Title: On the proper use of temperature screen-level measurements in weather forecasting models over mountains.
Author(s): Danaé Préaux et al.
MS No.: egusphere-2025-708

Summary

This manuscript investigates from different perspectives how the current usage of near-surface temperature observations, usually considered measured at 2-metre, impacts model evaluation and data assimilation in mountainous regions in an operational NWP model. The authors consider three different sources of uncertainties: 1) altitude difference between model grid-point and station height; 2) difference in the height above the surface of the sensor and the one of the model, usually assumed constant at 2-metre above the ground; 3) inhomogeneities in the observation's distributions.

This is an important topic as usually point 2) and point 3) are not considered in-depth during both model evaluation and assimilation of surface observations. A better usage of these observations can potentially improve forecasting of near-surface variables and give more meaningful indications on model developments pathways.

I find some of the conclusions worth publishing as they can be a good contribution to the present literature. In particular, the impact of the difference in the height above ground between the sensor and the model in case of snow cover is of particular importance, as this is usually neglected in operational data assimilation systems. However, there are major concerns (see "General comments" below) that should be addressed before publication.

General comments

The manuscript should be deeply revised before resubmission. There are typos, missing parentheses and many sentences that are very difficult to read in English, which I strongly recommend rephrasing. A few examples are given in the specific comments below, but I encourage the authors to edit the text throughout. Also, in some places the tone of writing is too "colloquial", for instance see lines 464-466. While posing questions can be a good way to engage the reader, an excessive use does not align with the scientific style of a research manuscript.

From a scientific perspective, I also strongly suggest revisiting the way the scientific questions are introduced and some of the results presented. In the introduction, it is hard to clearly get the scientific questions the authors would like to answer and what is the methodology used to address them. One way could be to clearly states what hypotheses and questions relate to model evaluation, and which ones to data assimilation. The last paragraph (lines 107 – 111) has elements of it, but it should be improved. For instance, the hypothesis that "In the present paper we examine the assimilation of mountain near-surface temperatures as a possible cause for the cold bias observed in Arome forecasts. ", in Sect. 2.3, should be introduced earlier in the text, possibly in the Introduction. Sect. 3.2 should be better introduced to guide the reader that the manuscript is now moving to the "assimilation" part and analysis of the sensitivity experiments.

Regarding the logical order, some of the “Discussion” sections seems more “Results” and should be rearranged accordingly. For instance, Sect. 4.2 seems better placed together with Sect.3.1.1. Figure 8 is discussed in Sect. 3.1.3 so it should be introduced accordingly in the text.

Some parts of the Discussion and Conclusions are too tightly related to AROME. This could be ok, as this is the numerical tool the authors are using, however I strongly encourage the authors to draw some generality out of some of these statements for the wider community, for instance talking about the physical reasons to use or replace a particular scheme rather than referring to a specific scheme’s name.

From the point of view of the methodology, I have a few issues with the sensitivity experiments performed and the analysis presented in Sect. 3.2 and Figure 7.
Firstly, it is hard to understand what is plotted in Figure 7: it should be better clarified when the difference between the sensitivity experiments and the control are plotted; I do not understand the reason to use a figure legend that is different from the sensitivity experiment names, which adds confusion to the reader. Secondly, it is not clear to me what the “mountain” line represents in Figure 7: the NO_NIGHT experiment already removes the assimilation of all surface observations during nighttime (at least this is the understanding from Sect. 2.3), so what is the reason to combine it with the NO_VALLEY experiment increments? Furthermore, if the aim of the authors was to check the impact of all mountain surface observations, an additional sensitivity experiment, in which all T2m observations were removed, would have been useful. The linearity assumption used to combine the increments can be hard to justify, given the high non-linearities present in NWP models. At least the authors should justify why this experiment was not performed.

Another issue is related to the analysis in Sect. 3.1.1, in particular the author’s statement that the approximation T2m ~ T5m is incorrect (line 306). From their results this conclusion is a bit misleading as on average, considering the whole winter periods, they show that this approximation is valid. The authors base their conclusion on the analysis of a specific case study in Figure 5, covering only a few days. To make their conclusion statistically stronger, I think the authors should also show a diurnal cycle of temperatures (or some statistics) computed only for anticyclonic conditions or clear-sky periods.

Specific comments

Abstract: “rôle” → “role”

Line 28: “… they are originally designed for Rudisill et al. (2024); Gouttevin et al. (2023).“ Some parentheses might be missing for the references.

Line 35: “While primarily strong over peaks and ridges, it often comes with a warm bias in valleys. “ is not clear. Do you mean ”…it is often associated with a warm bias in valleys”?

Line 38: I think a few more references on the literature review and/or the forecaster’s reports would be useful to better justify the 3 types of biases described.

Line 44: “wood fire heating Aymoz et al. (2007)” some parentheses are missing for the reference.

Line 46-52: Not clear, could you please reformulate? Also It seems that the reference justifying the argument is put at the end (Beauvais, 2018) but should it be placed at the beginning of the paragraph?

Line 75: what are Nivose stations? Please introduce them properly.

Line 83: “have been a preoccupation for numerous modelers”, not clear. Do you mean “have been an issue recognised by numerous modelers”?

Line 83: “kind of covering up the height-above- surface adjustments that we here mention. “, please reformulate this.

Line 99: “mountain” → “mountains”

Line 108: “pitfall” does not sound very scientific, do you mean “challenges”?

Line 142: “T5m_ mod refers abusively to the temperature..”, not clear. What do you mean by “abusively”? Also I think there is a trailing space between T5m_ and “mod”

Line 154: “This stage eliminates observations that are considered doubtful because they come from a non-qualified source or are too far away from the design.” Is not clear, please reformulate. Do you mean that differences with the first guess are larger than a certain threshold?

Line 154: “unfairly” is not a scientific term here.

Cost function Equation: Please add an equation number to the text.

Sect. 2.1.2: Why CANARI and the 2D-OI are introduced? It does not seem this is used at all in the manuscript, as the analysis focusses on 3D-Var, as far as I understood.

Line 161: “Increments are calculated for the surface observations and for the upper-air observations. Then, J is minimized using these increments.” I am not sure if here “increments” is the right terminology. Do you mean the x-xb terms in the cost function equation?

Line 168-172: This is not clear. How B is defined in the 2D OI equation? What do you mean “that very few observations are important in determining the analysis increment”? Please clarify and/or reformulate.

Line 177: “1D IO” I think should be “1D OI”.

Line 180: “(Figure 2: map on the right) “, please add a label to each panel and refer as “Figure 2b”.

Line 204: "…, with for the latter a measurement co-located with every other observation due to the high spatial variability of the snow height at this site." It is not clear, please reformulate.

Line 213: I think the authors refer to the Nivose stations in the text before these are described/introduced. Please adjust the text.

Line 221: This hypothesis would be better placed at the end of the introduction to better illustrate the scope of the work. I think this is somehow described at line 96-97, but it is not clear enough, in particular the link to the model cold bias. The authors should also explain why an assimilation deficiency should cause a forecast bias (at which lead times?).

Line 224: "These numerical simulations are be compared to a reference. ". Please reformulate.

Figure 3 caption: "exemple" → "example"

Line 283: "et" → "and"

Line 306-310: "Thus, the approximation is invalidated: …" please reformulate. It is not clear at this stage what approximation the authors are referring to, even though it is clarified in the remaining of the paragraph.

Line 306-310: Can this argument be better generalised for instance by computing a diurnal cycle only for the anticyclonic conditions in the considered 4-year period?

Line 343: The discussion in the text "jumps" from Figure 4/5 to Figure 8, which is a bit confusing. Please readjusts the figure order or the text.

Section 3.1.3 is difficult to follow. Could you please reformulate the Section to better distil the main message the authors would like to convey? For instance, reducing where possible references to "see above text" or if necessary, better explaining where a reader should focus, e.g. a figure or a particular subsection.

Line 369: "Secondly, we note that the analysed T5m is worse at Nivose stations". Worse than what?

Line 383: "esp. ". please correct this.

Line 387: "Mountainous areas are complex to instrument and model.", please clarify what "complex" means in this context.

Table 2: typos in "Biais", no closing parenthesis in figure caption.

Line 495-497: Please rephrase.

Line 513 – 520: This discussion is too much related to Meteo-France's models and using a lot of details that are not of interest to an external reader. The authors should

give the information that would be of interest for the community, not the namelist settings used in operational AROME.

Line 588: “Correcting T2m biases would also enable the model’s physical parameterisations to be improved.”. Is it not the other way around, improvements in physical parameterisations reducing the T2m biases in the first guess and hence reducing the “activity” of the data assimilation system?

Line 596: Is there any reference to previous work that could back up this hypothesis of misrepresented katabatic winds?

Line 599-606: This long paragraph mixes perspectives and hypotheses for future work with conclusions from this work. I would suggest reorganising it so that the conclusions from this work are clearly divided from the perspectives. E.g. the recommendation to properly consider the height above the surface of the observations in assimilation and evaluation should come as a conclusion, whereas reducing temperature biases through improved model physics (using Isba-ES-DIFF etc.) should come as a perspective.