Responses to comments from anonymous referee #1

This manuscript presents AGPC, a new annual 500-m gridded population dataset for mainland China covering the period 1990–2020. A key strength of the study is that it explicitly incorporates three-dimensional building volume dynamics into the population mapping framework, thereby addressing the density-saturation limitations of conventional two-dimensional population products in vertically developed urban environments. The authors provide comprehensive validation using census and statistical data at multiple administrative levels, and the results demonstrate that AGPC achieves improved spatial representation compared with several widely used population datasets. Overall, the manuscript is well organized, and the proposed dataset will be valuable for studies of population dynamics, urbanization, disaster risk assessment, and sustainable development. I have several comments and concerns that I believe will help further improve the clarity, reproducibility, and interpretation of the manuscript.

Response:

We thank the reviewer for the positive assessment of our manuscript and for recognizing the value of the AGPC dataset. We also appreciate the constructive comments, which have helped us improve its quality and clarity. Accordingly, we have substantially expanded the township-level validation samples, clarified the availability of the GUS-3D dataset and the temporal consistency of the POI and impervious-surface data, and refined the interpretation of SHAP importance across different years. We also clarified the county-level population constraints and corrected the terminology, and editorial issues identified throughout the manuscript. These revisions have improved the robustness, reproducibility, and clarity of the manuscript. Our point-by-point responses are provided below.

1. Section 2.1 indicates that approximately 6,800 townships were used for validation in 2010, whereas only about 3,830 were available for 2020. Please explain the reason for this substantial reduction (e.g., administrative boundary changes or data availability) and discuss whether the smaller validation sample affects the robustness and comparability of the reported validation results.

Response:

We agree that the unequal sample sizes may affect the comparability of the validation results. The difference in the number of township-level validation samples between 2010 and 2020 in the original manuscript was mainly due to data availability and difficulties in spatially matching tabular census records to township boundaries. In particular, duplicated township names, administrative name changes, and boundary adjustments required substantial manual verification. In response, we substantially expanded and manually verified the township-level validation datasets, with particular attention to improving geographical coverage and maintaining comparable sample sizes between the two years. The revised township-level validation samples now contain 10,834 records for 2010 and 10,336 records for 2020, with all provinces in mainland China represented. Accordingly, all township-level validation analyses, statistics, figures, and tables have been recalculated and updated. The expanded and more balanced datasets provide a more robust and comparable basis for evaluating AGPC across the two benchmark years. This response can be found in the revised manuscript, as follows:

(In section 2.1, Lines 191–199)

"…Approximately 10,000 township-level census records were compiled for each of 2010 (Fig. 1-c, 10,834 records) and 2020 (Fig. 1-d, 10.336 records). As the original township-level census records were primarily available in tabular form without explicit spatial information, they were spatially matched to township administrative boundaries based on administrative names, with ambiguous records manually verified to account for duplicated names, administrative name changes, and boundary adjustments. The resulting validation datasets provide broad geographical coverage across mainland China, with all provinces represented. These township-level records, representing a finer spatial scale than the county-level training data, were used to independently evaluate the accuracy of the gridded population estimates, with a particular focus on the model's ability to capture intra-county spatial heterogeneity."

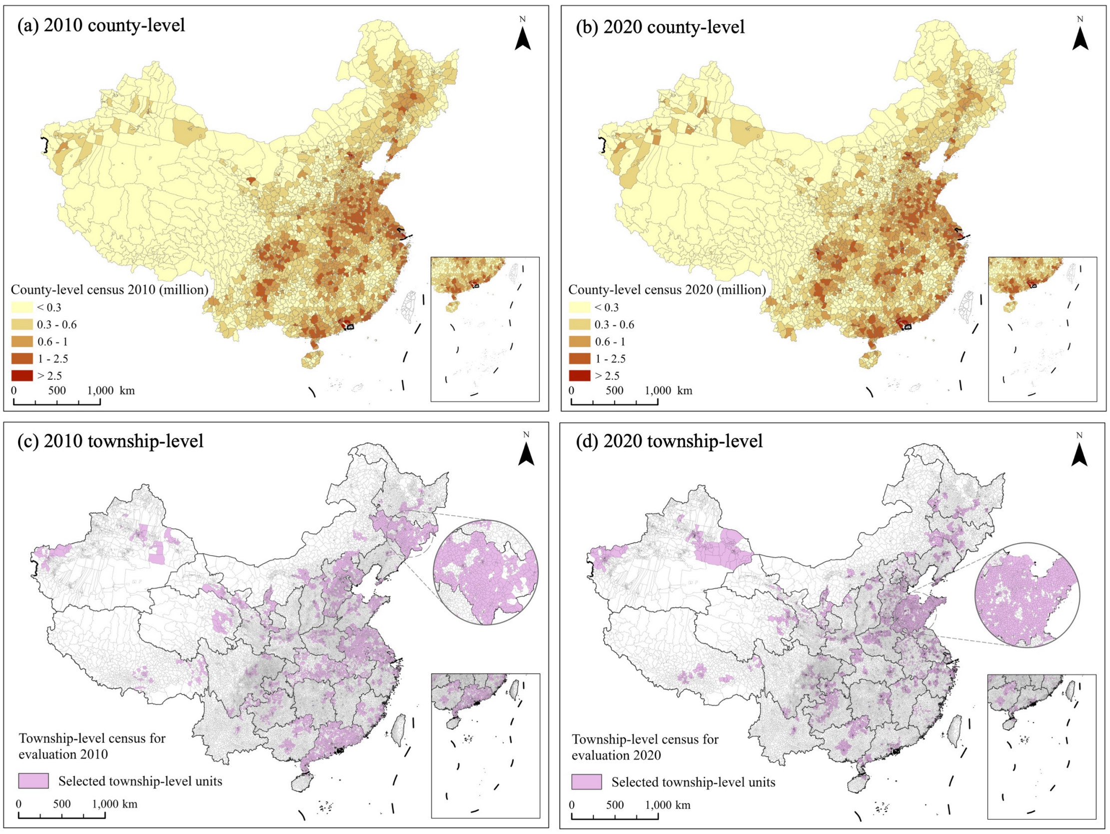

Fig. 1 Population census data from Sixth and Seventh National Population Census of China, including county-level population data for 2010 (a) and 2020 (b), and the spatial distribution of selected township-level census units for 2010 (c) and 2020 (d)

2. In Section 2.2.1 and the reference list, the GUS-3D dataset is cited as Liu et al. (2025) and Wu et al. (2026), with the latter marked as "Undergoing Review." Please clarify its current availability, public access protocol, or expected release timeline so that the reproducibility of AGPC can be properly assessed.

Response:

Thank you for raising this point regarding data availability and reproducibility. The manuscript describing the GUS-3D dataset (Wu et al., 2026) is currently under review and has therefore not yet been formally published. Nevertheless, to facilitate the reproducibility and assessment of AGPC, the GUS-3D data used in this study can currently be accessed through a private data link: https://figshare.com/s/9b7d1d40792daa9ab707. This link provides access to the dataset used for the reconstruction of the historical population distribution in AGPC. Upon formal publication of the GUS-3D manuscript and dataset, the current access information will be replaced or supplemented with the permanent public repository and corresponding DOI citation. The corresponding revisions can be found in the data description section and data availability section of the revised manuscript, as follows:

(In section 2.2.1, Lines 240–247)

"*...The 3D building variables were derived from the spatiotemporal GUS-3D dataset developed by our research team (Liu et al., 2025; Wu et al., 2026), which provides global coverage at a 500 m spatial resolution and includes building volume, building footprint ratio, and footprint-area–weighted mean building height for each grid cell. The dataset spans 1990–2020 at five-year intervals, enabling dynamic characterization of the spatiotemporal evolution of urban 3D structure. The manuscript describing the complete GUS-3D dataset is currently under review (Wu et al., 2026); however, the data used in this study are currently accessible through a private Figshare link provided in the Data Availability section. A permanent public repository and corresponding DOI citation will be provided upon formal publication of the GUS-3D dataset.*"

(In Data Availability, Lines 894–895)

"*...The GUS-3D dataset used to characterize the spatiotemporal evolution of urban 3D structure is currently accessible through a private link: https://figshare.com/s/9b7d1d40792daa9ab707.*"

3. Please clarify the temporal consistency of the POI dataset. Section 2.2.2 states that the POI data were collected from Gaode Map but does not specify their acquisition year. These POIs are subsequently treated as temporally invariant throughout the study. If they represent recent urban conditions (e.g., 2022–2023), their application to earlier years such as 1990 or 1995 may introduce temporal inconsistency, as also acknowledged in Section 4.1.3. Please specify the acquisition year of the POI dataset and further discuss the potential influence of this temporal mismatch on the early-period population estimates.

Response:

Thank you for pointing out this issue. The POI dataset used in this study was collected in 2020 through the Gaode Map API, which has now been explicitly stated in Section 2 of the revised manuscript. Because comparable historical POI datasets are unavailable, the 2020 POI density surfaces were treated as temporally invariant covariates throughout the reconstruction period. We acknowledge that this assumption may introduce temporal inconsistency, particularly for earlier years and areas that experienced substantial changes in urban functions. In our framework, this potential influence is partly mitigated by the use of time-varying built-up areas and 3D building characteristics, together with year-specific population constraints. Nevertheless, the lack of temporally consistent POI data remains a limitation in reconstructing historical changes in urban functions.

In the revised manuscript, we have therefore clarified the acquisition year and temporal treatment of the POI data in Section 2.2.2, and expanded the limitation discussion in Section 4.7, including potential improvements using multi-temporal POI archives or other temporally explicit functional data, as follows:

(In section 2.2.2, Lines 252–262)

"POIs serve as spatial proxies for urban functional facilities and socioeconomic activities and can partially reflect the intensity and type of human activities within cities. POI data were collected from Gaode Map (https://amap.com) through its API in 2020, comprising more than 34 million records. To better capture major urban activity patterns, the original 23 POI categories defined by Gaode Map were aggregated into eight functional types: residential, recreation, life services, education, commercial, office, medical, and open space. Details of the reclassification and record counts are provided in Table 2. Kernel density estimation (KDE) was applied to each POI category to generate density surfaces at a 500 m spatial resolution. A quartic kernel function (Silverman, 2018) with a uniform bandwidth of 2,000 m was used for all KDE calculations. Because comparable historical POI archives extending back to the 1990s and early 2000s are generally unavailable, the 2020 POI density surfaces were treated as temporally invariant covariates in the historical reconstruction. It should also be noted that the acquisition year of a POI record does not necessarily correspond to the establishment year of the associated facility."

(In section 4.7, Lines 872–886)

"…First, some explanatory variables, such as POI density and road network characteristics, were assumed to be temporally static due to the lack of long-term historical datasets. In particular, the POI data were obtained from a consistent 2020 dataset and applied throughout the reconstruction period, which may introduce temporal inconsistency for earlier years. Although POI density primarily represents urban functional intensity and spatial organization rather than the exact historical presence of individual facilities, the lack of temporally consistent historical POI data limits the ability to fully reconstruct past changes in urban functions. This mismatch may cause historical urban functions to be represented by their more recent spatial configuration, potentially affecting fine-scale population allocation in areas that experienced substantial functional transformation or rapid urban development. Nevertheless, historical population changes are also represented by time-varying built-up area and 3D building characteristics and constrained by year-specific statistical population totals, reducing the dependence of the reconstruction on temporally invariant POI information alone. Future studies could integrate multi-temporal POI archives, historical business registration data, or other temporally explicit functional land-use information to better capture the evolution of urban functions and further improve the temporal consistency of population reconstruction."

4. Section 3.1 states that the GAUD impervious surface dataset was combined with GUS-3D to delineate inhabited areas. However, the cited GAUD product only covers 1985–2015. Please specify how inhabited areas were derived for 2020 and clarify whether the 2015 GAUD layer was directly used as a proxy or whether an alternative dataset or updating strategy was adopted for the post-2015 period.

Response:

Thank you for pointing out this issue. The 2015 GAUD layer was not used as a proxy for 2020. Instead, the inhabited-region mask for 2020 was constructed using the corresponding 2020 impervious surface layer extracted from the 30 m Annual Global Land Cover (AGLC) dataset (Li et al., 2026), together with building information from GUS-3D. The AGLC dataset extends the temporal coverage of the impervious surface information originally provided by GAUD (Liu et al., 2020) to 2022, allowing year-specific impervious surface information to be used throughout the study period.

We have revised Section 3.1 to clarify this procedure and the relationship between the GAUD and AGLC datasets. The citation and Data Availability section have also been updated to include the AGLC dataset, which is publicly accessible through the National Tibetan Plateau Data Center (https://doi.org/10.11888/Terre.tpdc.302477). The corresponding revisions in the manuscript are as follows:

(In section 3.1, Lines 312–335):

"…Inhabited regions were primarily identified based on building presence, supplemented with impervious surface data to capture small and fragmented built-up areas. To capture the spatiotemporal dynamics of inhabited regions over the study period (1990–2020), building footprint ratio data from the GUS-3D dataset (Liu et al., 2025; Wu et al., 2026) were integrated with annual impervious surface information extracted from the 30 m Annual Global Land Cover (AGLC) dataset (Li et al., 2026). The AGLC dataset extends the temporal coverage of impervious surface information originally provided by the Global Annual Urban Dynamics (GAUD) dataset (Liu et al., 2020) to 2022, enabling year-specific impervious surface layers to be used throughout the study period. Because small and fragmented impervious features were partially filtered during the construction of GUS-3D, impervious surface fractions were recalculated at the 500 m grid level from the corresponding annual 30 m AGLC layers and merged with the GUS-3D building footprint ratios to construct a more complete inhabited-region mask for each year. In addition, fragmented settlements not captured by the AGLC dataset, such as small rural villages in mountainous areas (e.g., parts of Sichuan Province), were manually delineated using high-resolution satellite imagery from Google Maps. Building height and volume for these supplementary settlements were assigned based on visual interpretation and surrounding settlement patterns. This procedure further improves the spatial completeness and continuity of inhabited regions, particularly in sparsely populated and topographically complex areas."

(In Data Availability, Lines 932–933)

"The annual global land-cover dataset AGLC used in this study is publicly accessible through the National Tibetan Plateau Data Center at https://doi.org/10.11888/Terre.tpdc.302477."

The citation of AGLC dataset was added to the reference list accordingly:

"Li, B., Xu, X., Zhuang, H., Zhang, H., Che, Y., Zheng, Y., Cai, Y., Liu, X., and Li, X.: Annual global land cover mapping at 30 m resolution from 1985 to 2022 with high temporal consistency, Sci Bull, 71, 2726-2730, https://doi.org/10.1016/j.scib.2026.01.027, 2026."

5. Section 4.4 shows that the SHAP importance of CommercialPOI_den increases from approximately 28% in 2010 to more than 35% in 2020. Since the POI layer is assumed to be

temporally invariant, please clarify how this increase should be interpreted. Does the higher SHAP importance in 2020 reflects a stronger dependence of the model on commercial POIs? Please provide a clearer mechanistic explanation to avoid misinterpretation.

Response:

Thank you for this important comment. We agree that the increase in the SHAP importance of CommercialPOI_den from 2010 to 2020 requires careful interpretation, particularly because the same 2020 POI layer was used in both models. The higher SHAP importance of CommercialPOI_den in the 2020 model should not be directly interpreted as evidence that population distribution became more strongly dependent on commercial POIs over time. SHAP importance is model- and data-specific: it quantifies the average contribution of a feature to the predictions of a particular model given the distribution of all predictors and the response variable, rather than representing an intrinsic or temporally comparable importance of the feature itself. In our study, the 2010 and 2020 models were trained independently using population data and other predictors corresponding to their respective years. Although the CommercialPOI_den layer remains unchanged, the population distribution, 3D building characteristics, and distributions and interactions of the other predictors differ between the two years. These differences can change how the independently trained models use CommercialPOI_den and consequently result in different SHAP importance values. Therefore, the increase from approximately 28% in 2010 to more than 35% in 2020 indicates that CommercialPOI_den accounts for a larger share of the model attribution in the 2020 model under its specific predictor and population distributions, rather than demonstrating a temporal strengthening of the underlying population–commercial POI relationship. In particular, because the POI layer is temporally invariant, the difference in SHAP importance cannot by itself be used to infer changes in the actual influence of commercial activities on population distribution between 2010 and 2020.

We have revised Section 4.4 to clarify this interpretation and explicitly caution against interpreting differences in absolute SHAP importance as temporal changes in the underlying population–POI relationship. Instead, the cross-year comparison focuses on the general consistency of feature rankings, contribution patterns, and the directionality of SHAP values. The corresponding revision in Section 4.4 is as follows:

(In Lines 721–752)

"To improve the interpretability of the XGBoost-based dasymetric mapping model, the SHAP framework was employed to quantify the contribution of each feature and characterize its relationship with population density. The SHAP summary plots for 2010 and 2020 (Fig. 10) exhibit generally consistent feature rankings and contribution patterns despite the models being trained independently for the two years. Although differences in relative SHAP importance are observed, these values should not be interpreted as direct temporal changes in feature importance. SHAP values quantify the average contribution of each feature to predictions for a specific model and dataset, rather than representing an intrinsic importance of the feature itself. For example, the relative SHAP importance of commercial POI density increases from approximately 28 % in 2010 to more than 35 % in 2020. Because the same POI layer was used for both years, this increase does not necessarily indicate a stronger population dependence on

commercial POIs in 2020, but may reflect differences in population distribution, other temporally varying predictors and their interactions, and how the independently trained models utilize these predictors We therefore focus on the consistency of feature rankings, contribution patterns, and the directionality of SHAP values rather than directly comparing their absolute magnitudes between years.

Commercial POI density is ranked as the most influential feature in both models, with higher densities consistently associated with positive SHAP values, suggesting a strong relationship between commercial intensity and population concentration. This contribution may partly reflect the high spatial representativeness of commercial POIs, which are generally more comprehensively documented than residential population carriers. Moreover, POI density primarily serves as a proxy for the spatial organization of urban functions rather than a direct indicator of the historical evolution of individual facilities, whose establishment times are often unavailable. Its importance should therefore be interpreted as representing urban functional intensity rather than the temporal development of commercial activities.

In contrast, slope generally exhibits negative SHAP values, implying that steeper terrain constrains population concentration, although a limited number of high-slope samples show positive contributions. 3D building volume consistently ranks among the most influential predictors, with larger building volumes corresponding to positive SHAP contributions. Building-function variables further distinguish different urban land-use characteristics: higher residential building fractions generally increase predicted population density, whereas commercial and industrial building fractions tend to contribute negatively.

These results highlight the importance of incorporating vertical urban morphology and functional characteristics into fine-scale population mapping. The persistent contribution of 3D building volume indicates that population-carrying capacity in dense urban environments cannot be adequately represented by conventional 2D built-up indicators alone. Building-function variables also provide complementary information beyond POI-based indicators. The relatively limited explanatory power of residential POI density may result from the incomplete representation of residential communities in POI datasets despite their dominant role as population carriers. Incorporating building-based functional information therefore helps better characterize residential land use and reduce potential biases in population spatialization."

6. Figure 3 includes a "District-level Census Data" component under the Population Constraints module, yet district-level census data are rarely mentioned elsewhere in the manuscript. Please clarify whether district-level constraints were incorporated into the baseline population allocation.

Response:

Thank you for pointing out this inconsistency. The label "District-level Census Data" in Fig. 3 was not sufficiently precise and may have caused confusion. The population constraints were derived from county-level administrative units in China, which include counties, districts, county-level cities, and other county-level administrative divisions. To ensure consistency with the terminology used throughout the manuscript and to more accurately reflect the administrative level of the census data, we have revised the label in Fig. 3 from "District-level Census Data" to "County-level Census Data." The figure and the corresponding text have been updated accordingly, as follows.

(In section 2.1, Lines 173–177):

"……Here, county-level administrative units refer collectively to counties, urban districts, county-level cities, and other administrative divisions at the equivalent level in China. Population statistics for 2,845 such county-level units across mainland China were digitized and linked to corresponding county boundary data acquired from the Tianditu platform (www.tianditu.gov.cn)"

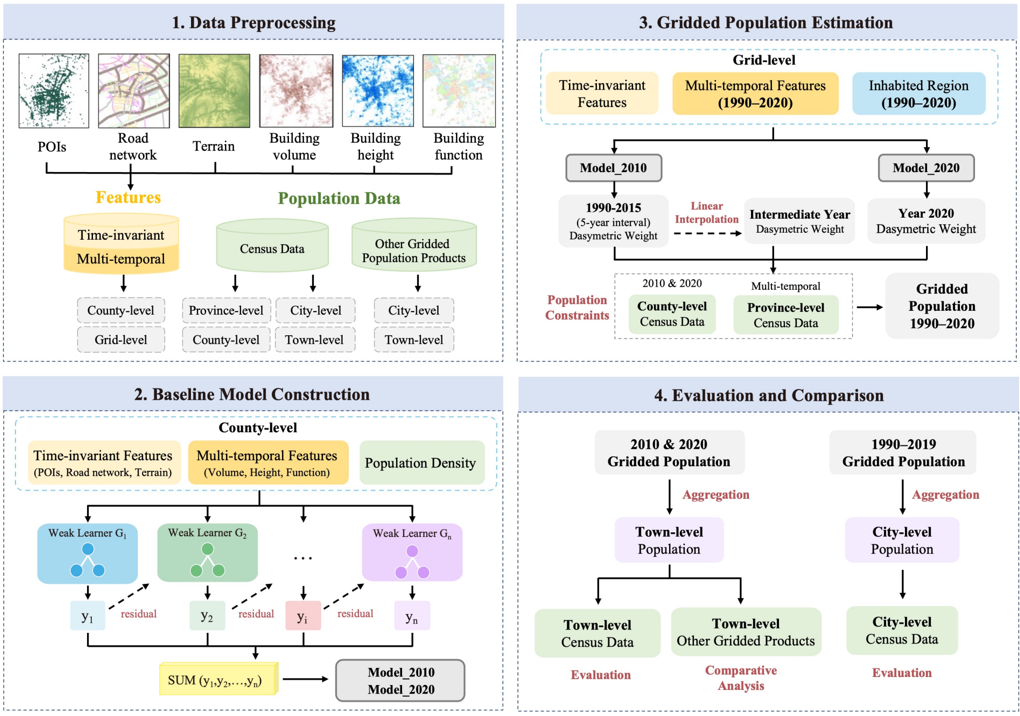

Fig. 3 The overall workflow of deriving the AGPC dataset

7. The word "Grided" appears repeatedly in the title, abstract, and throughout the manuscript (e.g., "500 m Grided Population"). The correct spelling is "Gridded". Please correct this throughout the manuscript.

Response:

Thank you for pointing out this spelling error. We have corrected "Grided" to "Gridded" in the title, abstract, and throughout the manuscript. Please refer to the revised manuscript for more details.

8. In the reference list, "China Statics Press" should be corrected to "China Statistics Press".

Response:

We have corrected "China Statics Press" to "China Statistics Press". Please refer to the updated reference list in the revised manuscript.