Response to Referee Comments on Manuscript ESSD-2026-447

Dear Editor and Referee,

We sincerely thank the Editor-in-Chief, Dr. Baptiste Vandecrux, for the careful handling of our manuscript, the continued guidance throughout the review process, and the opportunity to further improve our work. We are also grateful to the referee for the careful re-evaluation of the revised manuscript and for the two constructive final comments. We have addressed both points through targeted textual revisions. Our point-by-point responses are provided below.

Referee #1:

Comments 1:

Conditional nature of the reported accuracy. The quoted validation statistics (R^2 = 0.720 for the 10° × 10° block cross-validation, 0.780 for LOSO-CV) describe agreement within the station network that was used for calibration, and the Discussion already notes that performance decreases with stronger spatial separation. What is not yet visible to a reader of the Abstract is the magnitude of this conditionality: Table S4 shows that RMSE rises from 26.46 cm for ALT < 100 cm to 84.04 cm for ALT > 300 cm, with a bias of −68.12 cm in the deepest class, the latter being based on only 84 records from 19 sites. Please add one or two sentences to the Abstract or to the Conclusions stating that the reported R^2 values refer to agreement within the calibration network and degrade for deep ALT and in data-sparse regions, so that they are not read as a uniform per-pixel accuracy.

Response 1:

Thank you for this important clarification. We have revised both the Abstract and the Conclusions to make this limitation explicit. The revised text now clarifies that the reported cross-validation statistics summarize model performance across the available monitoring network and that prediction errors increase substantially under deep-thaw

conditions and in sparsely represented environmental settings. We also emphasize that deep-thaw and extrapolated conditions should be interpreted with additional caution.

Comments 2:

Scale of applicability: regional gradients versus local variability. Section 3.6 correctly states that the nominal 1 km grid does not imply a 1 km effective resolution, and lists microtopography, ground ice, soil stratigraphy, snow redistribution and local hydrological variability as unresolved controls. The predictor set provides an additional and more fundamental reason that deserves to be stated explicitly: among the 12 predictors (SR, MAT, MAP, SDE, TI, NDVI, Biome, Elev, Slope, BD, CF and Silt) there is no variable representing soil water content, even though soil moisture is a primary control on local ALT. The SHAP ranking (elevation 34.2%, bulk density 13.1%, MAT 12.6%, TI 8.8%, slope 7.6%) confirms that the model is driven mainly by large-scale climatic and topographic gradients. The product can therefore be expected to reproduce hemispheric-scale structure - the increase in ALT from the Arctic towards the southern permafrost limit, and its decrease from lowlands towards high mountains such as the Mongolian and Tibetan Plateaus - while kilometre-scale contrasts within a region are largely damped, with mapped values tending towards the regional mean. Please state this distinction explicitly and characterize the mapped values as regional-scale estimates rather than local predictions.

Response 2:

Thank you for this important suggestion. We have revised Section 3.6 to clarify this distinction. Specifically, we now state that fine-scale controls on ALT, including local soil-moisture conditions, microtopography, ground ice, soil stratigraphy, snow redistribution, and local hydrological variability, are not fully resolved by the current predictor set. As a result, kilometer-scale contrasts within individual landscapes may be damped. We have also explicitly characterized the mapped ALT values as regional-scale estimates represented on a 1 km grid rather than as local predictions at the individual-pixel scale. The Conclusions have been revised accordingly to reinforce this interpretation.