Dear Reviewer,

Following your suggestion, we have compared the performance of all soil properties of interest in this study by evaluating both the mean predictions using Random Forest (RF) and median predictions using Quantile Regression Forest (QRF). The table below presents a 10-fold cross-validation performance comparison for each method—mean prediction by RF and median prediction by QRF—under the 'All Data', 'High Values' and 'Low Values' conditions. Specifically, the 'All Data' condition evaluates performance across the entire training set, while 'High Values' and 'Low Values' conditions focus on prediction accuracy for the top 10% highest and bottom 10% lowest values, respectively.

The 'Prediction method' column documents the models constructed to generate the final nationwide 90-meter resolution predictions for each soil property. In selecting the models, we considered their performance in mean predictions and their ability to capture extreme values (both maximum and minimum). Additionally, we observed a consistent trend in model performance across different depth layers for each soil property (i.e., for all layers of a specific property, either RF's mean predictions or QRF's median predictions consistently outperformed the other, as shown in the figure below). Consequently, for each specific soil property, only one optimal prediction model was ultimately selected to develop the 90-meter resolution soil maps. Therefore, we have presented only the performance metrics for the 0-5 cm surface depth in the table to streamline the comparison.

Thank you once again for your insightful comments, which helped us improve the clarity of our work.

We have added the following to the manuscript:
Modification:

Although the performance differences between mean predictions using RF and median predictions using QRF are minimal, their ability to capture extreme values was considered. In this study, we evaluated the performance of RF and QRF models by not only the overall statistical metrics but also their capacity to predict extreme values (i.e. both high and low values), to determine the most suitable model for generating national gridded soil maps of various soil properties at a 90-meter resolution. As shown in Table S7, soil properties such as soil pH, silt, clay, TP, Red (R) of wet soil color, Blue (B) of wet soil color, Red (R) of dry soil color, and Blue (B) of dry soil color were modeled using median predictions from QRF, as this approach better captured extreme values. Similarly, the study by Helfenstein et al., (2024) also assessed mean predictions by RF and median predictions by QRF, highlighting that for certain soil properties, median predictions are more appropriate than mean predictions. For most other soil properties in this study—such as sand, BD, OC, gravel, AN, TN, CEC, porosity, TK, AK, AP, Green (G) of wet soil color, and Green (G) of dry soil color—mean predictions from RF were used to generate the 90-meter resolution soil maps. The better model was consistent across different depths for

the same soil property; thus, Table S7 only presents the performance comparison of mean and median predictions for the surface layer (0-5 cm depth interval) and either the mean or the median is used for the mapping of a soil property for all depths.

When developing the 90-meter resolution soil maps in this study, either mean or median predictions were selected for storage efficiency. However, for lower-resolution maps provided at 1 km and 10 km, in addition to mean and median predictions, we also included prediction maps for the 0.05 and 0.95 quantiles. These additional maps are helpful for illustrating data uncertainty.

Helfenstein, A., Mulder, V. L., Teuling, K., Walvoort, D. J. J., Heuvelink, G. B. M., Wageningen, A., and Wageningen, R.: BIS-4D: mapping soil properties and their uncertainties at 25 m resolution in the Netherlands, 2024.
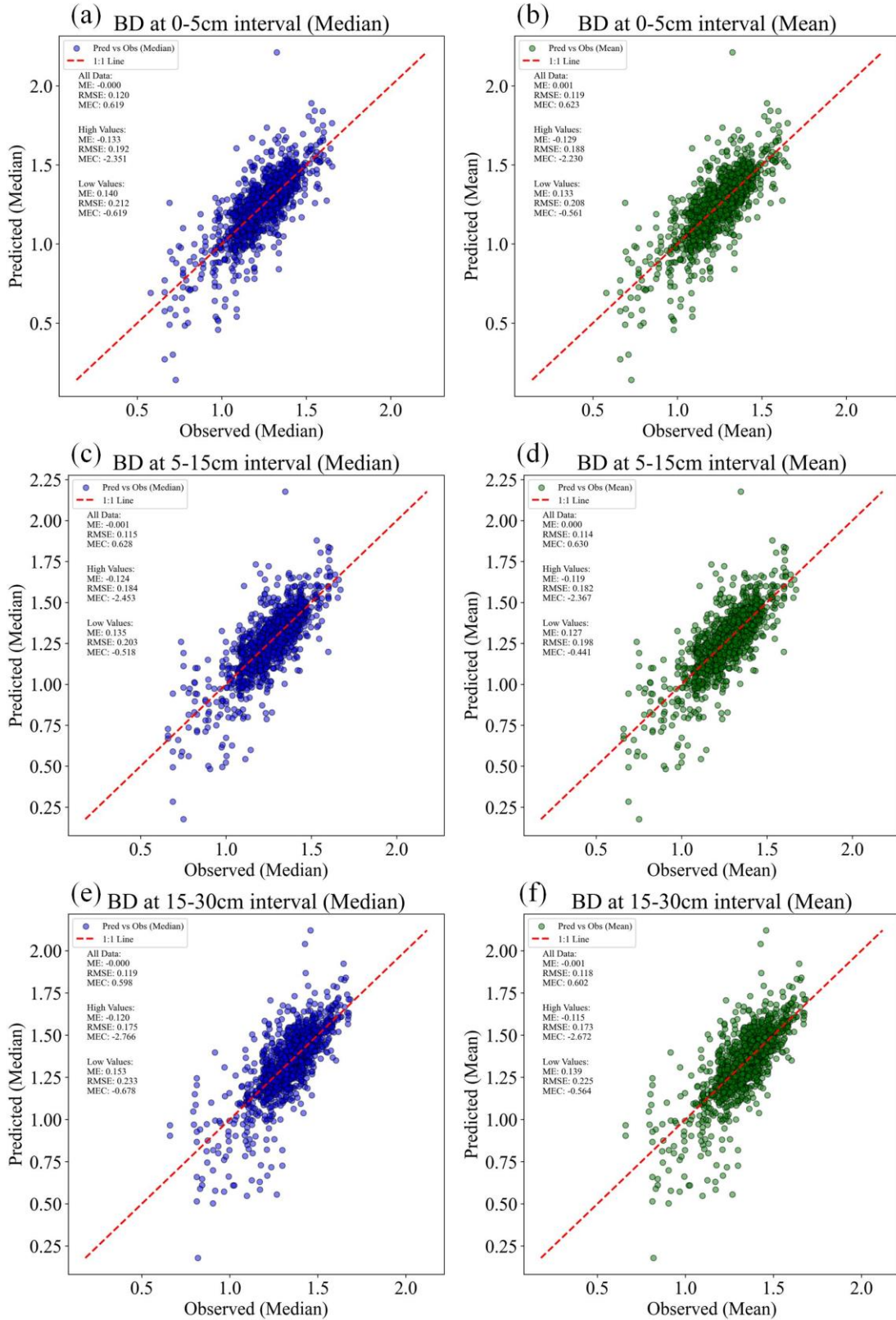
The information added in the supplementary material is presented in the table below:

Table S1. Comparison of predictive performance for mean predictions using random forest model and median predictions using quantile regression Forest model across different soil properties under 'All Data,' 'High Values,' and 'Low Values' conditions based on 10-fold cross-validation. The 'All Data' condition evaluates performance on the full training set, while 'High Values' and 'Low Values' assess prediction accuracy for extreme high and low values within the training set, respectively. The 'Prediction method' column documents the models constructed for generating final national-scale predictions at a 90-meter resolution for various soil properties.

property	Statistic Validation	All Data			High Values			Low Values			Prediction method
		MEC	RMSE	ME	MEC	RMSE	ME	MEC	RMSE	ME	
pH	Mean	**0.693**	**0.706**	**0.001**	-2.923	0.786	-0.564	-7.196	1.023	0.830	**Median**
	Median	0.690	0.709	-0.012	**-2.871**	**0.781**	**-0.557**	**-5.958**	**0.943**	**0.730**	**(QRF)**
sand	Mean	**0.670**	**12.161**	**0.056**	**-8.000**	**22.178**	**-16.260**	-9.543	12.961	8.400	Mean
	Median	0.667	12.231	-0.734	-8.612	22.919	-16.560	**-9.070**	**12.021**	**7.299**	(RF)
silt	Mean	0.615	9.825	0.023	-4.324	15.014	-10.967	-8.659	15.652	11.526	**Median**
	Median	0.614	9.840	0.003	**-4.240**	**14.895**	**-10.789**	**-8.838**	**15.796**	**11.139**	**(QRF)**
clay	Mean	**0.629**	**6.749**	**0.019**	-1.328	12.197	-8.221	-23.577	8.543	6.279	**Median**
	Median	0.626	6.771	0.018	**-1.281**	**12.071**	**-7.919**	**-23.416**	**8.515**	**6.088**	**(QRF)**
BD	Mean	**0.623**	**0.119**	**0.001**	**-2.230**	**0.188**	**-0.129**	**-0.561**	**0.208**	**0.133**	Mean
	Median	0.619	0.120	-0.000	-2.351	0.192	-0.133	-0.619	0.212	0.140	(RF)
OC	Mean	**0.570**	**2.043**	**0.028**	**0.089**	**5.382**	**-2.647**	-98.229	1.056	0.559	Mean
	Median	0.556	2.075	-0.225	-0.071	5.836	-3.455	**-69.297**	**0.889**	**0.464**	(RF)
gravel	Mean	**0.494**	**13.010**	**0.066**	**-5.133**	**24.486**	**-19.554**	-150.920	10.572	8.463	Mean

	Median	0.483	13.152	-1.542	-5.427	25.067	-19.771	**-103.985**	**8.789**	**6.430**	(RF)
AN	Mean	**0.535**	**96.580**	**1.489**	**-0.671**	**224.419**	**-155.231**	-91.610	89.083	58.083	Mean
	Median	0.528	97.276	-8.873	-0.882	238.166	-171.510	**-80.097**	**83.362**	**51.067**	(RF)
TN	Mean	**0.437**	**0.153**	**0.003**	**-0.602**	**0.403**	**-0.249**	-63.525	0.090	0.066	Mean
	Median	0.411	0.157	-0.024	-0.950	0.445	-0.310	**-37.921**	**0.069**	**0.050**	(RF)
CEC	Mean	**0.342**	**8.516**	**0.168**	**-1.280**	**20.586**	**-15.277**	-47.768	7.887	6.714	Mean
	Median	0.322	8.644	-1.273	-1.706	22.427	-17.579	**-31.976**	**6.486**	**5.280**	(RF)
porosity	Mean	**0.286**	**5.496**	**-0.028**	**-6.014**	**9.548**	**-8.380**	**-10.608**	10.167	9.167	Mean
	Median	0.283	5.507	0.064	-6.041	9.566	-8.436	-10.728	10.219	9.236	(RF)
TK	Mean	**0.254**	**0.569**	**0.004**	**-6.439**	**1.133**	**-0.985**	-7.496	-0.921	0.842	Mean
	Median	0.251	0.570	-0.022	-6.856	1.164	-1.007	**-6.626**	**0.873**	**0.772**	(RF)
TP	Mean	0.039	0.153	**0.001**	-0.073	**0.471**	**-0.114**	-45.798	0.047	0.040	**Median**
	Median	**0.042**	**0.153**	-0.012	-0.092	0.475	-0.136	**-23.025**	**0.034**	**0.029**	**(QRF)**
AK	Mean	**0.161**	**169.589**	**1.120**	**-0.250**	**484.127**	**-235.202**	-74.502	91.844	77.809	Mean
	Median	0.130	172.666	-24.174	-0.413	514.801	-285.971	**-46.213**	**72.628**	**61.121**	(RF)
AP	Mean	**0.137**	**10.600**	**0.284**	**-1.000**	**29.102**	**-21.562**	-217.302	6.999	6.334	Mean
	Median	0.075	10.976	-2.468	-1.470	32.340	-25.594	**-90.100**	**4.521**	**4.074**	(RF)
R (Wet)	Mean	**0.275**	**33.108**	**0.032**	-10.615	**56.055**	-50.741	-10.427	54.593	49.311	**Median**
	Median	0.271	33.212	0.081	**-10.481**	55.730	**-50.198**	**-10.363**	**54.441**	**48.539**	**(QRF)**
G (Wet)	Mean	**0.258**	**32.333**	**0.076**	**-12.180**	**55.557**	**-51.001**	-24.998	52.446	48.498	Mean
	Median	0.244	32.639	-0.777	-12.543	56.317	-51.137	**-24.089**	**45.522**	**46.730**	(RF)
B (Wet)	Mean	**0.205**	**34.046**	**0.021**	-9.174	**57.428**	-52.629	-75.942	54.758	50.755	**Median**
	Median	0.193	34.305	0.934	**-8.686**	56.034	**-50.974**	**-74.629**	**54.168**	**49.383**	**(QRF)**

R (Dry)	Mean	**0.256**	**34.204**	**0.041**	-11.524	58.243	-51.861	-11.524	56.236	50.954	**Median**
	Median	0.249	34.331	0.095	**-11.142**	**57.531**	**-51.256**	**-11.321**	**56.112**	**50.364**	**(QRF)**
G (Dry)	Mean	**0.269**	**31.238**	**0.067**	**-11.173**	**54.248**	**-50.843**	-23.128	50.571	46.368	Mean
	Median	0.254	31.854	0.421	-11.534	55.658	-50.994	**-22.451**	**46.358**	**43.589**	(RF)
B (Dry)	Mean	**0.213**	**33.224**	**0.020**	-9.854	56.552	-52.223	-74.642	53.775	49.228	**Median**
	Median	0.204	33.612	0.635	**-9.347**	**55.012**	**-50.128**	**-73.734**	**53.127**	**48.581**	**(QRF)**

(a) BD at 0-5cm interval (Median)
(b) BD at 0-5cm interval (Mean)
(c) BD at 5-15cm interval (Median)
(d) BD at 5-15cm interval (Mean)
(e) BD at 15-30cm interval (Median)
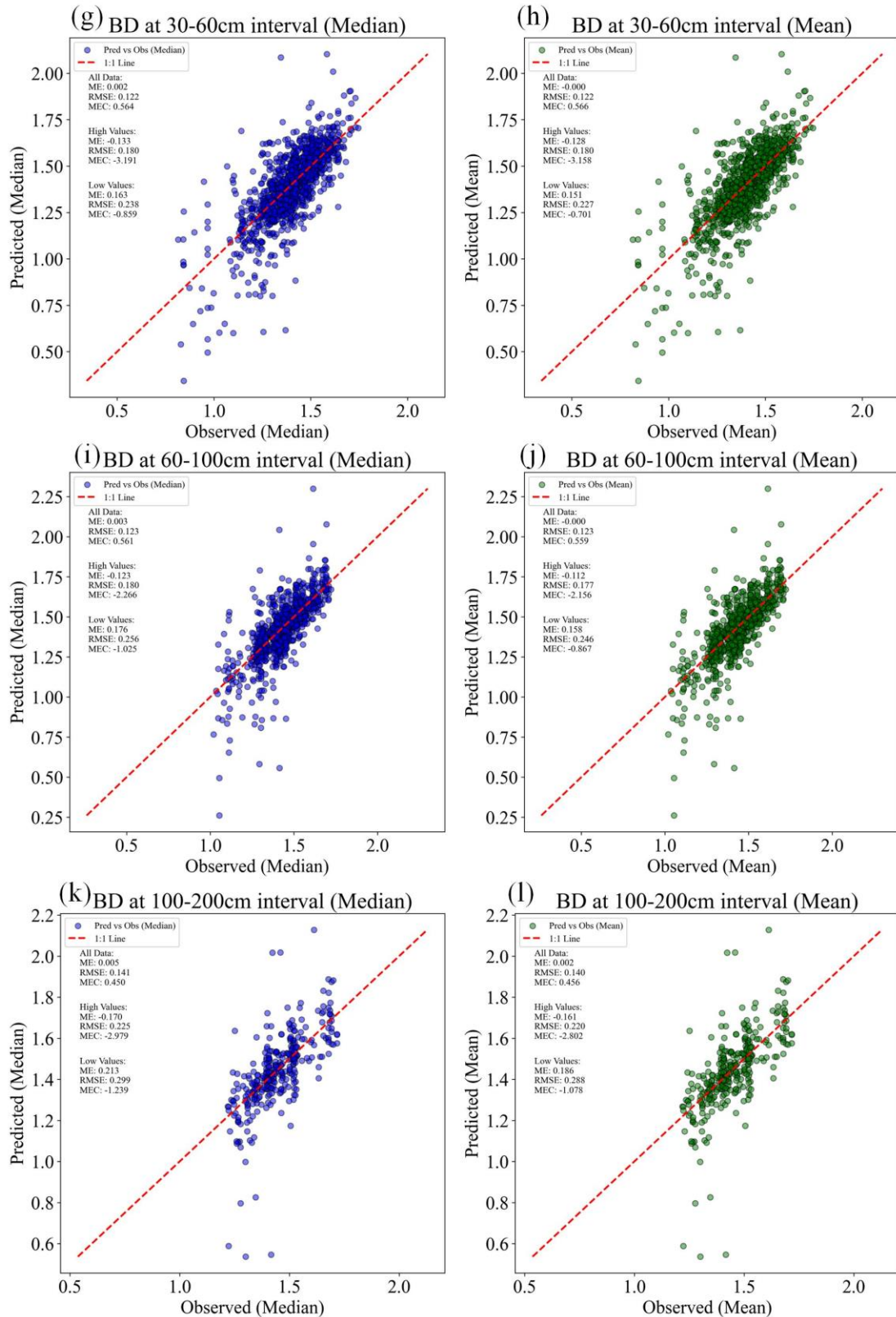(f) BD at 15-30cm interval (Mean)

Continue

Figure. Predicted median (a, c, e, g, i, k) and mean (b, d, f, h, j, l) bulk density (BD) at various depth on the y-axis vs. measured BD content on the x-axis. Accuracy plots and metrics (ME, RMSE and MEC) were computed using 10-fold cross-validation.