Dear Editor and Reviewers,

We revised the manuscript in accordance with the reviewers' comments and carefully proofread the manuscript to minimize typographical, grammatical, and bibliography errors. Here below is our point-by-point reply to the comments.

Reply to Reviewer #1

This paper introduces the aboveground biomass data of 250 m spatial resolution grassland on the Qinghai-Tibetan Plateau from 2000 to 2019. The data are of great significance for studying the grassland carbon budget, the interaction between grassland vegetation and climate change, and the construction of ecological civilization on the Qinghai-Tibetan Plateau. A few suggestions or comments on the data paper.

Response: We appreciate your insightful comments on our paper. The comments provided have been extremely helpful to us. We have revised the manuscript in response to your comments and carefully proofread the manuscript to minimize typographical, grammatical, and bibliography errors. The point-to-point responses to your comments are listed below in **blue**.

Point 1. Complementing the sources of study region boundary data, the boundary data from different sources are slightly different.

Response: Thank you for your comments. As you suggested, we have added the sources of boundary in section 2.1 (Lines 90-91):

"In this study, the boundary of the QTP of China (Zhang et al., 2014) was downloaded from the National Earth System Science Data Center, National Science & Technology Infrastructure of China (http://www.geodata.cn)."

Point 2. The study area of this data article is the Qinghai-Tibetan Plateau in China, and it is suggested that the area outside of China should also be taken into account to calculate the grassland aboveground biomass as a whole.

Response. Thank you for your comments. In this study, we only conducted field surveys on the Qinghai-Tibet Plateau in China and did not collect aboveground biomass samples in other regions. In the future, we will continue to expand the sample collection area to realize the inversion of grassland AGB across the Third Pole. To avoid confusion, we have modified the title to (lines 1-3):

"A 250m annual alpine grassland AGB dataset over the Qinghai-Tibetan Plateau (2000-2019) in China based on in-situ measurements, UAV images, and MODIS Data."

Point 3. How were sparse grassland, Alpine grassland, and Alpine meadow classified in Figure 1? What was the basis or source of the classification? Similarly, how was the distribution of grasslands on the Qinghai-Tibetan Plateau determined in this data article? The topography of the Qinghai-Tibetan Plateau is complex, and it is a critical task to determine the extent of the 250-m resolution grassland distribution, which is also related to the use of this data. If the research results of others were used, please give the relevant sources and accuracy.

Response: Thank you for your comments. As you suggested, we have added the detailed information on the grassland classification map in Section 2.1 (lines 92-96):

"Grassland type data was derived from the 1:1000000 Chinese digital grassland classification map provided by the China Resource and Environmental Science and Data Center (https://www.resdc.cn/). This data set, generated through field surveys in the 1980s and supplemented by satellite and aerial imagery, is the most detailed grassland-type map available. For comparison with others, we combined the grassland types into three categories: alpine meadow, alpine grassland, and sparse grassland, and resampled to 250 m (Table A1)."

Table A1. Combined grassland types

New grassland type	Original grassland type
Alpine meadow	Alpine meadow, Lowland meadow, Montane meadow,
Alpine steppe	Temperate steppe, Alpine steppe, Alpine meadow steppe
Spare grassland	Temperate steppe desert, Alpine desert

"

Point 4. The content and depth of discussion in this data article is far from adequate and it is suggested to be strengthened. Scholars have published a large number of reviews and research results in the field of estimating grassland biomass using remotely sensed data. It is suggested to read carefully to enhance the discussion of this data article.

Response: Thank you for your comments. As you suggested, we have made the following changes.

- We have modified the original Figure 8 by adding an eco-geographic region layer, and conducted a statistical analysis of interannual trends in eco-geographic regions. In addition, we have added Table A5 in the Appendix section to describe the abbreviations and full names of

the eco-geographic regions (lines 328-333).

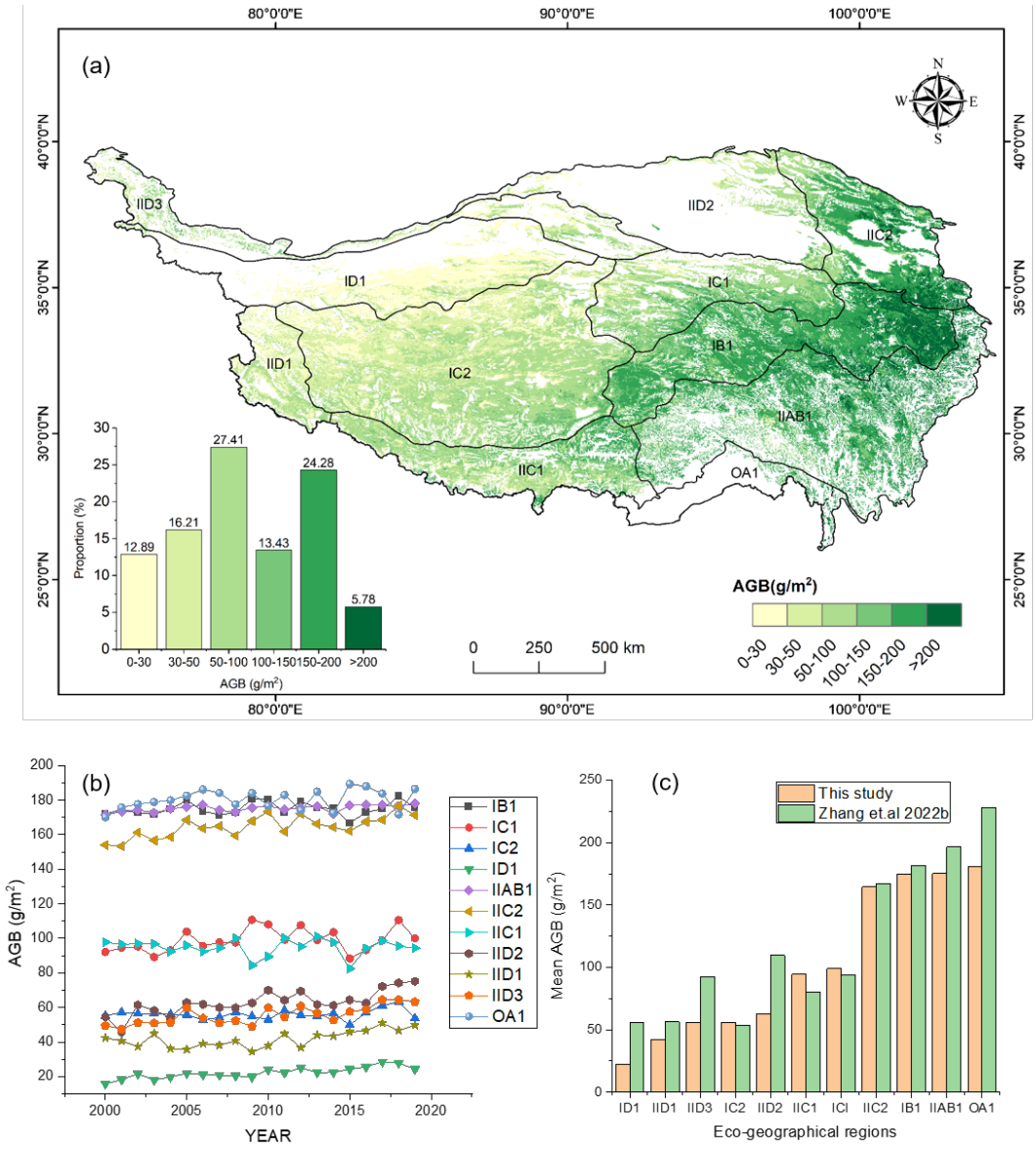

Figure 8. (a) The spatial distribution of average grassland AGB on the QTP from 2000 to 2019. IID1, IID2, IID3, ID, IIC1, IIC2, IC1, IB1 IIAB1, and OA1 are the eco-geographical regions of the QTP(Zheng, 1996). The full names of each eco-geographical region were listed in Table A5. (b) AGB values of each eco-geographical region from 2000 to 2019. (c) Comparison of multi-year AGB averages in the different eco-geographical regions.

Table A5: List of abbreviations of eco-geographical regions and the mean AGB of the QTP

Abbreviation	Full name
IB1	Golog-Nagqu high-cold shrub-meadow zone
IIAB1	Western Sichuan-eastern Tibet montane coniferous forest zone
IC1	Southern Qinghai high-cold meadow steppe zone

IC2	Qiangtang high-cold steppe zone
ID1	Kunlun high-cold desert zone
IIC1	Southern Tibet montane shrub-steppe zone
IIC2	Eastern Qinghai-Qilian montane steppe zone
IID1	Nagri montane desert-steppe and desert zone
IID2	Qaidam montane desert zone
IID3	Northern slopes of Kunlun montane desert zone
OA1	Southern slopes of Himalaya montane evergreen broad-leaved forest zone

- We analyzed the changes of grassland AGB on the Qinghai-Tibet Plateau from 2000 to 2019 using the Theil-Sen Median trend analysis and the Mann-Kendall test. And Section 2.7 was added to introduce the trend analysis method (lines 246-253):

  "2.7 Trend analysis of grassland AGB
  This study combined the Theil-Sen median trend analysis and Mann-Kendall test to analyze the temporal variation characteristics of grassland AGB of QTP (Jiang et al., 2015). Theil-Sen median trend analysis is a robust trend statistical method with high computational efficiency, insensitive to outliers (Hoaglin et al., 1983). The Mann-Kendall test is a nonparametric test for time series trends, which does not require the measurements to follow a normal distribution and is not affected by missing values and outliers. The Theil-Sen Median trend analysis and Mann-Kendall trend test have been widely used to analyze vegetation index, cover, and biomass(Gao et al., 2020; Jiang et al., 2015; Fensholt et al., 2009). The formulas for the Theil-Sen median trend analysis and the Mann-Kendall method are detailed in Jiang et al. (2015)."

- In Section 3.4, we have added a description of the spatial distribution and trends of the AGB of the QTP from 2000 to 2019 based on the eco-geographical regions (lines 315-325):

  "The spatial distribution of the average grassland AGB on the QTP from 2000 to 2019 was calculated (Figure 8). The AGB gradually increased from west to east. As shown in Figure 8b, the average biomass of eastern OA1, IIAB, IB1, and IIC2 eco-geographical regions ranged from 150 to 190 g/m^2, and the average AGB of IC1 and IIC1 ranged from 80 to 110 g/m^2. The average AGB of IID2, IID3, IC2, and IID1 in the west was relatively low, ranging from 35 to 75 g/m^2. The ID1 region was dominated by sparse grassland with the lowest average interannual AGB values, which fluctuated around 20 g/m^2 (Figure 8b). The average AGB of QTP showed an insignificant increasing trend between 2000 and 2019, with an average growth rate of 0.22 $gm^{-2}a^{-1}$ (Figure 9a). The overall mean AGB of the QTP was 103.6 g/m^2, with 151.85 g/m^2, 60.85 g/m^2, and 28.91 g/m^2 for alpine meadow, alpine steppe, and sparse grassland, respectively (Figure 9b). In addition, the temporal trend of grassland AGB in each pixel was analyzed. As shown in Figure 10, the IID3, ID1, IID2, and IIC2 eco-geographical regions of the northern QTP showed an increasing trend from 2000 to 2019, while the IC2, IB1, and IIC1 regions showed some degradation. Therefore, there was spatial heterogeneity in the temporal variation."

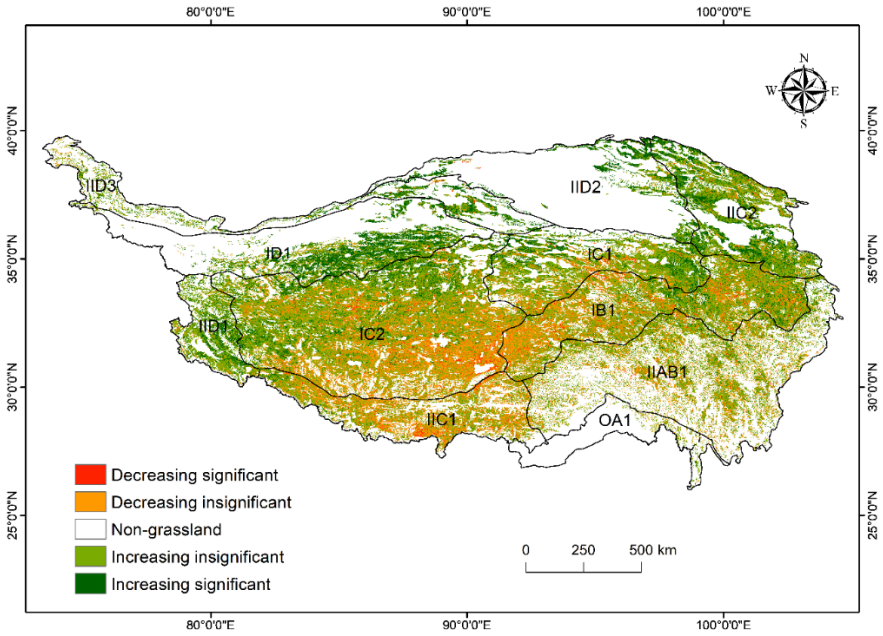

Figure 10. Spatial trends of grassland AGB on the Tibetan Plateau from 2000 to 2019. IID1, IID2, IID3, ID, IIC1, IIC2, IC1, IB1 IIAB1, and OA1 are the eco-geographical regions of the QTP (Zheng, 1996). The full names of each eco-geographical region are listed in Table A5.

- In section 4.1, a discussion on spatial scale matching has been added (lines 346-350).

  "In previous studies, the AGB value of a satellite pixel was represented by the average value of 3-5 quadrat-scale samples, so there is a large spatial gap between the ground samples and the satellite pixel (Yang et al., 2017; Yang et al., 2009; Meng et al., 2020). The spatial gap between ground samples and satellite indices can impact the accuracy of grassland AGB models. The smaller the spatial gap between the two, the higher the accuracy of the model (Morais et al., 2021). We address this issue using the UAVs as a bridge to reduce the spatial gap."

- In Section 4.3, the original Table 4 was updated. The "Data Source" column in the original table was changed to "input parameter". In addition, three new references were added to the list (lines 403-404).

Table 5: Comparison of AGB estimation results of different studies on the QTP

Mean AGB (g/m^2)	Alpine steppe (g/m^2)	Alpine meadow (g/m^2)	Study period	Approach	Input parameter	References
68.8	50.1	90.8	2001-2004	Linear regression	EVI	(Yang et al., 2009)
__	22.4	42.37	2000-2012	Linear regression	NDVI	(Liu et al., 2017)
120.73	__	__	1980–2014	Exponential regression	NDVI	(Jiao et al., 2017)
78.4	__	__	1982-2010	RF	NDVI, climate	(Xia et al., 2018)
77.12	76.43	154.72	2000-2014	RF	NDVI, EVI, climate, terrain	(Zeng et al., 2019)
59.63	42.75	77.56	2000-2017	RF	NDVI, climate	(Gao et al., 2020)
102.4	__	__	2000-2020	RF	climate,soil,and terrain	(Zhang et al., 2022b)
70.00	__	__	1960–2002	Century	climate and soil data	(Zhang et al., 2007)
119.78	__	__	2002–2004	Orchidee	climate, soil and LAI data	(Tan et al., 2010)
103.6	60.85	151.85	2000-2019	RF	MODIS	this study

[1] ZHANG, X., LI, M., WU, J., HE, Y., and NIU, B.: Alpine Grassland Aboveground Biomass and Theoretical Livestock Carrying Capacity on the Tibetan Plateau, Journal of Resources and Ecology, 13, 129-141, 2022b.4

[2] Zhang, Y. Q., Tang, Y. H., and Jiang, J. A.: Characterizing the dynamics of soil organic carbon in grasslands on the Qinghai-Tibetan Plateau, 2007.4

[3] Tan, K., Ciais, P., Piao, S., Wu, X., Tang, Y., Vuichard, N., Liang, S., and Fang, J.: Application of the ORCHIDEE global vegetation model to evaluate biomass and soil carbon stocks of Qinghai-Tibetan grasslands, 2010.4

- In Section 4.3, the following four aspects of discussion have been added.

  ✧ First, the comparison of scale matching between dependent and independent variables was added (lines 390-395).

  "At the pixel scale, compared with other studies, this paper achieved the spatial scale matching of independent and dependent variables during the modeling. In previous studies (Yang et al., 2009; Yang et al., 2017; Meng et al., 2020), they constructed the models from the measured AGB values at the quadrat-scale and the spectral indices of the satellites without considering the spatial scale difference. It partly explained why the R^2 of the AGB linear model constructed by Yang et al. was only 0.4 (Yang et al., 2009). Our results confirmed that the R^2 of the linear model could be increased from 0.29 to 0.78 after reducing the spatial gap between measured AGB and NDVI (Figure 7)."

  ✧ Second, a comparison of model validation methods has been added (lines 395-402).

  "In addition, thanks to the rapid sampling of UAV AGB, a total of 2602 samples matching the pixel scale were collected during 2015-2019. It allowed us to perform cross-year validation to assess the robustness of the model over time, which has rarely been performed

in previous studies. Our results showed similar validation results for 2017-2019 (R^2=0.85, p<0.001) despite different sample sizes and spatial distributions (Figure 1, Table 1). But in 2015-2016, R^2 was relatively low, at 0.63 and 0.77, respectively (Table 3, Figure 6). The reason was that during 2015-2016, some photos with abnormal white balance were obtained due to improper settings, which reduced the estimation accuracy (Figure A5). The validation results showed that the pixel-scale AGB estimation model had good robustness in different regions and times when the photo quality was acceptable."

✧ Third, we added the spatial trend comparison of AGB on QTP at the eco-geographical region scale (lines 410-419).

"The spatial distribution of AGB was consistent with previous studies, showing a west-to-east increasing trend (Zhang et al., 2022b; Xia et al., 2018). Specifically, the average AGB of OA1, IIAB, IB1, and IIC2 eco-geographical regions in the east was significantly higher than that of IID2, IID3, IC2, IID1, and ID1 regions in the west (Figure 8). In general, the average AGB estimates for each eco-geographical region in this paper were not much different from those of Zhang et al. (2022b). Among them, our average AGB estimates for ID1, IID1, IID3, and IID2 regions were slightly lower, but our values were closer to the measured values of these regions (Figure 8c). The reason may be that they calculated the potential AGB, while we calculated the actual AGB, so our estimate was relatively low. In terms of spatial and temporal trends, the data results showed that the eco-geographical regions in the northern part of the QTP demonstrated an increasing trend (IID3, ID1, IID2, and IIC2), while the IC2, IIC1, and IB1 regions exhibited significant or non-significant decrease, which was consistent with the results of others (Gao et al., 2020; Liu et al., 2017)."

✧ In addition, the reasons for the differences between the results of this study and those of previous studies were discussed. We discussed this through three aspects: the measured samples, the model input parameters, and the modeling approach (lines 421-440).

"The difference between our estimated grassland AGB and previous studies might be due to differences in data sources and modeling methods. Firstly, the sample size and spatial distribution of ground samples were different. The number of ground samples is the most important variable affecting the accuracy of the grassland AGB estimation model (Morais et al., 2021). Unlike previous studies, we collected ground validation data by combining the traditional sampling method and UAVs. The newly proposed method could overcome the shortcomings of traditional samplings (time-consuming and labor-intensive). It no longer takes years to obtain spatially representative, large-scale ground validation data (Yang et al., 2017). With UAV sampling, ground observations matching the satellite pixel scale can be obtained in only 15-20 minutes, which is difficult to achieve in traditional surveys. Our new sampling method not only accelerates the sampling speed and increases the sample size, but also improves the spatial match between ground samples and satellite

pixels. As a result, our ground validation data is superior to previous studies in terms of quantity and spatial match to the satellite data. Secondly, the input parameters of AGB estimation models were different. Some scholars used only a single vegetation index (NDVI or EVI), while others combined the vegetation index with meteorological, soil, and terrain indices to construct the AGB estimation models (Table 5). In this study, NDVI, kNDVI, EVI, DEM, and PREC were used as the final predictor variables to construct the AGB estimation model at the pixel scale (Table 2). Thirdly, modeling methods might also affect the simulation results. As shown in Table 5, the overall AGB averages of the QTP estimated based on different methods (such as linear or nonlinear regression, machine learning, and ecological process model methods) varied considerably. Yang et al.(2017) found that the model performance of ANN was much better than the linear regression model when using the same dataset to estimate grassland AGB in the Three-River Headwaters Region of China. Jia et al.(2016) reported that the model forms could bring 13% uncertainty to the AGB estimation. Wang et al. compared the RF with the support vector regression (SVR) machine learning algorithm and found that the RF yielded the best performance in grassland biomass estimation (Wang et al., 2017)."

- In Section 4.4, we have added a discussion of the underestimation of the AGB estimation models, as well as possible reasons and the future research direction (lines 443-456):

"We acknowledge that there are some shortcomings in this study. 1) The predicted values of the quadrat-scale model were underestimated when the measured biomass values were greater than 250 g/m^2 (Figure 6). One reason may be that the number of samples greater than 250 g/m^2 was relatively small, accounting for only 5.18 % of all samples. Another reason may be that for high biomass grasslands, a single UAV RGB photo can only reflect information such as vegetation cover and greenness, but not height information. This feature is very unfavorable for estimating AGB in grassland areas with high vegetation coverage and height. Studies have shown that adding vegetation height information can help improve the estimation accuracy of grassland AGB (Zhang et al., 2022a; Lussem et al., 2019; Viljanen et al., 2018). In future work, an affordable DJI Zensil L1 Lidar drone will be introduced to invert the height of grassland. 2) At the pixel scale, limited by the estimation accuracy of AGB from UAV, there was also some underestimation in the high biomass area. Although the MODIS index closest to the sampling time was chosen in the construction of the AGB estimation model, there was still a time difference between the measured samples and the MODIS indices, which might lead to estimation errors. In addition, the NDVI saturation problem was not considered in this study, which might affect the AGB estimation accuracy of QTP (Tucker, 1979a; Gao et al., 2000; Mutanga and Skidmore, 2004; Tucker, 1979b). In the next step, we will continue to collect samples with high biomass and try to correct the NDVI saturation problem to optimize the simulation accuracy of the data set."

Point 5. Accuracy evaluation should not only pay attention to R^2 and RMSE, but also pay attention to the relationship between the regression line and 1:1 line. In the regression analysis in Figure 6, the relationship between the regression line and the 1:1 line was not discussed, although the authors gave the 1:1 line. It should be noted that the regression line is meaningful only if there is no significant difference between the regression line and the 1:1 line. This information was not given in the paper, and there was a significant lack of accuracy evaluation in the article. This is directly related to the reliability of the regional results.

Response: Thank you for your comments. As you suggested, we changed as following:

- In the revised version, we have added the formulae for calculating R^2 and RMSE in Section 2.5.1 (lines 204-208)

  "

  Statistical metrics R^2 (Eq.1) and RMSE (Eq.2) were used to evaluate the performance of the model.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(\hat{y}_i - \bar{y}_i)^2} \quad (1)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}} \quad (2)$$

  where n is the number of samples, y_i and $\hat{y}_i$ represent the measured and the predicted AGB value, respectively, $\bar{y}_i$ is the mean value of measured AGB samples.

  "

- We modified Table 3 by adding significance level information of R^2 (lines 284-285).

Table 3: Validation results of AGB models at quadrat and pixel scales

Scale	Year	Training set		Validation set	
		R^2	RMSE(g/m2)	R^2	RMSE(g/m^2)
Quadrat-scale	2019	0.94	20.18	0.73 ***	32.94
Pixel-scale	2019	0.96	10.68	0.85 ***	23.36
	2018	—	—	0.85 ***	24.83
	2017	—	—	0.85 ***	23.83
	2016	—	—	0.77 ***	31.28
	2015	—	—	0.63 ***	34.07

'***' significant at $p<0.001$

- Moreover, regression analysis was performed to verify the strong linear relationship between the predicted and measured AGB values. In the revised version, we have added a new Table A4 to describe the slope and intercept coefficients of the regression model and the t-test results for these coefficients in the Appendix section (line 526):

Table A4: Regression analysis for AGB estimation models at quadrat and pixel scales

Model name	Coefficient	Value	Standard Error	t-Value	p-value
2019_Quadrat-scale	Slope	0.67	0.016	42.58	9.05e-194
	Intercept	20.10	1.49	13.59	5.96e-37
2019_Pixel_scale	Slope	0.84	0.03	31.59	2.75e-73
	Intercept	23.20	4.04	5.74	4.24e-8
2018_Pixel_scale	Slope	0.73	0.02	45.81	8.28e-157
	Intercept	20.43	2.74	7.46	6.01e-13
2017_Pixel_scale	Slope	0.75	0.01	59.13	1.98e-260
	Intercept	13.89	2.04	6.82	2.19e-11
2016_Pixel_scale	Slope	0.94	0.02	40.45	4.69e-157
	Intercept	2.48	3.75	0.66	0.03
2015_Pixel_scale	Slope	0.82	0.04	18.88	2.59e-47
	Intercept	9.50	5.25	1.81	0.04

- In this study, we used the RF model to predict AGB values rather than through a linear regression equation. The linear regression equation constructed in Figure 6 was only used to assess whether there was a significant linear relationship between the estimated and measured AGB values at the quadrat or pixel scale. In the new version, a Student's t-test was used to test whether there was a statistically significant difference between predicted and measured AGB values at the 95% confidence level (Table 4, lines 288-289).

"Table 4: T-test results between the predicted and measured AGB values for the modes at quadrat and pixel scales

Validation model	Measured mean	Predicted mean	t	df	p-value
2019_Quadrat-scale	51.57	54.35	-0.66	939.35	0.51
2019_Pixel_scale	136.68	137.7461	-0.15	340.78	0.88
2018_Pixel_scale	152.49	131.48	4.01	723.81	6.63e-05
2017_Pixel_scale	141.42	120.60	5.48	1225.2	5.26e-08
2016_Pixel_scale	149.56	142.70	1.68	961.99	0.09413
2015_Pixel_scale	108.65	98.23	1.96	1225.2	0.05

"

- In this version, we have modified Figure 6. Indexes such as R^2, p-value, RMSE, and the number of validation samples (N) have been added to quantify the difference between the predicted and measured AGB values.

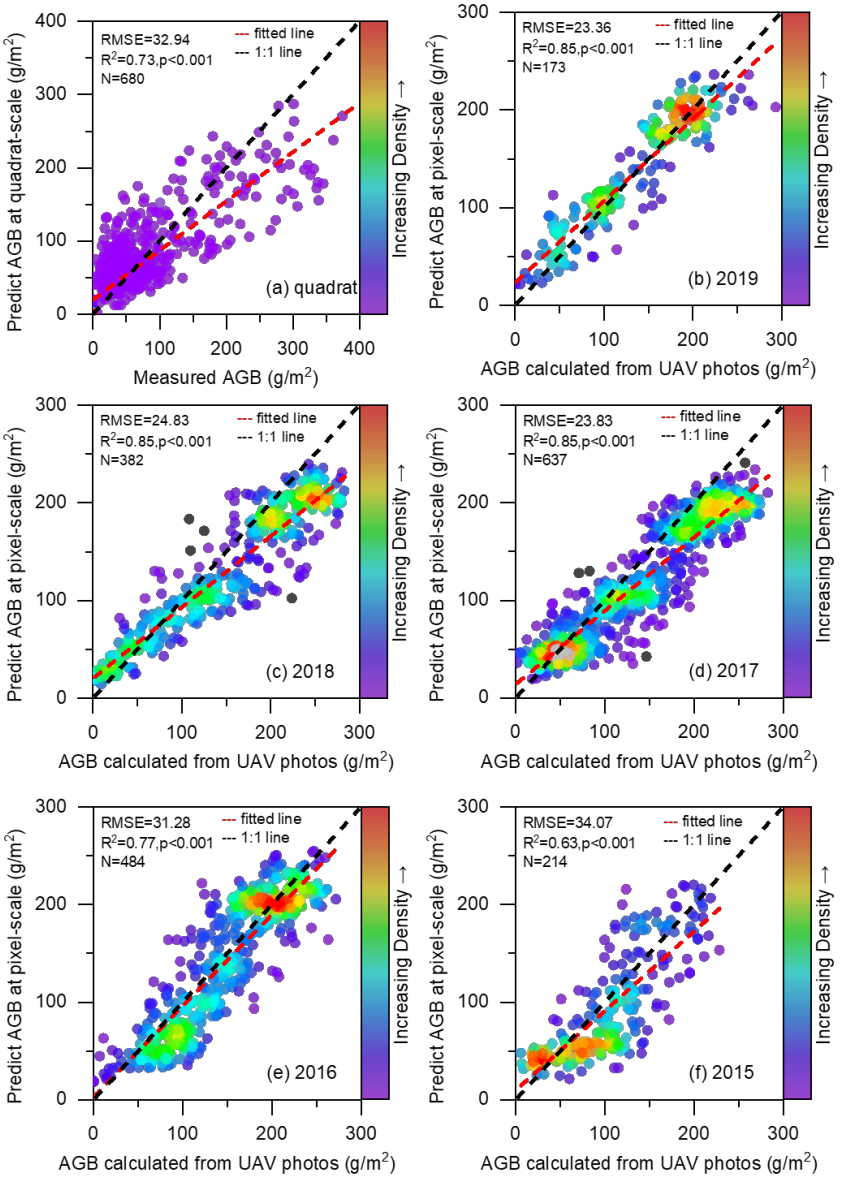

Figure 6. Validation results of the AGB estimation models at the quadrat (a) and MODIS pixel scale for 2015-2019 (b~f).

- We have also modified the text in section 3.2 (lines 266-281):

"For the AGB estimation model at the quadrat scale, the results of 10-cross validations showed that there was a significant linear relationship between the estimated and the measured values (R^2=0.73, p<0.001, Table 3, Table A4). The student's t-test was also used to assess whether there was a significant difference between the predicted AGB values and the measured values at a confidence level of 95%. As shown in Table 4, there was no significant difference

($p=0.51>0.05$) with an RMSE of 32.94 g/m^2. The scatter plot showed that the model predicted well when the measured biomass was less than 150g/m^2, but showed some underestimation when it was more than 200g/m^2 (Figure 6a). This may be because the number of samples more than 200g/m^2 is relatively small, accounting for only 8.50% of all samples (Figure 5a). Although the sample size of UAVs varied from year to year, most of the AGB values estimated from photos ranged from 0 to 300 g/m^2 (Figure 5b).

For the pixel-scale AGB estimation model, there was a strong linear relationship between the predicted AGB and UAV estimates for 2015-2019 (Table A4). The fitting coefficient R^2 was 0.85 for 2017-2019, slightly lower for 2015-2016 at 0.63 and 0.77, respectively (Table 3, Figure 6b-f). The RMSE of the pixel-scale model ranged from 23.36 to 34.07 g/m^2 (Table 3). In addition, we found no significant differences between the predicted and measured average AGB values except for 2017 and 2018 (Table 4). While the average model projections for 2017 and 2018 were 14.72% and 13.78% lower than the UAV estimates, they were within acceptable ranges. Therefore, the constructed pixel-scale AGB estimation model had good performance and robustness in different years (Figure 6b~f).”

Point 6. The English of the article is seriously substandard, with problems such as too many grammatical errors, and even some sentences lack obvious sentence components.

Response: Thank you for your comments. Based on your suggestion, we have carefully proofread the manuscript to minimize typographical, grammatical, and bibliography errors. For specific changes, please refer to the revised version with change tracking.

Reply to Reviewer #2

This paper developed a new annual 250-m grided AGB dataset from 2000 to 2019 over the Qinghai-Tibetan Plateau by using ground in-situ measurements, multi-years UAV images and MODIS datasets. UAV images were used as the bridge to overcome scale mismatches between ground samples and coarse MODIS satellite pixel scales. Many efforts have been devoted on UAV observation works as well as field sample works at large region scales. In general, I think this is a good paper and is within the scope of ESSD. I have a few comments for authors' consideration.

Response: We appreciate your insightful comments on our paper. The comments provided have been extremely helpful to us. We have revised the manuscript in response to your comments and carefully proofread the manuscript to minimize typographical, grammatical, and bibliography errors. The point-to-point responses to your comments are listed below in **blue**.

Point 1. In table 1, the acquisition times of UAV sampling and field sampling of AGB in 2019 were mainly in the growth season and therefore the UAV estimations of AGB also in this season. However, the MODIS pixel level vegetation indices were composited by MVC method which can reflect the best grown condition in the whole year. Aboveground biomass may still be available for several months after sampling. The temporal mis-matches between field work and MODIS composites may lead to estimation errors.

Response: Thank you for your comments. As you pointed out, the UAV photos and measured AGB samples we acquired in 2019 were mainly from July to August.

- In constructing/validating the AGB estimation model at the pixel scale, the MODIS NDVI, EVI, and kNDVI indices closest to the sampling time were chosen to minimize the difference between sampling and satellite acquisition times.

  As you suggested, we added an explanation in section 2.4.2 (lines 162-166):

  "After that, the corresponding vegetation indices closest to the time of the UAV sampling were extracted to construct/validate a pixel-scale AGB estimation model. In addition, the kNDVI index was calculated to overcome the NDVI saturation issue based on the equation kNDVI= TANH ($NDVI^2$) (Camps-Valls et al., 2021)."

  We also explained it in section 2.5.4 (lines 232-234):

  "The MODIS vegetation indices and other spatial metrics corresponding to each GRID or RECTANGLE mode were then extracted using the ArcGIS software. Here, the MODIS NDVI, EVI, and kNDVI indices closest to the sampling time were chosen to minimize the time

difference between sampling and satellite overpass."

- The annual maximum vegetation indices were used as the input parameters for the AGB inversion of the entire QTP region. As in previous studies, we used MVC method to obtain the maximum value of annual vegetation index (NDVI, EVI and kNDVI), and assumed that AGB corresponding to this period also reached the maximum value in the growing season. We explained this in Section 2.4.2 (lines 165-166).

  "The annual maximum vegetation indices were calculated by the maximum value composition (MVC) algorithm to estimate the spatial AGB distribution of QTP from 2000 to 2019 (Holben, 1986; Wang et al., 2021; Gao et al., 2020)."

- Although the vegetation index closest to the sampling time was selected when constructing the pixel scale estimation model, there was still a time difference between the ground samples and MODIS indices, which would lead to estimation error. We pointed out this limitation in section 4.4 (lines 451-453)

  "Although the MODIS index closest to the sampling time was chosen for the construction/validation of the AGB estimation model, there was still a time difference between the measured samples and the MODIS indices, which might lead to estimation errors."

Point 2. In section 2.3.2, the BELT flying mode were used for three GRID routes and four ground sampling quadrats were sampled in the BELT routes. However, in Figure 2 and Section 2.5, how were the BELT images at 2-m height used was not introduced. It seems only 20-m UAV images were used for development of UAV estimation model.

Response: Thank you for your comments. As you mentioned, we only used 20m high UAV photos when constructing the AGB estimation model at the quadrat scale.

- However, in fieldwork, the BELT flight mode served as a bridge between traditional and UAV sampling at 20m. Compared with GRID mode, the size of BELT was relatively small (40 m×40 m), and the flight altitude and speed were set to 2m and 1m/s, respectively. It ensured the staff had enough time to place a sampling frame directly below the aircraft lens to capture it in the 2m and 20 m UAV photos. We explained this in section 2.3.1 (lines 129-133).

  "The BELT mode is similar to GRID, but is designed to obtain near-ground UAV image data with higher resolution (Figure 3b). It can be combined with the traditional sampling method to ensure the consistency of UAV images with the ground samples (Figure 3d). Typically, the BELT size is set to 40 m × 40 m, and the flying height and speed are set to 2 m and 1m/s to ensure that field crews have enough time to place sampling frames under the UAV waypoints. As with the GRID mode, 16 UAV images can be captured in a single flight."

- Although our previous research confirmed that 2m photographs could be used to model sample-scale grassland AGB, the reasons why 20m high UAV photographs were chosen for the quadrat-scale model in this study were as follows: Firstly, the spatial coverage of a 20m-high UAV photo (26 m×35 m) is much wider than a single 2m-high UAV photo, making it easier to match to the MODIS pixel scale. We have added an explanation in section 2.5.2 (lines 210-216) :

  "Since the spatial coverage of a 20m-high UAV photo (26 m×35 m) is much wider than a single 2m-high UAV photo, making it easier to match to the MODIS pixel scale. Hence, the 20m-high UAV photos containing the sample frames were chosen for constructing the quadrat-scale AGB estimation model. A total of 906 pairs of quadrat-scale UAV-field AGB observation data were collected, with good spatial representativeness (Figure 1 a, red dots). The observed AGB values ranged from 0 to 450 g/m^2, with mean and median values of 59.75 g/m^2 and 33.04 g/m^2, respectively, most of which were less than 100 g/m^2 (Figure 5a). The cropped 20-meter-high UAV image indices and the measured AGB values were used as the independent and dependent variables to build the RF model (Figure 2).
  "

- Secondly, using the 20m-high UAV photo containing the sample frames to construct the quadrat-scale model ensures an exact match of spatial scales between the independent and dependent variables. Furthermore, it facilitates the estimation of the AGB of an entire 20m-high UAV photo. We explained this in section 4.1 (lines 350-357)

  "Spatial scale matching of dependent and independent variables was achieved in estimating AGB values at different scales. First, at the quadrat scale, the independent variables were all derived from cropped 20-meter-high UAV images corresponding to the ground samples (Figure 3e). Then, the 20-meter-high UAV image was cropped into ~2000 quadrat-sized patches to ensure consistency with the quadrat-scale model, and the average of these patches was used as the final AGB at the photo scale. Finally, by averaging the AGB of 16 or 12 UAV photos within the MODIS pixel, the AGB value matching the MODIS pixel scale was calculated (Figure A1). With these three steps, we successfully upscaled the measured AGB from the traditional quadrat scale (0.5 m×0.5 m) to the photo scale (26 m×35 m) and MODIS pixel scale (250 m×250 m).

- In addition, we discussed the quadrat-scale AGB estimation models based on the 2-meter and 20-meter UAV photos in section 4.3 (lines 384-388):

  "At the quadrat scale, consistent with our previous study, we further confirmed that the UAV RGB images could be used to estimate grassland AGB (Zhang et al., 2022a; Zhang et al., 2018). Similar to the 2-meter-high UAV image, the indices from the 20-meter-high UAV image could be used to estimate the grassland AGB at the quadrat scale (R^2=0.73, RMSE=44.23 g/m^2, Figure 6a). Compared with the 2-meter-high UAV image, the 20-meter-high UAV image is more suitable for matching the MODIS pixel due to its wider spatial coverage (26 m ×35 m)."

Point 3. For one MODIS pixel, it seems more than 16 UAV images at 20-m height are needed to cover the whole pixel. I'm not quite understanding the GRID, RECTANGLE and BELT flight modes. Does it mean the UAV only take pictures in the waypoints and there are gaps among those pictures? The authors can explain more about how it works as traditionally we will make overlaps among pictures.

Response: Thank you for your comments. As you mentioned, all three flight modes (GRID, RECTANGLE, and BELT) are designed to shoot only at the pre-set waypoints.

- Compared to the commonly used MOSAIC flight mode (which requires a guaranteed overlap rate between photos to obtain a full view of an area), our designs are more in line with the traditional eco-sampling concept, which can better balance the spatial representation and accessibility of samples for efficient sample collection. An explanation of the three flight modes was added in Section 2.3.1 (lines 126-135):

  "GRID, RECTANGLE, and BELT are the most commonly used flight modes in the FragMap software. GRID and RECTANGLE modes have 16 and 12 waypoints for capturing UAV images within a MODIS pixel range (Figure A1). Their flying height and speed are set to 20 m and 3m/s, respectively. The spatial coverage of a 20-meter-high UAV photo is about 26 m × 35 m. The BELT mode is similar to GRID, but is designed to obtain near-ground UAV image data with higher resolution (Figure 3b). It can be combined with the traditional sampling method to ensure the consistency of UAV images with the ground samples (Figure 3d). Typically, the BELT size is set to 40 m × 40 m, and the flying height and speed are set to 2 m and 1 m/s to ensure that field crews have enough time to place sampling frames under the UAV waypoints. As with the GRID mode, 16 UAV images can be captured in a single flight. Compared with the MOSAIC flight mode (which requires a guaranteed overlap rate between photos to obtain a full view of an area), our design is more in line with the traditional ecological sampling concept. It allows for a better balance of spatial representation and accessibility of samples, resulting in efficient sample collection."

- Thus, for the plot size of the MODIS pixel (250 m × 250 m), it typically takes 40 minutes to complete sampling using MOSAIC mode, while it takes only 10 minutes using GRID and RECTANGLE flight modes. Our flight modes significantly reduce the sampling time and offer the possibility of obtaining more samples matching the MODIS pixel size.

- In addition, to further clarify the impact of the number of UAV samples on the AGB estimation at the MODIS pixel scale, we conducted a comparative analysis of the AGB estimation results using 1 photo to 16 photos in a step-by-step incremental manner. As shown in Figure 7a, when the number of UAV photos increased to 4, the growth rate of the correlation coefficient slowed down and tended to be stable. It indicated that although the 16 photos could not cover the entire MODIS, they were sufficiently spatially representative. The relevant results for this part were

presented in Section 3.3 (lines 294-300):

"Moreover, the correlation between NDVI and UAV-estimated AGB increased with the number of UAV photos. It increased rapidly as the number increased from 1 to 4 (from 0.74 to 0.86), then slowed down and stabilized (from 0.87 to 0.88). In addition, we compared the scatter plots and fitting lines between NDVI and different AGB estimation methods (Figure 7b-f). The results showed a weak linear relationship between the traditionally measured AGB and NDVI, with an R^2 of 0.29. With the UAV sampling method, the linear relationship was greatly improved and increased with the number of photographs. The fit coefficient R^2 increased from 0.54 to 0.78, much higher than the traditional sampling method (Figure 7).

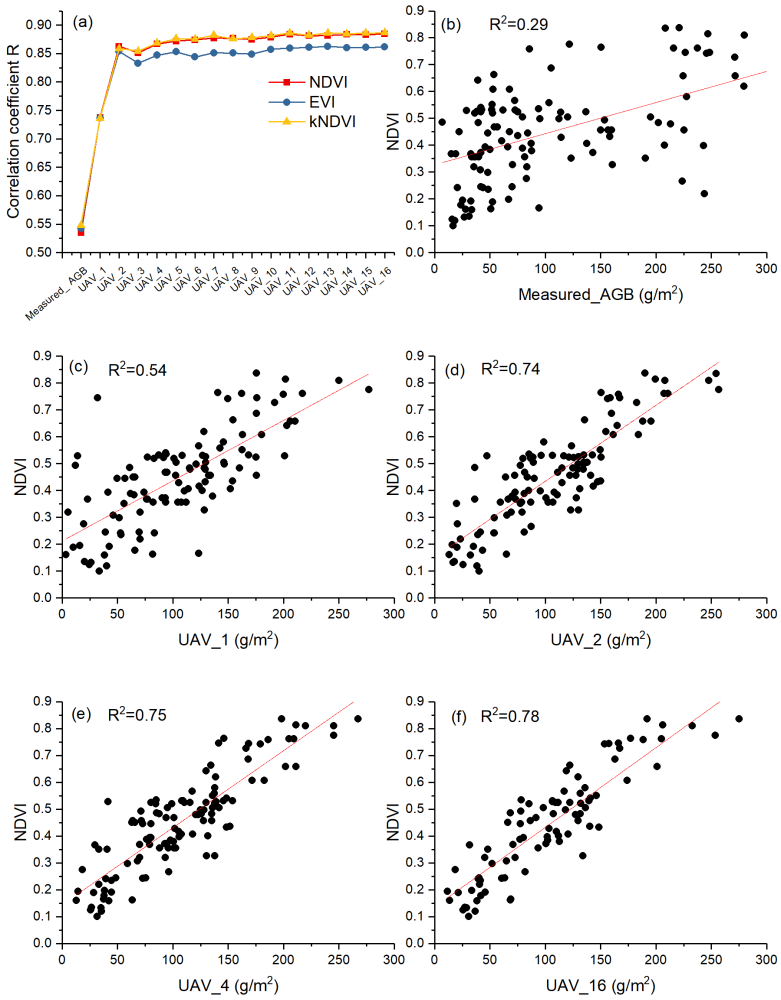

Figure 7. Correlation between MODIS vegetation indices and different AGB estimation methods (a); scatter plots of NDVI with different AGB estimation methods (b-f). UAV_x, x represents the number of UAV photos used to estimate the average AGB at the MODIS pixel scale. Here, x ranges from 1 to 16.

"

- Admittedly, neither GRID nor RECTANGLE flight mode captures the entire MODIS pixel. We discussed this limitation in Section 4.4 and proposed a solution for the next step (lines 461-463):

"Moreover, although the UAV images in GRID or RECTANGLE mode could cover most areas of a MODIS pixel, full pixel coverage was still not achieved. Therefore, we will gradually upscale to MODIS pixels by combining UAVs with Sentinel-2 or Landsat images."

Point 4. Page 18, line 324, "The reason was that the GIRD mode could obtain 16 photos in the MODIS pixel at a time, while the RECTANGLE mode could only take 12 photos". Figure A1(b) should be cited to explain the RECTANGLE model.

Response: Thank you for your comments. As you suggested, we modified it as (lines 364-365):

"The reason is that GIRD mode can take 16 pictures within a MODIS pixel, while RECTANGLE mode only takes 12 pictures (Figure A1)."

Point 5. Page 22, line 407, "AGB_2000.tif represents this TIFF file describing the alpine grassland AGB condition of QTP in 2005" should be in 2000.

Response: Thank you for your comments. As you suggested, we modified it as (lines 468-469):

"For example, AGB_2000.tif represents this TIFF file describing the alpine grassland AGB condition of QTP in 2000."