

Response to referee #3 (report #1)

I would like to thank the authors for comprehensively addressing the reviewer comments. I find the manuscript sharper, more concise, and better organized. I have only one remaining comment which I expect authors to address, as it has also come up in the last review round, and still hasn't been comprehensively addressed.

Thank you very much for the positive feedback. We considered the last remaining comments in the revised manuscript. Please find the point-by-point response to the comments in the following sections.

Major comment

L207-L210: I do not understand how the scatter between Figure 2d and Figure 2f can be compared as the y-axis in Figure 2f is massively zoomed in (variation of ONLY 0.04C in Figure 2f vs variation of >5C in Figure 2d). Simply mentioning this in the figure caption does not solve this problem as the linear relationship is later used in the manuscript to justify ETA 20 (see L207-210). Reviewer #4 pointed this out clearly in the previous review round, however this has not been adequately addressed. I particularly find L209-210 incorrect: "Above eta level 20, the scatter progressively increases, and by eta level 49, no clear relationship with surface elevation remains". At eta level 49, the absolute temperature difference between different land use classes are most likely smaller than at eta level 20. The larger scatter in Figure 2f is because of excessive zooming, making the corresponding text in the manuscript incorrect.

Yes, we apologize, the scaling of the y-axis should have been homogenized earlier. We fully agree that the individual-scaled y-axis limits comparability, so we have revised the axis range (see Fig. 1).

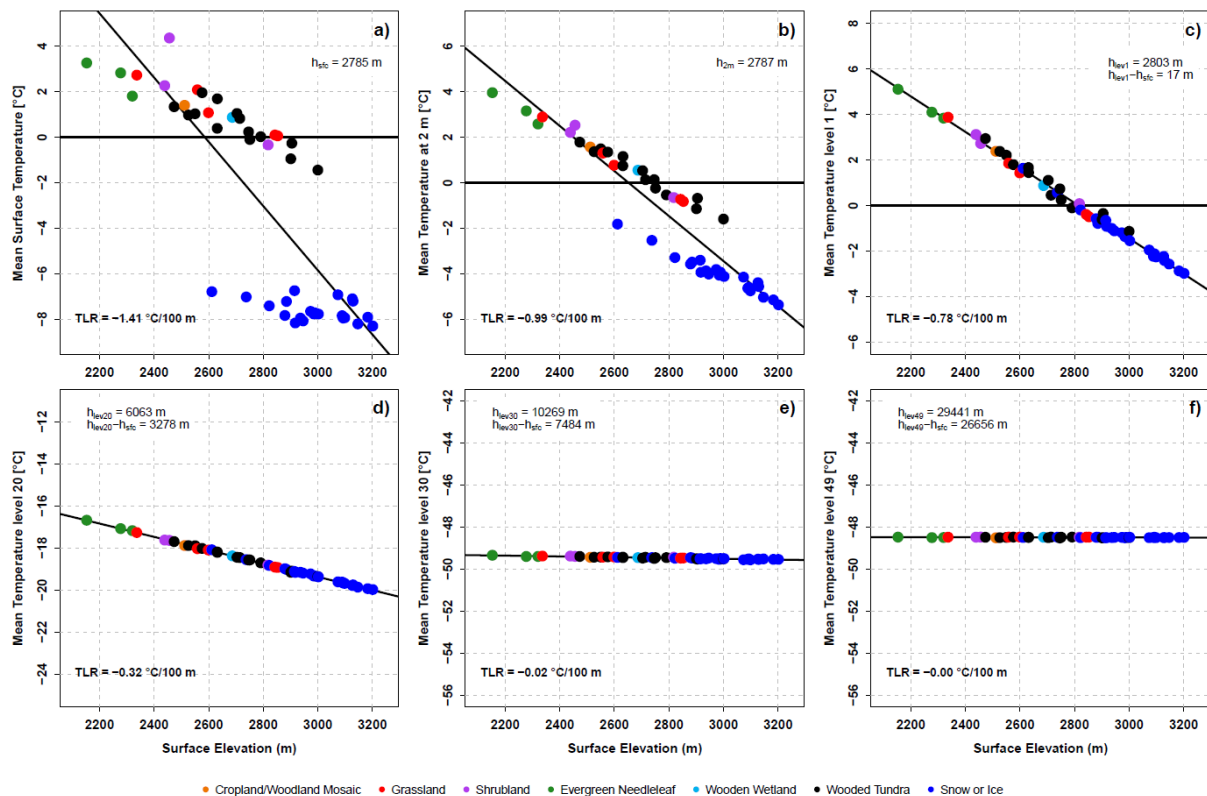


Figure 1: Mean WRF temperature for different land use categories and surface elevations (m) for the Kauner Valley domain covering a test period of three months. The diagonal black line shows the linear gradient between elevation and temperature obtained by least-squares fitting to all data points. The numbers in the subplots indicate the respective geopotential height, the height above ground, and the

temperature lapse rate (TLR). The horizontal black line indicates the freezing level at 0 °C. The y-axes of all six panels share the same 14 °C y-axis span, each centered on that panel's own mean.

We have also revised the corresponding paragraph in the Methodology section (lines 206:217):

“From eta level 1 to eta level 20 (Fig.2c,d), WRF temperature exhibits an almost perfectly linear relationship with surface elevation and very little scatter. This indicates that temperatures at these levels are predominantly governed by a nearly uniform domain-wide lapse-rate structure, with only minor influence from land-cover or snow-cover differences. Because all panels in Fig.2 share a common y-axis span, this weakening relationship is directly visible as a progressive flattening of the fitted slope with height: above eta level 20 the regression line becomes increasingly shallow, and by eta level 30 and eta level 49 the data points collapse onto an almost horizontal band, indicating that surface elevation no longer exerts a discernible control on temperature at these levels. This behavior reflects the decreasing influence of terrain with height: the lower eta levels closely follow the terrain, whereas higher eta levels become increasingly horizontal and sample the free atmosphere, where temperature variations are controlled more by large-scale atmospheric structure than by the underlying surface elevation. The lapse rate at eta level 1 equals $-0.78\text{ }^{\circ}\text{C }100\text{ m}^{-1}$, i.e., it is higher than in the free atmosphere (which corresponds to $-0.65\text{ }^{\circ}\text{C }100\text{ m}^{-1}$) for the three-month test period, and it decreases steadily with height, reaching $-0.32\text{ }^{\circ}\text{C }100\text{ m}^{-1}$ at eta level 20, and approaching zero ($-0.02\text{ }^{\circ}\text{C }100\text{ m}^{-1}$ and $-0.00\text{ }^{\circ}\text{C }100\text{ m}^{-1}$, respectively) at eta levels 30 and 49.”

Minor comments

L404-406: Thanks for explaining the rationale for choosing eta levels 1-20. As the cutoff of ETA level 20 is not based on any statistical analysis conducted in this study, I would suggest adding a reference to support this statement.

We thank the reviewer for this comment. We agree that the selection of eta level 20 should not be interpreted as a statistically optimized or universally applicable cutoff. Rather, it represents a pragmatic choice for the present WRF configuration, based on the sensitivity analysis presented in this study. Specifically, it balances the stronger sensitivity associated with the lowest atmospheric levels against the need to retain a coherent representation of the broader lower-atmospheric temperature gradient. Because the relationship between eta levels and near-surface temperature gradients depends on the model configuration and local atmospheric conditions, we consider a specific cutoff to be configuration-dependent rather than generally transferable. We have therefore revised the corresponding section of the manuscript to clarify the empirical and pragmatic basis for adopting eta levels 1–20, rather than presenting level 20 as a statistically derived or generally applicable threshold (lines 405:414):

“Although individual eta-level combinations yielded slightly better performance at specific stations, the best-performing combinations differed among sites, indicating limited spatial transferability. Therefore, a bulk lapse rate derived from eta levels 1–20 was adopted as a pragmatic compromise between reducing the sensitivity to individual near-surface levels and retaining a coherent representation of the lower-atmospheric temperature gradient. This choice should not be interpreted as a statistically optimized or universally applicable cutoff, but rather as a configuration-dependent choice based on the sensitivity analysis presented here. Because the relationship between eta levels and near-surface temperature gradients depends on the model configuration and local atmospheric conditions, the selected range is not intended to represent a generally transferable threshold. This approach avoids station-specific optimization and provides a consistent framework for temperature extrapolation across catchments with differing topographic and climatic characteristics.”

L353-354: I believe this sentence should be mentioned in Methodology and not Results. Sentence: “The uncorrected WRF temperature corresponds to the hourly WRF 2-m air temperature (T2) of the respective 2-km grid cell and is adjusted for the height difference between WRF and station elevation (Eq. 6).”

Thank you for pointing out this inconsistency. We integrated this and the following sentence in section 2.4.1 of the Methodology (lines 276:279).