the NortheastChinaSoybeanYield20m generated in this study accurately estimated soybean yield at both field and regional scales (Fig. 5 and 6). ***Compared to existing studies combining remote sensing with process-based model for soybean yield estimation (e.g., Baup et al., (2015), reporting estimation errors of 2 -18%), our method achieved comparable ccuracy. Notably, our framework outperformed existing approaches at both field and regional scales. At the field scales, the assessments obtained RMSE of 287.44 kg ha^{-1} that surpass the RMSE of 400.946 kg ha^{-1} reported by Ren et al., (2023a). At the municipal scale, the accuracy was RMSE of 272.36 kg ha^{-1} that was lower than RMSE of 420.00 kg ha^{-1} reported by Von Bloh et al., (2023). Additionally, the NortheastChinaSoybeanYield20m dataset demonstrated superior accuracy compared to similar resolution soybean yield datasets from other countries (Song et al., 2022). While UAV -based RGB data achieved higher point-scale accuracy (Li et al., 2021, 2024), their reliance on costly, localized data acquisition limits scalability. The method developed in this study strikes a balance between computational efficiency, spatial resolution and accuracy, offering a practical solution for large-scale yield estimation.***

13. By the way, Line 240:" 3.1.2 Multi-scenarios crop simulations", author said:" The four different types of model parameters were arranged and combined to generate various simulation scenarios". Where could I read the scenario settings and the results of this part in the manuscript?

Reply: Thanks for your suggestion.

In the revised version, we have provided a more detailed description of Section 3.1.2: Multi-scenario crop simulations. The study simulated different soybean growth scenarios by fully configuring four input parameters of the WOFOST model: meteorological parameters, soil parameters, crop-specific parameters, and agro-management parameters. The meteorological parameters were derived from observational data collected over 42 years from 51 meteorological stations, while 4 soil types, 5 crop varieties, and 4 agro-managements were defined. By combining different parameter types (similar to a lookup table approach), we inputted these parameter combinations into the WOFOST model to simulate various scenarios. To more clearly describe the scenario simulation process, we have added a Table 3 in the revision, which outlines the scenario settings in detail. Furthermore, we have corrected a numerical error in the revised version. The total number of simulated scenarios generated by the parameter combinations exceeds 170,000, rather than 80,000 as previously stated. We sincerely apologize for the oversight in the earlier version.

We have conducted a thorough review of the revised version to ensure that similar errors have been avoided and the accuracy of the content is maintained.

After parameter preparation, a soybean growth dataset was constructed through model simulations which accounted for the multi-scenarios in agricultural production. The four different types of model parameters (meteorological parameters, soil parameters, crop-specific parameters and agro-management parameters) were arranged and combined to generate various simulation scenarios (Table 4). The scenarios were then put into the model for simulation. Finally, a dataset containing more than 170,000 (51×42×4×5×4) available simulations were generated.

Table 3 Scenarios for WOFOST simulations

Parameters	Number of categories	Details
Meteorological parameters	51×42	Meteorological data from 51 stations over 42 years (1980 – 2021)
Soil parameters	4	Sandy loam, light loam, medium loam and heavy loam
Crop-specific parameters	5	Early maturity, medium-early maturity, intermediate maturity, medium-late maturity and late maturity
Agro-management parameters	4	Four planting dates 20 April, 30 April, 10 May, and 20 May