REVIEWER 2

COMMENTS:

1. *L202 Can you comment if 70 predictors are in normal range for this type of models?*

   Answer:

   The number is more than usually reported in literature. Typically, there are around 10-20 features. These are mostly a few basic meteorological variables (air temperature, precipitation) and a set of environmental variables (related to topography and vegetation). A great majority of features occurring in literature is included in our model. There are a few reasons that our dataset is bigger. At first, predictions of our RF model are time independent and therefore, some variables have been aggregated in different time intervals. Also, temporal variables had to bee added, which is not the case when an LSTM is run (e.g., Cui et al., 2023). Secondly, some input variables are highly correlated, despite their significant score in variable importance metric. These are for example: snow density and snow albedo, 850hPa air temperature and 2-m air temperature. Last but not least, some variables have to be doubled (despite their high collinearity) since we take two reanalyses on input (e.g., snow depth, model elevation). An interesting approach was adopted by Tanniru et al. (2025) – they initially prepared 40 predictors, however, they excluded those which were highly correlated (R>0.80), so they ended up running a model with 16 predictors.

   A short explanation has been added in Section 2.5.

2. *L204 Why was number of trees set to 100?*

   Answer:

   You must have mistyped the number of trees as in the paper stands 1000, not 100. The parameter is equal by default to 500 and was later doubled as more and more predictors where included into training. Boehmke and Greenwell (2019) suggest it should be around 10 times their number. Additionally, increasing the number of trees ensures more stable estimation of the out-of-bag error and variable importance (at the expense of longer training time though). A short justification was added in the manuscript.

3. *L214 Can you elaborate bit more why no spatial split of data was tested? Even for just few years? It would be interesting to see how model performs in new locations.*

   Answer:

   Thank your for this valuable suggestion! The training strategy has been extended so that, beside temporal cross-validation, also spatial split is considered. Fur this purpose, the study area has been split into five longitudinal blocks (2° each). Such a split ensures that every block includes both lowland stations with rare snow occurrence as well as elevated stations with deep snow cover. Then, cross-validation was performed with 4 blocks being a training set, and the remaining one – a test set. The variability of the RMSE training error across different combinations is limited which confirms solid generalization skill of the RF model. Section 2.5 has been updated on the above mentioned information.

4. *Figure 2 The caption text is little confusing, maybe clarify that left side are the input predictors and right is the target (station data) used for training.*

   Answer:

   The caption has been corrected according to Your suggestion.

5. *L 248 I recommend adding subsections to the result section to make it easier to follow.*

   Answer:

   The section have been split into four subsections.

6. *L248-263 This paragraph is quite long, considered dividing into two parts*

   Answer:

   The suggested change have been introduced.

7. *L260 What about station near the sea level but away from the shore?*

   Answer:

   In order to examine it, a plot of MAE as a function of elevation was constructed considering only stations located below 150 m a.s.l. (see Fig. 2 below). Additionally, distance to the sea coastline (in km) is indicated for each station. In fact, there are only a few stations situated below 50 m a.s.l. that are not at the coast. An increase of MAE for coastal stations is clearly evident if only synops are considered (red dots). In case of lower-rank stations, two climate stations (blue dots) exhibit an increased MAE, one of them lies just 2 m a.s.l. but 18 km away from the coastline. However, their MAE values do not stand out that much from other, more inland climate stations. Consequently, there is not enough evidence to conclude that stations lying near sea level but away from the coastline have increased errors.
   A minor change have been made in the manuscript to emphasize that this concerns mostly synop stations.

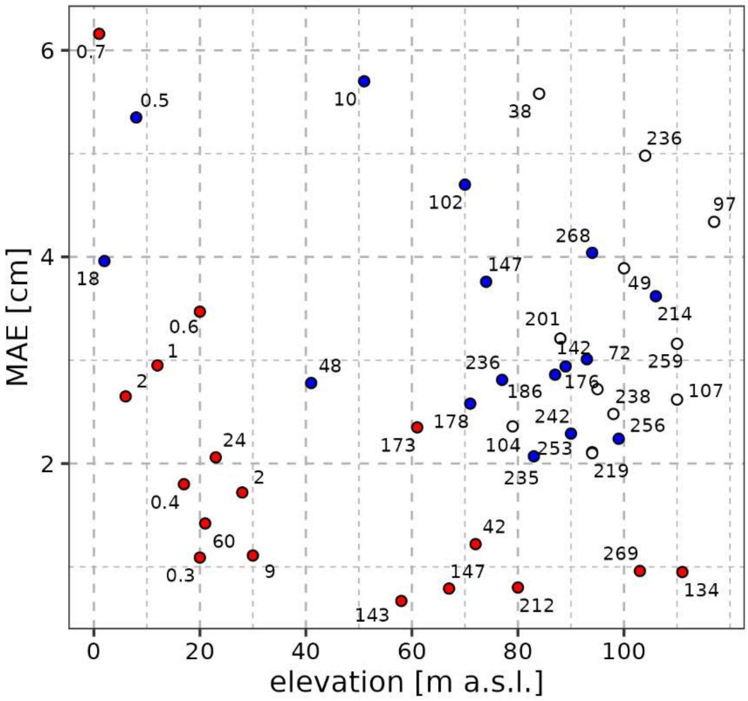

Fig. 1 MAE of snow depth in relation to elevation in ERA5 for stations lying not higher than 150 m a.s.l. Distance to the Baltic coast is showed as a label next to every point (in km). Stations are coloured according to their rank (red - synop, blue - climate, white – precipitation).

8. *Figure 7, Quite hard to see different years in figure (even bolded ones)*

Answer:

The figure has been significantly modified. Instead of spaghetti plot, a ribbon (plume) plot have been used so the readability is enhanced.

9. *L312 Might be good idea to remind reader about the resolutions of the grids here (or in figure 8 caption)*

   Answer:

   Thank you for this relevant suggestion. Information about horizontal resolution has been added in brackets in the indicated sentence as well as in the preceding one.

10. *L315 Can you clarify if data from the stations used to validate downscaling was also used for training model?*

    Answer:

    The training dataset does include data from these stations, however, not from this winter season. The predictions were generated through the strategy of temporal split (19 seasons of training, one for testing). We laid particular stress to the fact that a test set includes solely unseen data. We can assure you that there is no data leakage.

11. *L424 There are two commas after "all"*

    Answer:

    The redundant comma has been removed.

12. *L455 While it's clear to most readers what "this part of Ventral Europe" means, might be good idea to be bit more precise here*

    Answer:

    Thank you for this relevant suggestion! We have added a directional term "north-eastern" to be more precise.

REFERENCES:

Boehmke, B. and Greenwell, B.: Hands-On Machine Learning with R, 1st ed., Chapman and Hall/CRC, https://doi.org/10.1201/9780367816377, 2019.

Cui, G., Anderson, M., and Bales, R.: Mapping of snow water equivalent by a deep-learning model assimilating snow observations, Journal of Hydrology, 616, 128835, https://doi.org/10.1016/j.jhydrol.2022.128835, 2023.

Tanniru, S., Singh, D. K., Singh, K. K., and Ramsankaran, R.: Exploring Machine Learning's Potential for Estimating High Resolution Daily Snow Depth in Western Himalaya Using Passive Microwave Remote Sensing Data Sets, Earth and Space Science, 12, e2024EA003849, https://doi.org/10.1029/2024EA003849, 2025.