

1 **RESPONSE TO REVIEWER #1**

2 The manuscript proposes a high-resolution forest-specific mapping approach for predicting soil bulk
3 density and pH across China. It presents a substantial body of work and addresses a topic of interest,
4 which has the potential to contribute to the field. However, in my opinion, the current manuscript
5 requires major revision before it can be considered for publication.

6 Overall, the manuscript is informative and holds value but requires further refinement. The authors
7 are encouraged to more clearly emphasize the importance and novelty of their work, revise
8 redundant descriptions in Results while focusing on demonstrating statistical significance. With
9 careful revision, this manuscript has considerable potential to make a meaningful contribution to
10 the field.

11
12 **Response**

13 Dear reviewer #1,

14 We sincerely thank you for your insightful and comprehensive comments, which have been very
15 helpful in improving the quality and clarity of this manuscript. In response to your suggestions, we
16 have undertaken substantial revisions.

17 Specifically, we revised the Introduction to more clearly articulate the scientific motivation,
18 urgency, and novelty of forest-specific, high-resolution mapping of soil bulk density and pH across
19 China. We also carefully refined the Results section to reduce redundancy and to emphasize
20 quantitative interpretation supported by appropriate statistical evidence, avoiding subjective or over-
21 interpreted statements.

22 We believe that these revisions substantially enhance the rigor, transparency, and
23 interpretability of the manuscript. Our point-by-point responses to all comments are provided below,
24 and all corresponding revisions are marked in **blue** in the revised manuscript.

25
26 Best regards,

27 Jizhen Chen

28
29 **Major comment 1**

30 First of all, after reading the Introduction, I wasn't fully convinced of the necessity and urgency of
31 this study. The Introduction section begins with very basic background information on forest soil,
32 which is too general to establish a compelling rationale. The excessive introduction about
33 methodology doesn't effectively build a case for the study's significance, either. For instance, the
34 entire second paragraph is basically saying "a lot of people have done this", which may justify
35 methodological reliability but not why this work is needed. The fourth paragraph focuses on the
36 historical development of methodologies, which isn't the main goal of an Introduction. While
37 building a nationwide forest soil profile database is potentially valuable, the current Introduction
38 does not sufficiently highlight how this study advances beyond simply extracting forest-covered
39 data from existing maps.

40
41 **Response**

42 We thank the reviewer for this constructive and important comment. We agree that the original
43 Introduction did not sufficiently establish the necessity and urgency of this study.

44 In response, we have substantially revised the Introduction to adopt a more problem-driven

45 structure. General background information on forest soils and the historical development of digital
46 soil mapping methodologies has been condensed (*Lines 37-60*). The revised text now explicitly
47 emphasizes the limitations of existing national and global soil bulk density and pH products, which
48 are largely derived from mixed-ecosystem samples and therefore fail to capture the distinct spatial
49 heterogeneity and vertical structure of forest soils (*Lines 61-73*).

50 Importantly, we now clarify that simply extracting forest-covered pixels from existing soil
51 maps is insufficient. Instead, we highlight the need for forest-specific modeling frameworks that
52 explicitly account for ecosystem-specific processes and depth-dependent variability. This rationale
53 is clearly articulated in *Lines 79-80*, where we emphasize the ecological importance, spatial
54 complexity, and current lack of high-resolution forest soil BD and pH estimates across China.

56 **Major comment 2**

57 Some findings are presented without statistical validation and therefore unconvincing. For example,
58 L255 “BD prediction accuracy...peaking at intermediate depths (15–30 cm: MEC = 0.657) with
59 lower accuracy in surface layers (0–5cm: MEC = 0.598) and deep layers (60–100 cm: MEC =
60 0.656)”. Without testing for statistical significance, how can 0.656 represent “lower accuracy”
61 compared to 0.657? Similarly, statements such as “all predictions maintained negligible bias ($|ME|$
62 ≤ 0.019) across depth intervals” lack a defined threshold for “negligible”. Descriptions like
63 “Conversely, pH predictions demonstrated superior accuracy: CV maintained strong performance
64 across depths” appear subjective, without definition for “superior” or “strong”.

66 **Response**

67 Thank you for your important comment regarding the interpretation of model performance
68 metrics. In response, we have revised the manuscript to avoid over-interpretation of small numerical
69 differences in performance indicators and to remove subjective descriptors that were not supported
70 by formal statistical testing.

71 Specifically, statements comparing prediction accuracy across soil depths (e.g., “higher” or
72 “lower” accuracy) have been removed, as differences in MEC values such as 0.656 versus 0.657 are
73 not statistically meaningful. Similarly, qualitative terms such as “superior,” “strong,” and
74 “negligible” have been replaced with objective descriptions based on the reported ranges of MEC,
75 RMSE, and ME values.

76 The revised text now focuses on presenting model performance in a descriptive and transparent
77 manner, emphasizing the overall consistency between cross-validation and independent validation
78 results, as well as the absence of systematic bias indicated by ME values close to zero.

79 These revisions can be found in *Section 3.2 (Lines 275–288)*.

81 **Major comment 3**

82 Similarly, in the Result section, the authors keep emphasizing that their “patterns align with former
83 maps”, which further raises questions about the novelty and importance of this work.

85 **Response**

86 We thank the reviewer for this insightful comment regarding the interpretation of similarities
87 between our results and existing soil datasets. In response, we have revised the manuscript to avoid
88 overemphasizing pattern agreement with previous products in the Results section.

89 Specifically, statements such as “Macroscale patterns align with existing maps” have been
90 removed from the Results, as simple visual consistency alone does not sufficiently reflect the
91 novelty or contribution of this work. Instead, we have substantially revised *Section 4.1 (Lines 451–*
92 *511)* to provide a more detailed and quantitative comparison with existing datasets (CSDLv2,
93 ChinaSoilInfoGrids, and SoilGrids 2.0).

94 This revised discussion focuses on ecosystem-specific differences in both the vertical
95 distribution and magnitude of forest soil BD and pH, highlighting discrepancies that are not captured
96 by generalized soil products. In particular, we demonstrate that existing datasets fail to fully
97 represent the non-linear depth-dependent pattern of forest soil BD and systematically predict higher
98 BD values in deeper layers (60–100 cm), which may lead to overestimation of forest soil carbon
99 stocks. Similarly, our results indicate consistently lower forest soil pH values compared to existing
100 datasets, suggesting that ecosystem-specific acidification processes in forest soils are
101 underestimated in generalized products.

102 These revisions clarify that the contribution of this study lies not in reproducing existing spatial
103 patterns, but in providing forest-specific, high-resolution estimates that improve the representation
104 of soil properties and associated ecological processes.

106 **Major comment 4**

107 Many descriptions in the Results section are excessive or repetitive (e.g., L268–270, L274–279),
108 and some qualitative statements regarding spatial gradients (e.g., “BD values increase from the coast
109 inland” in L271) are unclear.

111 **Response**

112 We thank the reviewer for pointing out that parts of the Results section were overly descriptive
113 and repetitive, and that some qualitative statements regarding spatial gradients lacked clarity. In
114 response, we have substantially revised the Results section to improve conciseness and clarity, and
115 to shift from generalized qualitative descriptions toward a more quantitative and statistically
116 supported characterization of spatial patterns.

117 Specifically, repetitive regional descriptions have been removed, and ambiguous statements
118 such as “BD values increase from the coast inland” have been eliminated. Instead, we now quantify
119 spatial patterns using latitudinal and longitudinal gradients, together with regional statistical
120 summaries (boxplots), which provide a clearer and more objective representation of spatial
121 variability.

122 To support this revision, Figures 3 and 4 have been redesigned (now *Figures 4 and 5*) to
123 explicitly illustrate depth-specific latitudinal and longitudinal trends, as well as regional differences
124 in BD and pH. The revised Results section (*Section 3.3, Lines 290–345*) now presents the main
125 spatial features more concisely, while preserving the key information needed to interpret large-scale
126 patterns.

128 **Major comment 5**

129 Why is FRFS introduced in the Introduction section but QRF in the Method?

131 **Response**

132 We thank the reviewer for this comment regarding the structure and consistency of the

133 methodological description. We agree that the initial presentation may have given the impression
134 that FRFS and QRF were introduced at different conceptual levels.

135 In response, we have revised the Introduction to explicitly present FRFS and QRF as
136 complementary components of a unified modeling framework. Specifically, FRFS is introduced as
137 a feature selection strategy designed to reduce dimensionality and improve model parsimony and
138 interpretability, while QRF is described as the core predictive algorithm used to model soil BD and
139 pH and to quantify prediction uncertainty.

140 This dual-focused strategy is now clearly outlined in the Introduction (*Lines 46–60*). The
141 Methods section then follows this conceptual structure, first detailing the covariate selection
142 procedure based on FRFS (*Section 2.2.1*), and subsequently describing the implementation of the
143 QRF model as the primary predictive tool (*Section 2.2.2*). This revision ensures consistency
144 between the Introduction and Methods and clarifies the distinct but integrated roles of FRFS and
145 QRF within the overall DSM framework.

147 **Major comment 6**

148 Table 1 may be presented more clearly as a figure, and currently has a confusing caption.

150 **Response**

151 We thank the reviewer for this helpful suggestion regarding the presentation of Table 1. In
152 response, we have revised the manuscript to improve the clarity and interpretability of this
153 information. Specifically, Table 1 has been converted into a violin plot (now *Fig. 3*), which more
154 effectively illustrates the distribution of soil samples across depth intervals and highlights
155 differences between layers.

156 In addition, statistical tests have been applied to assess the significance of differences among
157 soil depths, and the corresponding results are now explicitly shown in the figure. The original table
158 has been moved to the Supplementary Information (*Table S5*) for reference. The Results section has
159 been updated accordingly to reflect the revised figure and the additional statistical information
160 (*Section 3.1, Lines 248–274*). These changes improve the clarity of data presentation and provide a
161 more informative summary of the sampling structure.

163 **Major comment 7**

164 Figure 6 might benefit from an overall analysis across depths, and consider adding relationships
165 between BD and MAP (or other key covariates) in supplementary materials.

167 **Response**

168 We thank the reviewer for this constructive suggestion regarding the analysis of Figure 6 and
169 the relationships between soil properties and key covariates. We agree that an overall comparison
170 across soil depths, together with a clearer interpretation of the relationships between BD/pH and
171 major environmental drivers, would substantially strengthen the manuscript. In response, we have
172 extended the methodological framework by introducing SHAP (SHapley Additive exPlanations)
173 analysis (*Section 2.3, Lines 222–229*).

174 This addition addresses a key limitation of the relative variable importance measures
175 previously derived from the QRF models, which reflect covariate importance only in a relative sense
176 within each depth-specific model and depend on the selected feature set.

177 Because FRFS yields different covariate subsets for different soil layers, these relative
178 importance values are not directly comparable across depths. SHAP provides independent, additive
179 contribution scores for each predictor, enabling consistent cross-depth comparison and allowing
180 both the magnitude and direction of covariate effects to be quantitatively interpreted. Based on this
181 approach, Figures 6 and 7 have been revised and are now presented as *Figures 7 and 8*, illustrating
182 depth-consistent importance patterns and the relationships between BD/pH and key covariates.

183 Accordingly, the previous Results subsection on variable importance has been fully revised
184 and replaced by *Section 3.5 (Lines 377–453)*, which now presents the SHAP-based analysis and
185 interpretation in place of the original QRF relative importance results.

186
187 **Major comment 8**

188 L85 & 91, QRF should be explained upon its first mention.

189
190 **Response**

191 We thank the reviewer for pointing out this issue. In response, we have revised the manuscript
192 to ensure that Quantile Regression Forest (QRF) is fully explained at its first mention (*Lines 51*). In
193 addition, we conducted a systematic check of abbreviations throughout the manuscript to ensure
194 consistent definition and usage upon first appearance.

195
196 **Major comment 9**

197 Abbreviations (including BD, SD and the abbreviations of models) in Tables and Figures should be
198 clearly defined in their captions to make them self-explanatory.

199
200 **Response**

201 We thank the reviewer for this helpful comment. In response, we have revised all figure and
202 table captions to ensure that abbreviations (including BD, SD, and model abbreviations) are clearly
203 defined upon first appearance, making the tables and figures self-explanatory. This change has been
204 applied consistently throughout the manuscript.

205
206
207 **Major comment 10**

208 L111 is redundant with L108.

209
210 **Response**

211 We thank the reviewer for carefully identifying this redundancy. In response, we have revised
212 the manuscript to improve clarity and conciseness by removing the sentence.

213
214 **Major comment 11**

215 L251, rephrase “conversely”.

216
217 **Response**

218 We thank the reviewer for this helpful suggestion regarding wording. In response, we have
219 revised the relevant sentence and removed the use of “conversely,” which was no longer appropriate
220 given the revised structure and interpretation of the Results section. Following the substantial

221 revision of *Section 3.2*, the description of pH model performance has been rewritten to avoid
222 subjective or contrastive wording and to present the results in a more neutral and consistent manner
223 based on the reported performance metrics.
224

225 **RESPONSE TO REVIEWER #2**

226 The authors address a significant topic by mapping key soil properties across China's forests using
227 a comprehensive dataset. The resulting high-resolution products have the potential to be a valuable
228 resource for the scientific community. However, several methodological and descriptive aspects
229 require substantial improvement to ensure the reliability and reproducibility of the findings.

230

231 **Response**

232 Dear reviewer #2,

233 We sincerely thank you for your insightful and comprehensive comments, which have been very
234 helpful in improving the rigor, transparency, and reproducibility of this manuscript. Following your
235 suggestions, we have undertaken substantial revisions.

236 Specifically, we strengthened the methodological description and expanded the Discussion to
237 more explicitly address sources of uncertainty (e.g., covariate resolution harmonization) and model
238 interpretability across soil depths. We also added new supplementary materials to improve
239 transparency, including covariate maps, sample distribution summaries, and additional quantitative
240 comparisons with existing BD and pH products.

241 Our point-by-point responses are provided below. All corresponding revisions are marked in
242 **blue** in the revised manuscript.

243

244 Best regards,

245 Jizhen Chen

246

247 **General comments:**

248 **General comments 1**

249 Resampling covariates with diverse native resolutions to a 90-m grid introduces significant
250 uncertainty, particularly for inputs derived from coarser scales. This issue warrants a detailed
251 discussion and quantification to assess the reliability of the final high-resolution maps.

252

253 **Response**

254 We thank the reviewer for raising this important point regarding uncertainty introduced by
255 harmonizing environmental covariates with heterogeneous native resolutions to a 90-m grid.

256 In this study, all covariate layers were projected to a common coordinate system and resampled
257 to 90 m using bilinear interpolation, following common practice in national-scale digital soil
258 mapping and consistent with established products (Poggio et al., 2021; Liu et al., 2022; Shi et al.,
259 2025). We acknowledge that resampling, particularly from coarser-resolution inputs, may introduce
260 scale-related uncertainty and smoothing effects, and that the effective spatial resolution of the final
261 predictions is constrained by the coarsest covariates.

262 In response, we have expanded the Discussion to explicitly address this limitation and its
263 implications for map interpretation (*Section 4.3, Lines 538–551*). As noted there, isolating and
264 quantifying uncertainty attributable solely to covariate resampling is methodologically challenging
265 at national scales and is rarely reported in existing large-area DSM studies. Therefore, the resulting
266 maps should be interpreted as conditional estimates that represent the most probable spatial patterns
267 given the available covariates and their effective spatial support, rather than as direct representations
268 of fine-scale soil heterogeneity at 90 m.

269 Reference:

270 Liu, F., Wu, H., Zhao, Y., Li, D., Yang, J., Song, X., Shi, Z., Zhu, A., and Zhang, G.: Mapping high
271 resolution national soil information grids of China, *Sci. Bull.*, 67, 328–340,
272 <https://doi.org/10.1016/j.scib.2021.10.013>, 2022.

273 Poggio, L., De Sousa, L. M., Batjes, N. H., Heuvelink, G. B. M., Kempen, B., Ribeiro, E., and
274 Rossiter, D.: SoilGrids 2.0: producing soil information for the globe with quantified spatial
275 uncertainty, *Soil*, 7, 217–240, <https://doi.org/10.5194/soil-7-217-2021>, 2021.

276 Shi, G., Sun, W., Shangguan, W., Wei, Z., Yuan, H., Li, L., Sun, X., Zhang, Y., Liang, H., Li, D.,
277 Huang, F., Li, Q., and Dai, Y.: A China dataset of soil properties for land surface modelling
278 (version 2, CSDLv2), *Earth Syst. Sci. Data*, 17, 517–543, [https://doi.org/10.5194/essd-17-517-](https://doi.org/10.5194/essd-17-517-2025)
279 [2025](https://doi.org/10.5194/essd-17-517-2025), 2025.

280

281

282

283 **General comments 2**

284 The manuscript provides insufficient discussion on the variable importance for BD and pH across
285 different soil layers. The underlying reasons for these variations require further elaboration.

286

287 **Response**

288 We thank the reviewer for this comment. We agree that the original manuscript did not sufficiently
289 elaborate on depth-dependent variations in variable importance and the underlying reasons for these
290 patterns.

291 In response, we substantially revised the interpretability analysis by introducing SHAP (SHapley
292 Additive exPlanations), which enables consistent, depth-comparable quantification of both the
293 magnitude and direction of covariate effects. This revision replaces the previous discussion based
294 on QRF-derived relative variable importance, which primarily supports within-model comparisons
295 and is not directly comparable across depths when the selected feature sets differ.

296 The revised *Section 3.5* now provides a depth-explicit interpretation of key drivers for both BD and
297 pH, and the corresponding methodological update is described in *Section 2.3*. The updated results
298 are presented in *Figures 7 and 8*.

299

300 **General comments 3**

301 Natural and planted forests possess distinct driving mechanisms. Developing separate models for
302 each forest type is advisable to accurately capture these specific variations.

303

304 **Response**

305 We thank the reviewer for this insightful suggestion. We agree that natural and planted forests often
306 differ in management history, stand structure, and species composition, which may influence soil
307 processes and their responses to environmental drivers.

308 However, in our compiled forest soil profile database, forest origin (natural vs. planted) is not
309 consistently documented in the original metadata, which prevents a reliable partition of the full
310 dataset by forest type. To evaluate the potential impact of forest origin, we conducted an additional
311 targeted test using an external forest-type dataset to assign forest origin at sampling locations (Cheng
312 et al., 2024), and performed separate modeling for the 60–100 cm layer as a representative case.

313 The results showed a pronounced decline in predictive performance when models were
314 separated by forest origin, primarily due to substantially reduced sample sizes at this depth (e.g.,
315 BD: planted MEC = 0.462, RMSE = 0.598; natural MEC = 0.302, RMSE = 0.632; pH: planted MEC
316 = 0.503, RMSE = 0.482; natural MEC = 0.404, RMSE = 0.402).

317 Moreover, evidence from the literature suggests that, at regional to national scales, the
318 dominant controls on soil bulk density and pH are largely consistent across natural and planted
319 forests. Broad environmental gradients—particularly climate, parent material, and long-term soil
320 development processes—have been shown to exert primary control on soil physical and chemical
321 properties, while forest origin mainly modulates response magnitude through differences in stand
322 structure or management intensity rather than altering the fundamental driver mechanisms
323 (Luysaert et al., 2008; Pretzsch et al., 2014).

324 Accordingly, separating models by forest origin at the national scale is unlikely to substantially
325 improve predictive performance for BD and pH, while potentially increasing uncertainty due to
326 reduced sample sizes. Given these considerations, together with the national-scale objective of this
327 study, we retained a unified modeling framework to ensure robust, spatially consistent predictions
328 across China's forest domain.

329 Reference:

330 Cheng, K., Yang, H., Guan, H., Ren, Y., Chen, Y., Chen, M., Yang, Z., Lin, D., Liu, W., Xu, J., Xu,
331 G., Ma, K., and Guo, Q.: Unveiling China's natural and planted forest spatial-temporal
332 dynamics from 1990 to 2020, *ISPRS J. Photogramm. Remote Sens.*, 209, 37–50,
333 <https://doi.org/10.1016/j.isprsjprs.2023.11.003>, 2024.

334 Luysaert, S., Schulze, E. D., Börner, A., Knohl, A., Hessenmöller, D., Law, B. E., Ciais, P., and
335 Grace, J.: Old-growth forests as global carbon sinks, *Nature*, 455, 213–215,
336 <https://doi.org/10.1038/nature07276>, 2008.

337 Pretzsch, H., Biber, P., Schütze, G., Uhl, E., and Rötzer, T.: Forest stand growth dynamics in Central
338 Europe have accelerated since 1870, *Nat. Commun.*, 5, 4967,
339 <https://doi.org/10.1038/ncomms5967>, 2014.

341 **General comments 4**

342 Providing spatial distribution maps for every covariate listed in Table S1 is advisable. The figures
343 need to clearly display the value ranges for continuous variables and the distinct spatial patterns for
344 each category within categorical variables. Specifying the number of sample points in both the
345 training and validation sets for each categorical variable is recommended.

347 **Response**

348 We thank the reviewer for this constructive suggestion regarding the presentation of environmental
349 covariates and sample stratification.

350 In response, we have added spatial distribution maps for all environmental covariates retained
351 after FRFS selection, which are now provided in the *Supplementary Information (Fig. S1)*. These
352 figures display the value ranges of continuous variables and the spatial patterns of categorical
353 variables, thereby improving transparency and interpretability of the covariate inputs.

354 In addition, we have quantified the number of soil sampling plots associated with each category
355 of the categorical covariates separately for the training and independent validation datasets, using
356

357 the surface soil layer (0–5 cm) as a representative overview. These statistics are summarized in
358 *Tables S2–S4 in the Supplementary Information.*

359 Together, these additions enhance the reproducibility of the study and provide a clearer basis
360 for assessing the robustness of the modeling framework.

361

362

363 **General comments 5**

364 Forest age represents a critical covariate. Incorporating this variable into the analysis is advisable to
365 improve model performance.

366

367 **Response**

368 We thank the reviewer for this important suggestion regarding the inclusion of forest age as a
369 covariate. We fully agree that forest age is a critical variable influencing forest structure,
370 productivity, and soil processes, and therefore has the potential to improve model performance when
371 reliable and spatially consistent data are available.

372 During the initial covariate selection stage, we carefully evaluated commonly used forest age
373 datasets for China (Cheng et al., 2024). However, currently available forest age maps are derived
374 products and still exhibit substantial uncertainty at local scales, particularly in cold-temperate
375 coniferous forests, where reported model performance remains relatively low ($R^2 = 0.47$).

376 More importantly, existing forest age products are developed based on forest domain
377 definitions and source data that differ from those adopted in this study. As a result, their spatial
378 coverage is not fully consistent with the forest domain considered here, leading to notable spatial
379 mismatches and discontinuities when the two datasets are overlaid. Integrating forest age products
380 into a national, wall-to-wall digital soil mapping framework would therefore introduce systematic
381 spatial bias and compromise prediction consistency. Given these considerations, we did not include
382 forest age as a covariate in the present study.

383 We note that digital soil mapping under conditions of incomplete or spatially inconsistent
384 covariate coverage represents a methodological challenge rather than a purely data-related issue.
385 This challenge is particularly pronounced in national-scale forest soil mapping studies in China,
386 where heterogeneity in data sources and definitions is common. Recent methodological advances,
387 such as the iPSM-based framework proposed by Fan et al. (2020), have demonstrated that covariate
388 incompleteness can be addressed by explicitly accounting for missing variables and their associated
389 uncertainty. Accordingly, we acknowledge this limitation in the Discussion (Show in **Section 4.3.**
390 **Line 553-568**) and consider the further development and application of such approaches to be an
391 important direction for improving future forest soil mapping frameworks.

392

393 **Reference:**

394 Cheng, K., Yang, H., Guan, H., Ren, Y., Chen, Y., Chen, M., Yang, Z., Lin, D., Liu, W., Xu, J., Xu,
395 G., Ma, K., and Guo, Q.: Unveiling china's natural and planted forest spatial-temporal
396 dynamics from 1990 to 2020, *ISPRS J. Photogramm. Remote Sens.*, 209, 37–50,
397 <https://doi.org/10.1016/j.isprsjprs.2024.01.024>, 2024.

398 Fan, N. Q., Zhu, A.-X., Qin, C.-Z., and Liang, P.: Digital soil mapping over large areas
399 with invalid environmental covariate data, *ISPRS Int. J. Geo-Inf.*, 9, 102, <https://doi.org/10.3390/ijgi9020102>, 2020.

400

401

402 **General comments 6**

403 Providing the original data is necessary to facilitate the reproducibility of the study by other
404 researchers.

405

406 **Response**

407 We thank the reviewer for emphasizing the importance of reproducibility. We would like to clarify
408 that the original forest soil profile data used in this study are subject to data confidentiality
409 agreements and access restrictions imposed by the institutions responsible for the national forest
410 soil survey in China, and therefore cannot be publicly released.

411 Nevertheless, we have made all data that can be legally shared openly available, and access to
412 restricted data may be requested by contacting the corresponding author, subject to the relevant data-
413 sharing policies. To ensure the highest possible level of reproducibility under these constraints, we
414 provide comprehensive documentation of the data processing workflow, quality control and
415 harmonization procedures, environmental covariates, modeling framework, and validation strategy.

416 In addition, all derived products generated in this study, including gridded maps of soil bulk
417 density and pH and their associated uncertainty estimates, are openly available through a public
418 repository.

419 We believe that this level of transparency allows other researchers to fully reproduce the
420 methodology and apply it to independent datasets, while respecting the access restrictions governing
421 the original observations.

422

423 **General comments 7**

424 Comparing the current results with existing soil BD and pH products is recommended. The
425 manuscript needs to clarify specific improvements and explain the reasons for these advancements.

426

427 **Response**

428 We thank the reviewer for this suggestion. Following it, we conducted additional quantitative
429 comparisons between our forest-specific predictions and existing national and global datasets,
430 including CSDLv2, ChinaSoilInfoGrids, and SoilGrids 2.0. These analyses explicitly quantify
431 differences in predicted BD and pH across depths and regions, and clarify where and why forest-
432 specific mapping provides added value.

433 The statistical comparisons are summarized in *Figures S5 and S6 at Supplementary*
434 *Information*, and spatial contrasts are shown in *Figures 9 and 10*. The results indicate that our
435 forest-specific maps better capture depth-dependent patterns and ecosystem-specific magnitude
436 differences that are not well represented by generalized products. The underlying reasons for these
437 improvements are discussed in *Section 4.1*, including the use of forest-only observations, forest-
438 specific covariate–response relationships, and depth-resolved modeling.

439

440 **General comments 8**

441 Presenting the spatial distribution of sample points for both the training and validation sets is
442 necessary. The manuscript should also address whether these distributions are spatially balanced.

443

444 **Response**

445 We thank the reviewer for this suggestion regarding the spatial distribution and balance of the
446 training and validation samples.

447 In response, we have provided a new figure (*Fig. S2*) that visualizes the spatial distributions of the
448 training and independent validation datasets for BD and pH across all soil layers, supporting the
449 spatial balance claim.

450 Additionally, we have added a statement in the Methods section (*Lines 97–98*) to clarify that the
451 spatial distributions of the training and validation subsets were explicitly examined to ensure that
452 both subsets maintain spatial balance across regions and soil depths.

453

454 **General comments 9**

455 A more detailed description of the raw data is necessary. The manuscript should specify the sample
456 sizes and spatial distributions across different temporal periods, soil types, and forest types.

457

458 **Response**

459 We thank the reviewer for requesting a more detailed description of the raw data. In response, we
460 revised the data description to explicitly report sample sizes and their distributions across forest
461 types, soil types, and temporal periods. Detailed counts by forest type, soil type, and related
462 categories are provided in *Tables S2–S5* (referenced in the main text; *Lines 112*).

463 In addition, we clarified the temporal coverage of the dataset by reporting the number of soil
464 profiles collected in each survey year from 2018 to 2023 (*Lines 119–120*). These revisions provide
465 a more transparent overview of the dataset composition in terms of time, forest types, soil types,
466 and spatial coverage.

467

468 **Minor comments:**

469 **Minor comments 1**

470 Lines 1-3: The general phrase "Soil properties" creates redundancy with the specific variables "Bulk
471 Density and pH," necessitating a more concise revision such as "High-Resolution, Multi-Depth
472 Mapping of Soil Bulk Density and pH in China's Forests Using Machine Learning".

473

474 **Response**

475 We thank the reviewer for this helpful suggestion. We revised the title by removing the general
476 phrase "Soil properties" and focusing on the specific variables examined in this study. The revised
477 title is: "High-Resolution, Multi-Depth Mapping of Soil Bulk Density and pH in China's Forests
478 Using Machine Learning." This change is reflected in *Lines 1–3*.

479

480 **Minor comments 2**

481 Lines 20-21: The claim of being 'first' is inaccurate due to the existence of prior 90-m products, so
482 the text should be revised to focus on the specific contribution to forest ecosystems instead. and

483 Lines 429-430: The use of "first" is an absolute claim that is prone to dispute.

484

485 **Response**

486 We thank the reviewer for pointing out that the term "first" is an absolute claim and may be
487 disputable given the existence of prior 90-m soil products. We agree and have revised the text to
488 avoid absolute wording, shifting the emphasis to the specific contribution of this study—namely,

489 forest-specific mapping and depth-resolved estimates of BD and pH. The relevant revisions have
490 been made in *Lines 18* and *Lines 598–599*.

491

492 **Minor comments 3**

493 Lines 73-74: The phrase "in heterogeneous" is grammatically incorrect.

494

495 **Response**

496 We thank the reviewer for identifying this grammatical issue. The incorrect phrase has been removed
497 during the revision of the Introduction, and the corresponding sentence has been rewritten for clarity.

498

499 **Minor comments 4**

500 Lines 109-111: This sentence is redundant and should be deleted.

501

502 **Response**

503 We thank the reviewer for pointing out this redundancy. We deleted the redundant sentence as
504 suggested and revised the surrounding text accordingly.

505

506 **Minor comments 5**

507 Lines 111-112: Specify the quality control and data harmonization methods.

508

509 **Response**

510 We thank the reviewer for this suggestion. We have added a clearer description of the quality control
511 and data harmonization procedures (*Lines 115–116*), including unit standardization, harmonization
512 of depth intervals, and reconciliation of metadata across survey sources.

513

514 **Minor comments 6**

515 Lines 139-155: The 41 environmental covariates lack necessary citations, and the sources or the
516 data itself should be made accessible to readers to ensure reproducibility.

517

518 **Response**

519 We thank the reviewer for this comment. In response, we added detailed citations and source
520 information for all 41 environmental covariates in *Supplementary Table S1* to improve transparency
521 and reproducibility.

522

523

524 **Minor comments 7**

525 Lines 159-160: The number of standardized soil layers should be corrected from four to five.

526

527 **Response**

528 We thank the reviewer for pointing out this discrepancy. We corrected the number of standardized
529 soil layers from four to five (*Line 141*).

530

531 **Minor comments 8**

532 Lines 222-223: *q50* denotes the median prediction.

533

534 **Response**

535 We thank the reviewer for this correction. We revised the manuscript to clarify that $q_{0.50}$ denotes
536 the median prediction and corrected the notation accordingly (*Line 185*). We confirm that this was
537 a notation issue only; all analyses were conducted using the median ($q_{0.50}$) prediction, and the
538 correction does not affect the results or conclusions.