“Comparison of temperature and wind between ground-based remote sensing observations and NWP model profiles in complex topography: the Meiringen campaign”

First, we thank the reviewer for the valuable, in-depth comments to our manuscript. Further analyses were done to explain that the wind from the Saneraatal can be explained by its bigger volume compared to the Haslital. This result improves the comprehension of the influence of the topography on the thermal wind system of the valley.

The answers to the comments and questions are written in italic thereafter. The explanations of this document cite the numbering of the figures in the first revision in accordance with the lines' numbers of the comments.

Answers to the reviewer 2 comments

Second review of egusphere-2023-1961

‘Comparison between ground-based remote sensing observations and NWP model profiles in complex topography: the Meiringen campaign” by Alexandre Bugnard et al.

Summary

The manuscript presents a comparison of campaign observations from the Meiringen Campaign in a narrow Swiss Alpine valley with the high-resolution 1-km KENDA analysis. The comparison focuses on temperature and wind profiles measured by a microwave radiometer and Doppler wind lidar, respectively, for ten months during 2021/2022. It is shown that observed and modeled seasonal climatologies of temperature and wind profiles agree well, although for specific situations, such as for example temperature inversions or foehn events the differences are relatively large. The manuscript also links the complex topography to thermal wind systems and presents cross- and along valley flow systems observed during the campaign period.

The manuscript presents valuable observations from an Alpine site and provides new insights in the quality of the high-resolution analysis in complex terrain and shows examples of how specific terrain-induced flow features can influence the differences between observations and the analysis.

The manuscript has been substantially improved during the first revision phase and the authors have addressed major reviewers' comments, i.e., the general structure was revised and the manuscript was streamlined. I support publication of this manuscript but I have several mostly minor comments that should be clarified and addressed prior to publication.

1 Comments

1. KENDA T bias

I still struggle to fully follow the discussion about KENDA temperature biases discussed in Section 3.1. l. 213 ff states: ”The main observed pattern is a general low altitude (< 1500 m) T underestimation from KENDA-1/MEE.” This cold bias pertains to all hours of the day and all months (Fig. 2b; except June). Subsequently, Fig. 4 shows that ”KENDA-1 overestimates the T during nighttime (+1.5C) in both cells and underestimates it during the day (-2C in MEE and -1.5C in MER).” (l. 240 ff). I understand that different data and levels are compared in Fig. 4, however, at nighttime Fig. 4 suggests that KENDA is larger than MWR. The latter is not visible in Fig. 2b. I would ask the authors to elaborate on and clarify the KENDA warm or cold bias, respectively, and streamline this paragraph. If the main reason of the KENDA-SMN bias results from differences in altitude between KENDA grid box and SMN observation and the frequent presence of inversions, Fig. 4 and the respective text may be misleading. From Fig. S5 I cannot induce if KENDA overall over or underestimates temperature.

The applied principle is to use the nearest data for each comparison. As already specified in the manuscript, the MER is at 574 m, so that the 2 m T data are observed at 576 m and the 10 m wind data at 584 m. MEE is at 589 m. The first MWR/MEE is at 625 m and the first DWL/MEE level at 775 m. The manuscript also specified the difference between KENDA-1 first level and the real topography (109 m at MER and 130 m at MEE) as well as the altitude of KENDA-1 first level (20m a.g.l.). The revised version specifies now in sect 2.2 that the first level is 705 m (574-109-20) for KENDA-1/MER and 739 m (589+130+20) for KENDA-1/MEE. As proposed, the observations' levels are now given in the figure or caption of Figs. 3, 4, 5, 7, 10 and 11.

Considering the applied principle, the first level of comparison of MWR/MEE and KENDA-1/MEE (Fig. 2) is 739 m, whereas Fig. 3 and 4 compare always the lowest level with SMN/MER. There is then an underestimation of KENDA-1/MEE compared to MWR/MEE between 739 and 1500 m. KENDA-1/MEE and KENDA-1/MER overestimate SMN/MER T (at 576 m) during night and underestimate it during day. Both Fig. 4 and Fig. 2 show that, during day, KENDA-1/MEE at 725 m modeled lower T than observed by MWR/MEE at 625 m. During night, KENDA-1/MEE often misses the T inversion leading to an overestimation of ground T and an underestimation of T below 1500 m. Fig. S5 presents explicit examples of this phenomenon leading to a negative difference KENDA-1/MEE-MWR/MEE at 705 m and a difference KENDA-1/MEE(at 705 m)-SMN/MER larger than MWR/MEE (at 625 m)-SMN/MER. In case of missed T inversion, KENDA-1/MEE extrapolated at 625 m (dashed red line in Fig. S5) would overestimate the T observed by MWR/MEE at the same altitude.

The text was adapted to better explain this apparent discrepancy between Fig. 2 and Fig. 4: "The missed T inversions by KENDA-1/MEE lead to both its important overestimation of the T at ground level (Fig. \ref{fig:boxplot_hr}) and its slight T underestimation between ~850-1200 m (Fig. \ref{fig:T_clim}). Detailed examples of T profiles during a day with missed T inversion by KENDA-1/MEE (Fig. S5) show these opposite T bias at several altitudes with SMN/MER and MWR/MEE observations."

2. Altitude depiction in observation and KENDA data

I would appreciate if the authors could provide information directly in the text, figures, and/or figure captions about the altitude of the used data. It is difficult to remember the altitude of lowest model grid box at MEE/MER as well as of MRW and DWL. I think this would facilitate following the storyline of the manuscript.

The altitude of the T and wind measurements were added in the captions of Figs. 3, 4, 5, 7, 10, and 11.

3. I would ask the authors to again double-check the manuscript for typos, missing words, and grammar mistakes (e.g., l. 36 ”Such inversions are favored in complex topography (Joly and Richard, 2018) and persist-s longer in deeper valleys, whereas inversion lifetimes converge to the one over a plain for wide valleys (Colette et al., 2003).”; l. 40 ”The quality of predictions for”, l. 392 ”of a monthly median values”). Please also consistently adjust the date/time format.

A colleague with high English skills reviewed the manuscript. In particular, the mentioned sentences were modified.

4. l. 6: Please introduce the acronyms COSMO-1E and KENDA-1 as not everyone may be familiar with the terminology.

Done

5. l. 19: ”of a model”: I would specify this and explicitly mentions ”KENDA-1”.

Done

6. l. 125: ”Vertical levels with spacings from 20 m at the surface”: What is the height of the lower-most level?

The lowest most level is 20 above the surface of the model’s terrain. It has been inserted in the text.

7. l. 176: ”from 200 m to 12000 m above ground”. Can the DWL measure successfully up to 12 km height?

*The DWL can effectively measure at 12km when cirrus clouds are present. This sentence is then correct even if such an altitude is not met in the absence of cirrus cloud*s.

8. l. 186: ”Even if SNM/MER surface observations are assimilated by KENDA-1, the comparison of the modeled and observed data allows evaluating the impact of the assimilation at MER.” Please clarify this sentence. From a comparison of the resulting KENDA analysis and the assimilated observations alone, the observation impact cannot be deduced directly, unless first-guess (as mention in l. 272 ll) is available.

The referee is right, the real impact of the assimilation cannot be estimated without a comparison with the first guess. The sentence was modified: “The comparison between KENDA-1 and observed data at MER allows evaluating the model's performances at a station, whose SNM surface observations are assimilated.”

9. l. 296: ”direction at low altitudes (800-1000 m) is mainly from W-SW”: I find this very difficult to see in Fig. 6, among others, as the 800-1000 m layer is very shallow and the colors are not distinct. l. 296: ”in the rest of the profile”: Please specify.

First, Fig. 6 was changed and no further distinction between wind speed lower or higher than 20 km/h is made. The text was also adapted considering your comment by mentioning first only the

W direction (blue color) and second the "the rest of the profile up to ridge height". The ridge height is also now added on the plots.

11. l. 338: "The comparison of the first level of KENDA-1/MER (Fig. 7.c)": Fig. 7c suggests that KENDA-1/MER is shown at 775 m. Which altitude is shown?

The altitudes given in Fig. 7 are right. To allow the comparison between the remote sensing and KENDA-1 at both sites, the same altitude of 775 m was chosen for all plots apart from SMN/MER.

12. l. 373: "Plots of radial winds perpendicular to the valley direction clearly present this circulation pattern both in presence of up and down valley winds around sunset (Fig. S9)." Please rephrase, e.g. Figure S9 shows ...

Done: "Fig. S9 shows radial winds perpendicular to the valley direction that clearly illustrate this circulation pattern observed in presence of both up and down valley winds around sunset."

13. l. 360 ff: "Finally, KENDA-1/MEE overestimates the influence of the synoptic winds leading to the absence of along valley wind in winter replaced by constant slow down valley winds below 1200 m and to higher up valley wind speed in spring and summer." (i) "along valley wind in winter replaced by constant slow down valley"; Do you mean "up valley wind" replaced by down valley wind (as "along valley" wind includes both, up and down valley wind)? I'm not sure if I understand the authors reasoning why the "influence of the synoptic winds" leads to down valley winds in winter and an overestimation of up valley wind in summer in KENDA? Could the authors please explain their reasoning.

The referee is right for both points. The observed low down valley wind replaces only the up-valley wind and this has nothing to do with the influence of synoptic winds, which is on the contrary visible at higher altitudes. The figure was also modified to represent only data present in both time series. The text was consequently modified: "Finally, from November bis February, KENDA-1/MEE overestimates the influence of the synoptic winds leading stronger up-valley wind presence down to 800 m and models continuous down-valley winds below 800 m with shallower diurnal cycle than observed by DWL/MEE. The foehn influence in March up to 2500 m is well modeled."

14. l. 370 "intense north-facing slope winds": Please rephrase; it is easy to confuse this with "north-facing" "slope winds" (i.e., south to north wind direction).

Yes, it is confusing. The sentence was modified:" Intense winds from north-facing slope ($>$ 25 km/h) are also observed between 1400 and 2000 m during some hours around sunset with a much lower intensity in May."

15. l. 389 ff: Please indicate where this information is shown.

These sentences were added to shortly describe the content of the section to streamline the manuscript. The additional wind observations in the Haslital are described with Fig. 10 and the one in the Sarneraatal in Fig. S8. The differences of the wind system at MEE, MER and the entire valley volume refer to the analysis bounded to Fig. 10. We do agree that the mention " to the entire valley volume" is exaggerated and now we only mention from the lake of Brienz to MER.

16. .. 391: Please correct ”SM/MER”.

Done

17. l. 451: ”Note that the KENDA-1/MER is in better agreement than KENDA-1/MEE with SMN/MER (not shown), which can indicate significant differences in the foehn influence at the two stations.” (i) ”not shown” Isn’t this information shown in Fig. 11? (ii) Can the difference also be related to differences in locations (as argued for MWR/MEE above)?

i) Yes, Fig. 11 allows to see this affirmation so that “not shown” was replaced by Fig. 11.

ii) Yes, we could consider that about 1° difference between KENDA-1/MEE and KENDA-1/MER can be attributed to the difference in location as for observations. Fig. 11a however shows that the difference is much less systematic than between MWR/MEE and SMN/MER. This is then a supposition that we prefer not to discuss in the paper.

18. l. 480 ff: I appreciate the summary paragraph. Out of curiosity, do the authors have any hypotheses about the reasons for KENDA wind speed overestimation and simultaneous temperature underestimation?

As probably guessed by the reviewer, we do not have hypotheses about the simultaneous wind speed overestimation and T underestimation by KENDA. We tried to figure out thermodynamic solutions, but without success.

19. l. 490: ”such wind speeds difference is subject to a discussion about a potential large overestimation of the winds at this location”: Do you here refer to an overestimation specifically during foehn events or during all conditions?

We are only referring to foehn events, which is now specified in the sentence.

20. l. 592 ff: Please replace ”daily cycle” by ”diurnal cycle”.

Done

21. l. 610: ”the NWP”: Please rephrase.

Done

22. Figure 1: I appreciate the revised map. I would suggest to increase the size of Fig. 1b, and would find it more intuitive if the x-axis were reversed to match panels a and c. In addition, I would find it helpful if the caption would indicate colors for up valley and down valley wind, respevtively (e.g. up valley wind (red), etc.). Similarly for slope winds.

The requested modifications were made in Figure 1. The caption was adapted to give the colors for along and cross valley winds: “The two cells of the model used are pink. Arrows representing up/down valley winds and north-facing/south-facing slope winds are colored respectively in red/blue.”

23. Figure captions: I would appreciate if the authors could revise figure captions (e.g. Fig. 3, 4, 7) and make sure to include the altitude of data which is shown.

The altitude of the data was added either in the figure or in the figure caption.

24. Figs. 8 and 9: Are the same sub-sets of dates/data points used in DWL and KENDA profiles (i.e., are KENDA data points removed from the analysis when no DWL observations are available)? It looks like KENDA includes more data points. In contrast, in Fig. 6 the NAN grid points appear to match.

You're right and we modified Fig. 8 and 9 so that only data present in both time series are now plotted.

25. Fig. 12: Please correct date and time in all panels.

Done. Sorry for the typo.

26. Fig. 12b: Please correct the colorbar labeling.

Done. Sorry for the typo.

27. Fig. S11b,c and Fig. S12: I would ask the authors to increase the label sizes.

Done

28. Supplement: I think Figs. S7 and S10 are not referenced in the manuscript. If they are relevant, please include a reference in the manuscript.

Figs. S7 and S10 were suppressed.

29. Title: Generally, abbreviations such as ”NWP” are avoided in the title. Moreover, did the authors consider adding ”Switzerland” in the title, as Meiringen is rather unknown?

NWP was spelled and we added the world alpine to situate the geographical area.

Answers to the reviewer 3 comments

Review of ’Comparison of temperature and wind between ground-based remote sensing observations and NWP model profiles in complex topography: the Meiringen campaign’ by Bugnard et al.

The study investigates wind and temperature in the boundary layer of the Haslital in Switzerland using ground-based remote sensing and in situ instruments deployed during the Meiringen campaign from October 2021 to August 2022. It compares the observations of a microwave radiometer, Doppler lidar, and surface meteorological station to the COSMO-1E model analysis (KENDA-1). The valley is rather narrow with 1.5 km width and one of the sites is affected by the low altitude Br¨unigpass to the north. By means of monthly composites, nighttime inversion and thermally driven wind systems were detected. Large model errors were found for nighttime temperatures on the average as well as on individual days.

This review is of the revised submission of the manuscript and is provided in view of the interactive public discussion. In the revised version, many of the comments of previous reviewers are addressed. For example, the manuscript was shortened and the appendix was moved to a supplemental. The gathered data provide a great opportunity to investigate the boundary layer conditions in a mid-sized Alpine valley and to evaluate the performance of KENDA-1. In my opinion, the issues with partial data assimilation at MER is of particular interest and may be relevant for other locations as well, leading to potential forecast improvements in complex terrain. While I believe that the manuscript is improved compared to the original version, I still have some major concerns about some of the aspects of the study, some new, some already raised by the previous reviewers. For example, the authors strongly focus on thermally driven flows, however, they don't distinguish between thermally and synoptically driven conditions when computing the monthly composites, which, in my opinion, masks many of the relevant features. KENDA-1 has a 1.1 km horizontal resolution, while the Haslital is 1.5 km wide. The authors discuss the difference in valley floor height between reality and the model, but they do not discuss the terrain in the model at all, such as shape and depth of the valley, and to what extent the Br¨unigpass is resolved or how many grid points are available in the valley. I provide more details on these two aspects as well as many others in my comments below. Overall, I think the manuscripts includes too much description and not enough analysis in its present form. I suggest that the authors consider my comments before the manuscript can be accepted for publication.

1 General comments

1. The manuscripts contains a very detailed description of conditions in the results section, but hardly any investigation of the relevant processes. It first presents a lengthy description of the temperature and wind field using monthly composites with lots of details on values in specific layers etc, followed by a description of a three-day case study when the flow from the Br¨unigpass affected the conditions in the Haslital and by a description of three foehn events. In my opinion, it currently is a mix between a campaign overview and some preliminary results. It is very descriptive without a clear story line. If the case studies are meant as teasers (for future manuscripts) they could be much shorter. No in-depth analysis of the case studies is presented and there are many open question, which arise to the reader. For example, why should a warm upvalley wind from the Sarneraatal descend into the Haslital? An in-depth analysis would probably be beyond the scope of the manuscript, but such contradictions to theory and open question should at least be mentioned. Also, I don't think that there is enough observational proof presented to conclude that the wind conditions in the Haslital are affected by the flow from the Br¨unigpass or that there is a vortex present. These findings should be formulated as hypotheses. A more detailed study, possibly using the 3D model output is needed to provide strong evidence for this and to investigate the physical processes.

> *As mentioned by the reviewer, an in-depth analysis of all the case studies presented in this manuscript is beyond the scope of this publication. Here we aim at providing a board overview of the complexity of wind regimes of a narrow Alpine valley and providing insight into how well this complexity is captured by an operationally used NWP model. Thus, we focus on highlighting challenges of the meteorological model and identifying potential (highly localized) phenomena which in future should be further assessed (e.g., during TEAMX). Thus, we do not claim to provide a complete explanation of all phenomena observed, but rather raise awareness for such local scale phenomena that*

might cause difficulties to operational weather prediction models and thus serve as a baseline for future measurement setups.

We agree that for both examples highlighted by the reviewer (upvalley flow from the Sarnertal and the influence of the wind regime in the Meiringen valley), the observational proof that can be underlined with measured wind lidar data is limited. This is a consequence of the setup used in this study. For a 3-dimensional representation of vortices, multiple systems (scanning wind lidars) need to be available to perform multiple doppler analyses that can yield estimations of 3-dimensional wind field information in addition to the orthogonal wind components. However, such 3D data is not available for the campaign presented here. But personal observations on site and various discussions with local farmers provide additional evidence for the hypothesis that the up valley wind from the Sarneraatal influences the wind regime in the Haslital. This condition is met if the synoptic condition is related to a Bisenlage, which corresponds to large scale north to north easterly winds. During such conditions, the wind in the Sarneraatal is directed from NNE to SSW, as indicated by the SMN measurements for the 22.11.2021 in the figure below.

As explained in the manuscript and used in this study, three meteorological stations are available for wind observations in the Haslital. One is located close to the village of Meiringen (MER, blue) and thus situated in the upvalley, i.e. to the east of the measurement setup during the campaign. A second station is located close to the runway of the airport and the measurement setup during the campaign (MEE. black), a third station is located to the west close to the lake of Brienz (BRZ, red). The arrows indicate the dominant wind direction during the period from 9:30 to 12:30 UTC (during which all the station have wind speeds of >1.5m/s). In the Hasli valley the wind direction is clearly separated with easterly wind in the western part (MEE & BRZ) and westerly winds in the eastern part (MER). Similarly in the snap shots of the RHI scans of the wind lidar, the regime difference is obvious in the along valley direction (left plot). In the direction of BRZ the radial wind are constantly directed away (red colors) whereas to east (MER) a low level radial winds are generally directed towards the instrument, whereas at slightly higher altitudes the radial wind has the opposite sign (directed away from the instrument). Unfortunately, the range is limited to 4km and thus the station MER (located at 5km) distance is not covered. Nevertheless, from this we conclude that there is an effect of the wind coming from the Brünig during these specific synoptic situations (Bisenlage). In addition, the RHI directed towards the Brünig pass (right plot in the lower figure) shows a distinct pattern of radial winds at low elevations along the RHI transect (from N to S) and a 2D vortex signature above the measurement location (indicated by the arrows), that corresponds to the middle of the valley. Such a structure in the two orthogonal planes of radial velocity do indicate a vortex like structure. Nevertheless, without a 3-dimensional observational dataset (e.g. by multi-Doppler analyses), a detailed analysis of the vortex structure is not straight forward and thus we agree with the reviewer's comment that this is not a proof of, but rather a hint for a vortex structure. However, this illustrates the complexity of wind regimes in narrow valleys in complex terrain and we believe it is worth to be mentioned this in the current manuscript to make other campaigns aware of such highly local phenomena, based on which the strategy of future campaigns could be better defined.

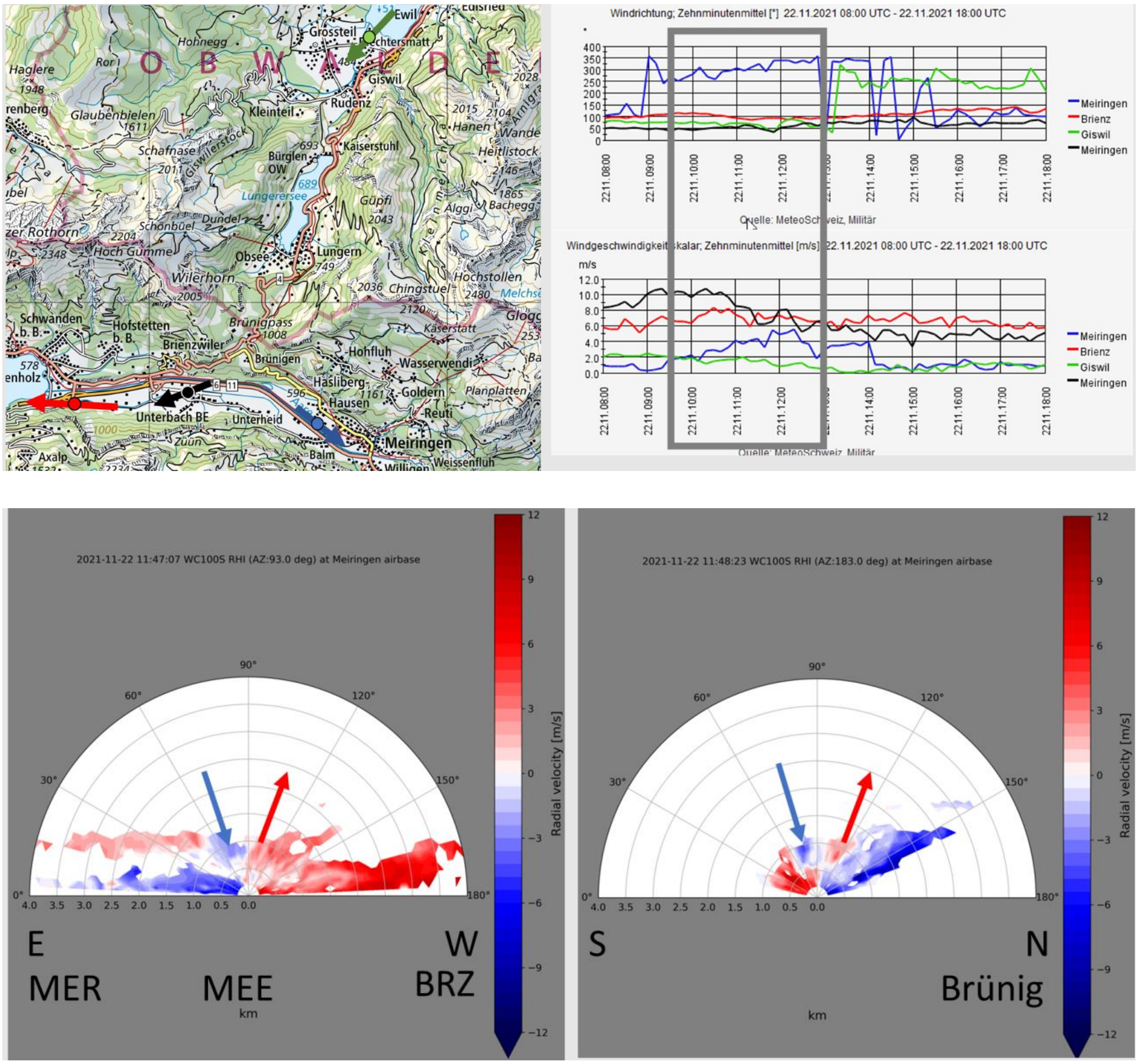


Fig. 1: a) map of the valley with wind direction during the event, b) wind direction and speed at MER, MEE and BRZ and c) radial wind compounds measured by DWL/MEE.

Furthermore, we investigate the pressure difference between Giswyl (GIH) in the Sarneraatal and MER in the Haslital. We used the pressure reduced at sea level to get rid of the altitude difference. Fig. 2 clearly shows that the mean monthly reduced pressure at GIH is higher than at MER for all months but the winter months. The highest difference (>1 hPa) is observed from mid-day and persists until the late afternoon or *even the early evening. The difference between potential T measured by MWR/MEE at the altitude of BRU (1010 m) and at BRU is also positive. Air masses from the Saneraatal are then colder than air masses in the Haslital not only in case of biselage. Such a phenomenon can be explained by the valley volume effect since the volume of the Sarneraatal is 1.7 time bigger than that of the Haslital. The heating of the air masses occurs then more slowly than in the Sarneraatal and induce the observed lower T. The up valley wind passing the Brünig Pass will thermodynamically tend to fall into the Haslital at MEE. This phenomenon is enhanced in the afternoon when the up-valley wind in the Sarneraatal is the strongest but also happens sometimes in the morning. This corresponds to the DWL/MEE measurements and seems to be well modeled by KENDA-1/MEE.*

A figure will be added to the manuscript and this phenomenon is now explained in section 3.3.

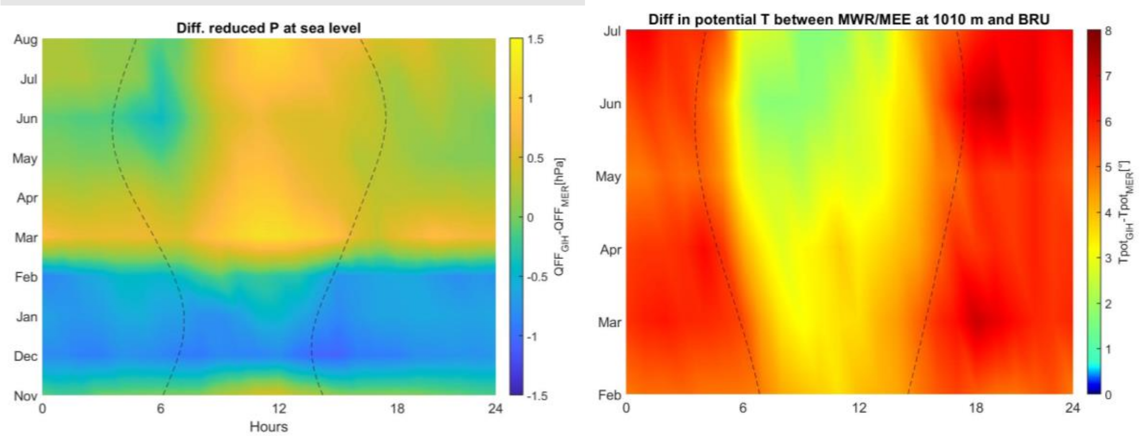
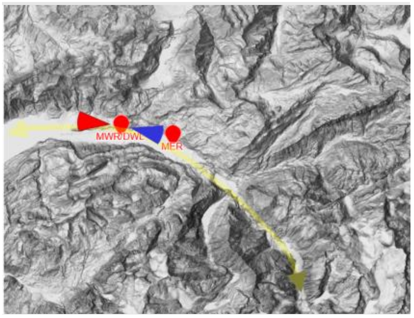


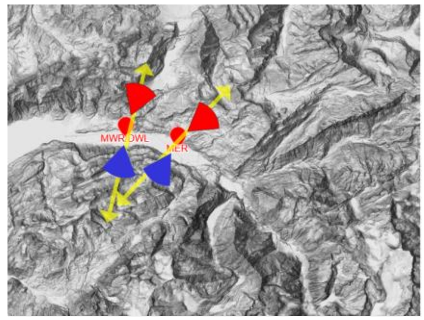
Fig. 2: a) Monthly diurnal cycle of the difference in pressure reduced at sea level between GIH and MER, b) Monthly diurnal cycle of the difference in potential temperature at 1010 m above MWR/MEE and at BRU and c) Monthly diurnal cycle of the difference in T at 1010 m above MWR/MEE and at BRU.

2. The analysis focuses very much on thermally driven flows, but the monthly composites are not separated for days that are dominated by large-scale conditions (frontal passages, foehn) or are affected by clouds (unfavorable for thermally driven flows) or are dominated by fair weather conditions (favorable for thermally driven flows). Computing composites over all days masks the signal of thermally driven flows (which primarily occur under fair-weather conditions, as correctly mentioned in l. 48). The authors still interpret the weak signals in wind in the composites and conclude that e.g. no thermally driven winds were observed in December and January. However, they might still be there just not in the monthly composites. Furthermore, including everything in the averages makes the comparison to other studies (Discussion section), in which days where filtered for thermally driven flows, not very meaningful. For example, conclusions on thermally driven flows are drawn from the composites of along valley wind component in Figs. 7 and 8. There are several features that are not typical at all for thermally driven flows and differ from theory, like the persistence of upvalley flow near ridge height and the decrease in downvalley wind strength during the night. This could be a result from the projection of the

flow on along valley wind direction or a result of sampling over all types of conditions. Either way this discrepancies need to be discussed and attempted to be explained.

First, the use of the word "projection" to describe how the along and across valley wind speeds were computed is misleading and we apologized for this misinterpretation. In fact, along valley wind are selected in a +-15° around the valley axis and cross valley wind in a +- 30° around the perpendicular to the valley axis. This is now described in the experimental section.





The described results are then not a result of the projection since, e.g. along valley does not contain any influence from the N-NE wind from the Brünig Pass.

Second, as answered thereafter (p. 11 of this document), a selection of very good (not shown) or good days (Fig. 3 this document) does not modify the main feature of the thermally induced valley winds. We think then the monthly composite still allows us to describe the main wind feature (first goal of the study) and to compare the modeled with the observed data (second goal). Finally, a comparison of the results of this study with other studies seems to us still worth, even if the selected weather conditions, the instrumentation and period of the year are not always identical.

Monthly plots for wind direction are separated using a wind speed threshold which seems a bit arbitrary and needs more justification. It is currently not clear at what height this threshold must be met and during what time period. Also, why do the authors not distinguish days for the other composites (along-valley and across-valley wind speed, temperature) as well? Since the focus is on thermally driven flows, the analysis could also be restricted to composites of days with thermally driven flows.

This would reduce the number of panels and streamline the manuscript. A better and more physical way to distinguish days with thermally driven flows from days dominated by synoptic flows could be by looking for a wind direction reversal twice per day and/or by considering cloudiness.

The arbitrary threshold of 20 km/h was indeed also questioned by the two first reviewers. We then decided to modify Fig. 6 and to plot and discuss first the monthly wind direction without wind speed threshold. As proposed by the third reviewer, plots of monthly wind direction for good weather conditions are now also available in the supplement. They were not inserted in the manuscript since the method to select days with thermally driven flows relays only on cloud coverage at the nearest station with this parameter (FRU). We also think that the criteria allowing thermal winds should however be studied before to be used. Such a study is beyond the scope of this paper, but the following questions could be adressed: which cloud cover during which period impedes thermally driven flows to occur? What is the influence of the cloud cover *during the night? Should we also include the influence of ground-based T inversions? Are these factors influenced by the seasons (i.e. the mean T or the snow cover)? Similar questions (e.g. time and altitude of wind direction reversal) are also raised concerning an automatic detection of thermally driven winds by wind direction reversal, particularly in the described complex terrain presented in this study. It has also to be considered that a selection of only "very clear days" misses information about thermally driven wind in "clear days". The provided new figures provided in the manuscript and in the supplement allow to answer some questions raised by the reviewer, but a complete analysis of the occurrence of thermal valley winds as a function of different synoptic situations, different cloud amount is beyond the scope of this study.*

Sections 3.2 and 3.2.1 were consequently modified to describe the new Fig. 6.

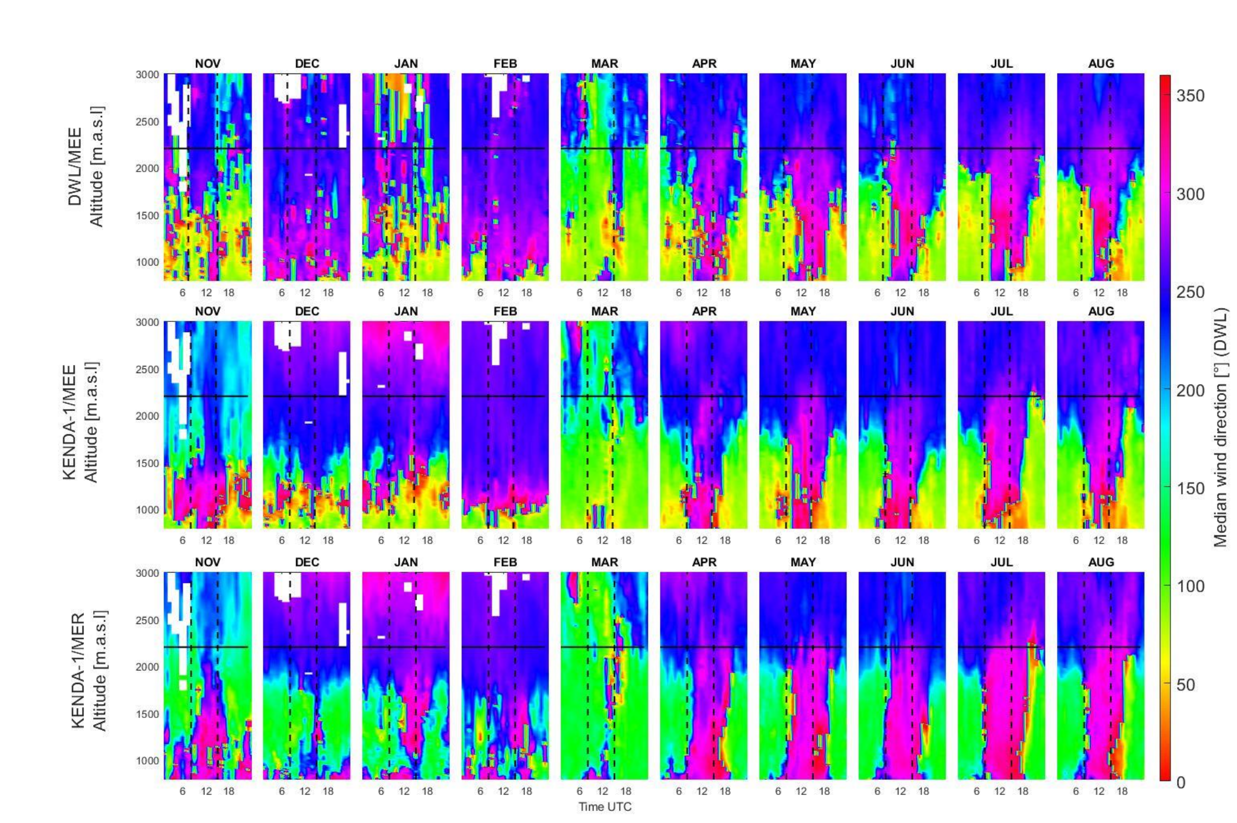


Fig. 3 (new Fig. 6): Monthly median wind direction [°] for a) DWL/MEE, b) KENDA-1/MEE and c) KENDA-1/MER (1.11.2001-23.08.2022).

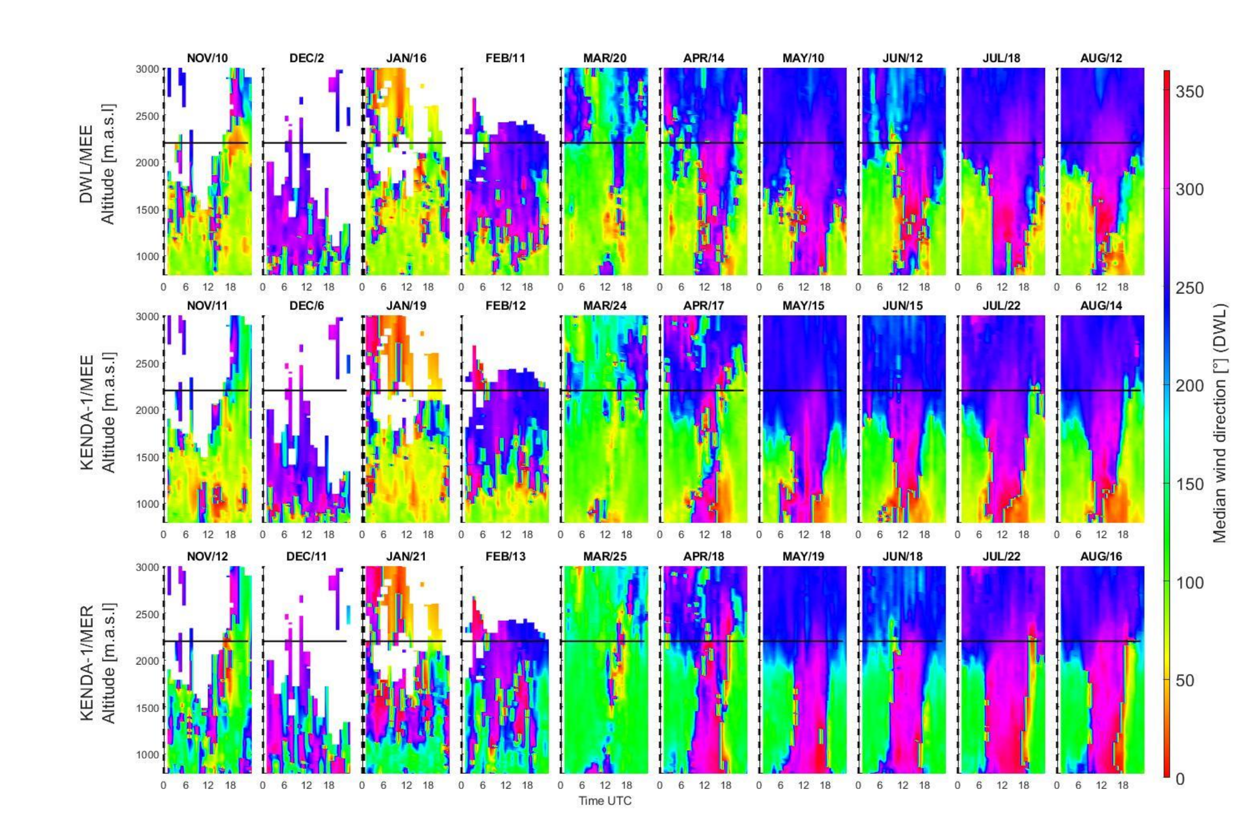


Fig. 4 (inserted in the supplement): Monthly median wind direction [°] for a) DWL/MEE, b) KENDA-1/MEE and c) KENDA-1/MER (1.11.2001-23.08.2022) for clear weather days. The clear weather days are determined by less than 5 octas of cloud cover measured at the SMN station of Frutigen.

3. Composite plots for wind (Figs. 6, 8, 9) are presented with 10 panels per row, which makes it impossible to see any details on the time axis and to follow the detailed description for specific hours. I highly suggest to think about another way of presenting these composites plots. I understand the desire to reduce the number of figures, but this should not be done at the cost of visibility.

Figure 6 was modified (see previous answer) and comprises now only three rows allowing more space for each pannel. The space between the plots was also reduced and we hope that this improves the global readability of the pannels.

4. A discussion of the model terrain is needed, which goes beyond the difference in valley floor height. For the capability of the model to simulate terrain induced features, the shape and depth of the valley and the numbers of grid points is most relevant. This is not discussed at all. The Haslital is 1.5 km wide and KENDA-1 has a 1.1 km resolution. What about the Sarnaraatal? Is the Br¨unigpass even resolved in the model? How are the lakes resolved? How much terrain smoothing is done? In l. 74, it is stated that the grid resolution should be about 10 to 20 times higher than the relevant topographic scale to fully capture the different exchange processes. This aspect is very important when interpreting KENDA-1 results, but is currently not considered at all.

The most important information on the model terrain such as the ridge heights, the height of the Brünning Pass and the position of the cell containing the MER and MEE were already given in the manuscript. Some points are further discussed as answer to the comment 12 and corresponding sentences have been added to the methods and discussion sections.

5. Some sentences are not very clear and perhaps the authors should consider using a professional editing service to remove these language issues. Examples include, l. 202-203: Layer with higher T develops gradually from sunset to sunrise to reach monthly-related maximal T and height. l. 250-251: Globally, the measured MWR/MEE first level T are closer to the SMN/MER T than the modeled T. l. 256-257: The analysis of the negative ground T difference between MER at 590 m and BRU at 998 m (horizontal distance = 3.7 km) shows that near ground T inversions are common during the night for all months in the study. l. 269-271: The missed T inversions by KENDA-1/MEE leads to both its important overestimation of the T at ground level (Fig. 4) and its slight T underestimation between 850-1200 m (Fig. S5 for detailed examples). l. 296-297: ..., whereas flows from W-NW are measured in the rest of the profile concerned by up valley winds (see further explanation in sect. 3.3). l. 303-305: The good KENDA-1/MEE performances comprise first the influence of the foehn up to 2500 m (ws >20 km/h) as well as the presence of valley wind below 1200 m (ws <20 km/h) in March. l. 345-347: Thermally induced wind height increases with temperature, reaching 1000 m in February, 1800 m in May and up to 2000-2200 m in July and August. l: 371-373: This suggests a circular motion with North updraft winds (median vertical velocity of 1 km/h) that cross the valley at a low altitude, rise against the north facing slope and come back at higher altitude with a South downdraft component l: 599-600: This is especially the case at the end of March, when enhanced night time radiative cooling and important global solar radiation form strong inversions.

A competent English writing person corrected the manuscript. All the mentioned sentences were modified.

2 Specific comments

1. l. 64: The classic work of Whiteman and Doran (1993) could be referenced here.

The Whiteman and Doran (1993) reference was introduced with a reference to the forced and the pressure-driven channeling mechanism.

2. l. 73-74: The classic work of Skamarock (2004) Skamarock and Klemp (2008) could be referenced here.

The citations have been inserted in the text.

3. l. 84-87: Very complicated sentence. Please rephrase.

Done

4. l. 88: Is the ”first objective of the campaign to study the seasonal and diurnal cycles” or is this the first objective of this study?

Right, this is the first objective of the study and not of the whole campaign comprising instruments (Radar and ceilometer) that are not used in this analysis. The sentence was consequently modified.

5. l. 93: I think a short introduction on what KENDA-1 is and why its evaluation is important needs to be added to the introduction (possibly where NWP models are mentioned).

An introductory sentence has been added.

6. l. 94: I find the acronyms MER for Meiringen and MEE for Unterbach a bit unfortunate, since they are very similar and easy to mix. Maybe just a matter of taste.

The acronyms were not chosen by the authors in the context of this study but are defined by the various networks in Switzerland. Even if more distinct acronyms would be helpful for the readers, the designation of the stations by their usual acronyms is a priority.

7. l. 104: Why are the times not given in local time? This is advantageous for a study that focuses on thermally driven diurnal wind systems. At a minimum the time difference between local time and UTC needs to be given.

UTC time was used since it corresponds to time used in our databank. The difference with local CET time is one hour (CET=UTC+1) and the difference between UTC and solar time at Meiringen (longitude= 8.1909°E) is about 40 minutes. These small differences have no impact on the comprehension of the diurnal cycles and sunrise and sunset times are most of the time represented in the figures. The difference between CET and UTC is now given in the manuscript: "Local time corresponds to Central European Time (CET), which is one hour ahead of UTC time (UTC+1)."

8. l. 105-106: The temporal resolution of the observations is higher than 1 hour. How is that considered? Are these values averaged to 1h values before computing monthly composites or are instantaneous hourly values used? When computing the composites how is data availability considered? Are model data only plotted for times and heights where observational data are available?

The measured data are first aggregated into hourly values, which are then used to resolve the diurnal cycles of the figures. Only time with observations are reposted in Fig. 6, whereas all data from KENDA-1 are reported in Figs 8 and 9. This is now corrected and the text was consequently adapted.

9. l. 110-112: The Gadmertal and Rychenbachtal are not shown in the map and not relevant for the study. Please remove or modify the map in Fig. 1 to include them. This was already a comment of a previous reviewer.

The mentions of the Gadmertal and Rychenbachtal have been deleted in the revised version of the manuscript.

10. Fig. 1: Fig. 1b is impossible to see. Please increase line width, axis labels, and legend to make it readable. I suggest flipping the x-axis, so that the location of the sites is consistent with 1a and 1c. Why are the dots in Fig. 1b not at the valley floor? Are the station heights in Fig. 1b from the model? In Fig. 1c, labels along VW and across VW are very hard to read.

Fig. 1 was modified as requested by both reviewers.

11. Section 2.2: This section sounds like COSMO-1E is used:" The NWP model used in the study is the limited-area non-hydrostatic atmospheric model from the Consortium for Small-Scale Modeling Model

(COSMO)”. I think it would be better to make clear from the beginning that the analysis KENDA-1 is used, which uses 1-h forecasts from COSMO-1D and observations. It would also be helpful to mention here how close to the investigation area the different observations are assimilated. SMN is assimilated at MER (this should already be mentioned here). But what about the sondes and profilers? How far away are they from the investigation area?

The manuscript has been adapted accordingly.

The distance between Meiringen and the assimilated radio-sonding at Payerne is of 94 km, whereas the distances to the three assimilated radar wind profilers are comprised between 75 and 110 km. It must be noted that all these profiling observations are situated on the Swiss plateau whereas Meiringen is in the Alps.

12. l. 141ff: Where are the MER and MEE model grid cells located in the model terrain? Are they at the valley floor in the model terrain? This is more meaningful than comparing the grid cells to the real terrain, which the model does not know.

As can be seen in the next figure, both MEE and MER are located at the valley floor in KENDA-1 DEM, so that the thermal valley wind system should be modeled correctly at the grid cells comprising MEE and MER. The narrow valley floor of about 1.5 km corresponds to one grid cell in KENDA-1 DEM at MEE and MER and enlarges to two grid cells 2 km after MEE in the vicinity of the Lac of Brienz.

The following sentence was added to the manuscript: “In the modeled terrain, both MEE and MER stations are situated in the cell grid corresponding to the valley floor.”

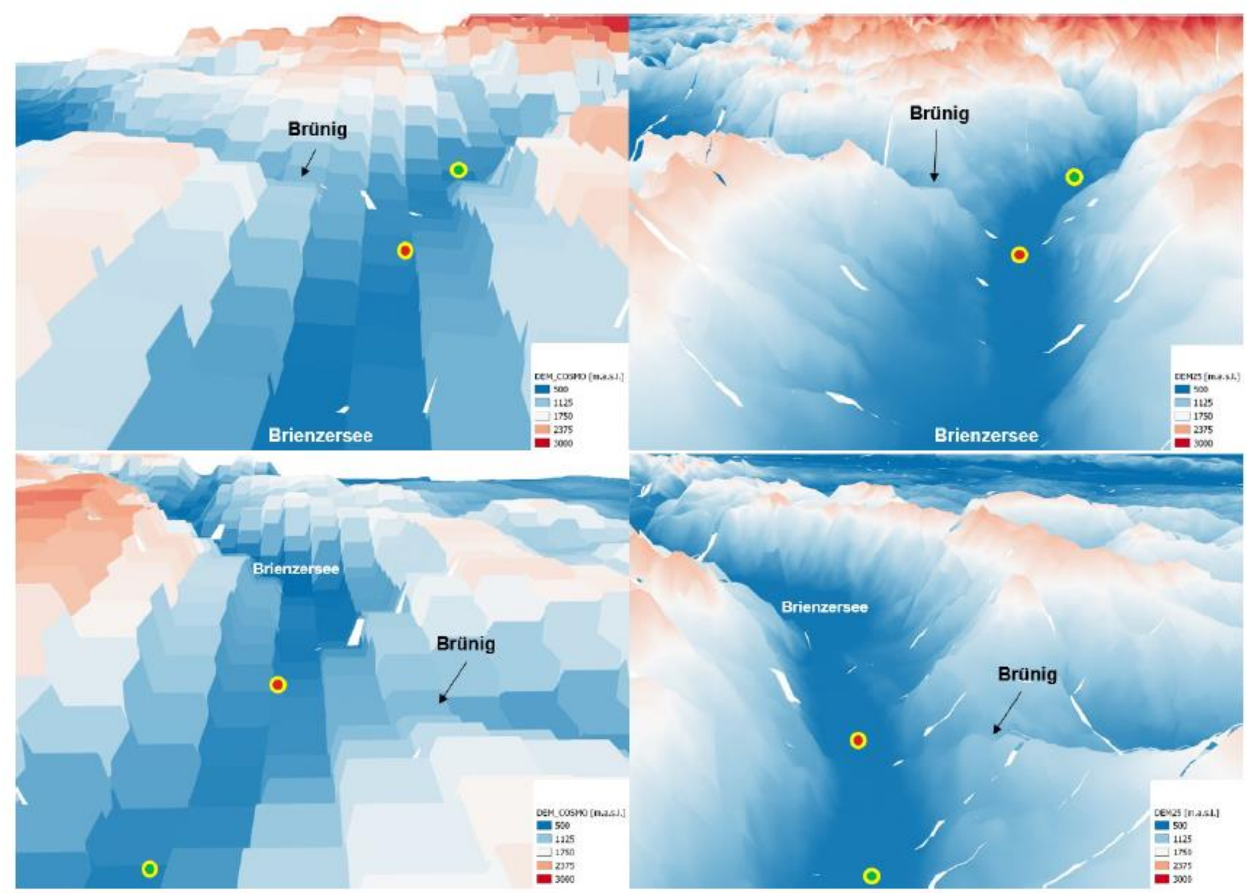


Figure 5: Comparison between KENDA-1 DEM (left) and the DEM25based on the 1:25'000 Swiss national map for the Haslital valley towards Lake Brienz to the west (bottom) and up the valley (top) to the east. The MEE and MER are represented by the red and green dots, respectively.

13. l. 150ff: I think the fact that MER data are assimilated in KENDA-1 is critical for the results and needs more attention. If the observations are assimilated during certains conditions (daytime) and not during others (nighttime inversions) this may affect the error distribution. In the response to reviewer 2, the authors included an example demonstrating how often MER data are rejected and argue that 'One of the reasons to compare observed and modeled data at MER is that the MER ground observations are only assimilated if the difference with the modeled data is inside a given threshold." This information as well as the threshold would be helpful to include in the manuscript. Furthermore, I think including statistics on how often MER data are assimilated during the whole campaign and at what times would be most helpful and interesting.

We added a description of the first guess check and a statistical evaluation of how often and on what daytime the rejections occurred.

14. Section 2.3.2.0: Was the RPG neural network retrieval used? As far as I know, HATPRO-G5 comes with a surface met sensor. Are this information used in the retrieval?

Yes RPG retrieval was used and the surface measurements are used in the retrieval.

If not, this surface met measurements at MEE could be an additional observational source to evaluate KENDA-1. Why was this not used?

Indeed, the Radiometer surface measurements could be used to evaluate Kenda 1 but they were not considered during this study.

What does 'line of sight of about 10 km' mean? Is this the line of sight in the direction of the low elevation scans? In which direction was this performed (up or down the valley)?

The line of sight of about 10 km means that there was no obstacle in the 10 km distance from the instrument in the direction of the low elevation scans. It was performed down the valley.

Temperature biases in microwave radiometer retrieval are highly instrument dependent, arising from spectral biases or liquid nitrogen calibration. The cold bias found by Hervo et al. (2021) is not necessarily transferable.

Indeed, the conclusion of the report might not be transferable. All mention to this report and the associated conclusions were then removed.

'The instrument at MER ..' isn't the MWR installed at MEE?

MER was replaced by MEE

15. Section 2.3.3: The scan strategy of the Doppler lidar is not clear. It sounds like the lidar did a combination of vertical stare (every 10 min for two minutes) and DBS (every 5 min). What did it do in between? The scan in Fig. S9 looks like it is an RHI scan, this needs to be mentioned here.

No time is left since the configuration of the Lidar for a 10-minute period was the following:

- *Starting with DBS for 2 minutes*
- *Fixed vertical scan for 2 minutes*
- *DBS for 3 minutes*
- *RHI parallel and perpendicular to the valley for 2 minutes*
- *DBS 1 minute*

The manuscript was consequently modified to be more explicit: " There are three measurement modes: 120 second zenith scans performed each 10 min to measure vertical wind speed, Range Height Indicator (RHI) scans for two minutes every 10 minutes to measure radial wind speed along and perpendicular to the valley (not used in this study). The rest of the time the instrument was measuring in Doppler Beam Switching (DBS) scans providing 7 independent wind profiles every 5 min to measure horizontal wind speed. In this analysis the wind profiles were averaged for each 5 minutes interval."

According to the Vaisala data sheet the maximum range for Wincube 100S is 3 km. Was this system modified to reach 12 km maximum?

The " Typical maximum operational range" provided by vaisala is limited by the presence of target. It is usually 3km but in case of elevated aerosol layers or clouds the 100S can measure higher. The data sheet also mentions the "Max acquisition range" that is around 12km for 100S.

16. l. 195: Heavy liquid precipitation ?

Yes, it is now specified in the text that the precipitation was liquid one.

17. l. 189ff: None of the information on the general conditions during the campaign are considered when presenting the results on temperature and wind. How do the boundary layer conditions differ over snow covered vs. snow free ground? Does KENDA-1 performance depend on snow cover? Because of the monthly composite and snow cover lasting until mid-December this is all mixed together. How do the heat waves reflect in boundary layer conditions and KENDA-1 performance? If the information on general conditions is included, it should be considered when desribing the results.

The general conditions are now referenced each time they allow to explain the results in the results and the discussion. As explained thereafter, a complete analysis and description of the ABLH is beyond the sclope of this paper.

18. l. 202-209: The description of temperature changes should be related to the boundary layer evolution. 'Layer with higher T' is the daytime boundary layer. Perhaps the authors could compute convective boundary layer height using the parcel method. I do not see the value of presenting temporal T gradients (Fig. S4), especially not for monthly composites. It is not surprising that T increases during the day and decreases in the evening. I suggest removing this to streamline the manuscript. In Fig. S4 temporal gradients are shown, in the text vertical T gradients are discussed (C/km).

The height of the convective boundary layer (CBLH) clearly depends on the T profiles described in Fig. 2 and § 3.1.1. CBLH computed by both the Parcel and the Bulk Richardson methods were computed from the MWR/MEE T and DWL/MEE profiles. During this study, we had not time to compute the Mixing Layer height (MLH) from the aerosol backscattering profile measured by the ceilometer. A complete study of the Mountain Boundary Layer (MOB) involving all potential ABLH computed from the MWR, the DWL (using backscattering and wind information) and the ceilometer during day and night associated with estimation of the stability of the atmosphere and the cloud cover would be very valuable. This study will perhaps be done in the future but is clearly beyond the scoop of this already very long manuscript.

The evolution of T in MEE is indeed not surprising but is necessary before to evaluate KENDA-1 performance. As requested, the § was shortened: "The evolution of T in MEE from February to July (Fig. 2.a) presents as expected clear diurnal cycle with a vertical extent depending on the season. Layer with higher T develops gradually from sunset to sunrise, persists during the first half of the night and gradually fades out towards sunrise. The time of the T maximum, the persistence of the warm layer and the extent of the warm layer are all enhanced during summer months. The maximal temporal T gradient usually follows sunrise and sunset (Fig. S4) and are confined below 1500 m with values up to +5°C/h in the morning and between -4 and -6.5°C/h in the evening."

Concerning the T gradients of Fig. S4, the units of the paper were falsely attributed to °C/km instead of °C/h. The manuscript is now corrected.

19. l. 213ff: Fig. 2b show the bias. Did the authors also investigate mean absolute error or root mean square error? Is mean absolute error also small when biases are small or are the small biases related to averaging artefacts (days with very large positive biases vs days with very large negative biases)? How

are observed and modeled profiles compared? Are modeled heights interpolated to observed heights or vice versa? Since the temporal resolution of the observations is higher than KENDA-1 are they averaged before comparing to the hourly model output? This information should probably be included in Sect. 2. I assume the authors refer to the study of Hervo et al. (2021) when saying 'The cold bias between the MWR and the radio sounding could however suggests a larger error of KENDA-1.' As mentioned before, biases are instrument specific and I do not think this is a valid conclusion.

No complete analysis of the uncertainties was done.

To allow the comparison between the modeled and observed data, all profiles were linearly interpolated at a vertical resolution of 10 m. This information was missing and is now given in sect. 2.: "Finally, all profiles were linearly interpolated at a vertical resolution of 10 m to allow comparison between the observed and modeled data."

As described at line 139, Kenda-1 data corresponds to instant hourly values, while the observed data are hourly medians as described at lines 105-10. The use of hourly medians for the observation was chosen to decrease the uncertainty of the measurements at the cost of introducing a further difference between the modeled and the observed data.

To avoid any confusion, we deleted the sentence "The cold bias between the MWR and the radio sounding could however suggests a larger error of KENDA-1."

20. l. 223: Please explain why observations at MEE are compared to MER (here and later in the manuscript). These are different sites affected by different physical processes.

The set-up of the campaign did not allow us to install the REM instruments at the SMN/MER station. The best compromise was to install them at MEE that is only 4 km apart in the same valley, and on the valley floor. A set-up with all instruments at the same place would have simplified the comparisons to only in-situ with REM observations and observations with modeled data. MEE is a station in the near vicinity of MER but the topography of the valley presents also different features at both stations, e.g. the presence of the Brünig Pass at MEE. This forces us to compare 1) in-situ observation at MER with both the REM observations at MEE and the modeled data at both MEE and MER, 2) REM observations at MEE with modeled data at MEE, but also 3) modeled data at MEE and MER in order to estimate if the differences between in-situ and REM observations are due to topographical features or to the instrumentation.

21. Fig. 2: Please add ridge height to panel a). To better see stablity it would be helpful to add potential temperature isentropes. For example, observed potential temperature isentropes could be added to a) and simulated potential temperature isentropes could be added to b).

The ridge height was added to Fig. 2a. We think that the addition potential T isentropes to T profiles would complexify the plots and require a further description in the paper.

22. Fig. 3: Please enhance label and legend size.

Done.

23. l. 234-239: What is the justification to use a mean environmental lapse rate of -6.5°C/km during nighttime temperature inversions? The height difference between the model grid point and real world is only one challenge when comparing observations to model output. The terrain in the model needs to be considered as well, i.e. is the grid point at the model valley floor or on the slope? What is the valley depth, etc. (also see my general comment on this).

A mean environmental lapse rate was only used to test if the mean discrepancies in ground T estimation could be explained by the altitude differences between the stations of MEE and MER. This seemed to us an important first result, which supports all other comparisons. Obviously, the use of a mean environmental lapse makes no sense for specific cases and particularly in presence of T inversions. A sentence explaining this in the first version of the manuscript was suppressed as required by the reviewers.As described in our answer to comment 12, the grid point of both sites are at the valley floor and not on the slope. This information is now mentionned in the manuscript. Moreover, the mean ridge's height being 2200 m, the mean valley depth is 1600 m. This information is now added to §2.1. The topography of KENDA-1 is already described in the manuscript at § 2.2: "the altitude difference between the valley floor and the crests is thus reduced of several hundred meters and, in particularly, the Brünig Pass is only 200 m higher than the valley floor."

24. Fig. 4: Are the number of samples the same for all time series, that is, are only time stamps used when all observations and model output was available?

Yes, the number of samples are the same for all observed and modeled time series. This is now specified in the figure caption.

25. l. 252ff: Using ground stations as pseudo-profiles can be affected by local impacts (e.g. solar heating during daytime, slope winds), even more since BRU is located at a pass. A brief discussion on potential error sources should be included (Whiteman and Hoch, 2014). How are temperature inversion determined? Just by computing the temperature difference between the upper and lower height and detecting negative difference? Was a minimum absolute value required? The authors compare temperature difference over layers which are not exactly the same depth and which introduce additional uncertainties. Why not compute gradients per fixed height interval? It is interesting that even though March was dominated by foehn events there is a very clear diurnal cycle in inversion frequency and amplitude. This should be discussed. In Fig. S5, there is absolutely no inversion visible in the KENDA-1 profiles. This is very strange and deserves further discussion, in my opinion. Furthermore, it has to be considered that the MWR profiles also have uncertainties and the smoothed shape of the profile may lead to an overestimation of inversion amplitude.

The potential sources of error when comparing ground-bases and free atmospheric observations are now mentioned in the paper: "T inversions observed on the ground may present an offset compared to observation by remote sensing in the free atmosphere due to the formation of nightly cold and daily warm surface layers or to different insolation or soil moisture depending on location. \citep{Whiteman_2014} observed differences generally within 1°C (standard deviation= 2-3°C) and a better agreement over steep slopes and during winter. BRU is influenced at least during daytime by colder up valley wind from the Sarneraatal (\ref{heteogenity_wind_valley}), which, however, also affect MWR/MEE and SMN/MER. (...)

Colder offset in BRU during night should lead to a higher frequency of T inversions observed from ground stations data, which is not the case ."

The temperature inversion is determined by detecting negative between the upper and lower T and no minimum absolute values were required.

As answered to comment §19, all the profiles were linearly interpolated with a resolution of 10m. The layers from observed and modeled profiles have then always the same depth. In that sense, the gradients are computed per fixed height intervals as suggested by the reviewer. The linear interpolation is well adapted to T profiles even if some uncertainties are still introduced.

In March, the 10 days with foehn corresponds to 96 hours, mostly during daytime (62.5%) so that foehn does not have a large impact on the T inversion frequency during night. The number of T inversions during daytime would have been perhaps larger without foehn events. The following sentence was added to the manuscript:" The foehn influence in March occurred mostly during daytime (8.1 \% of daytime and 4.8 \% of nighttime) and had then no direct influence on the T inversion frequency."

There is no inversion in the Kenda-1 profiles presented in Fig. S5 since it is indeed an illustration of a missed T inversion by KENDA-1. I do not see how it deserves the discussion on the weakness of model to predict ground-based T inversions.

Finally, the uncertainty of the MWR profile due to the smoothing is quite difficult to quantify and its impact on the detection of the T inversion is difficult to estimate.

26. Fig. 5: I find the amplitude plots confusing. I assume that sample size is not constant during the day and some of the spikes during summer daytime in the MWR data are probably caused by averaging over very few days. Perhaps a minimum sample size should be required for showing the amplitudes. The grid lines do not fit to the tick labels on the x-axis. What is meant by '10m spaced vectors'?

Yes, the sample size varies during the day and the number of samples per day can be estimated from Fig. 5a (100%= 28-31 depending on the month). The spikes clearly correspond to only 1-2 T inversions. For coherence between Fig. 5a and Fig. 5b, we will however keep the amplitude for very small sample size.

The mention of the 10 m grid lines spacing was removed since the interpolation of all profiles to 10 m vertical resolution is now described in section 2.

27. l. 272ff: As mentioned before, I find this issue with assimilation very interesting. Do the large differences between the observations and the 1h forecasts mean that there is no inversion in the model? What about later forecast hours? Is this maybe related to spin up time?

Yes,T inversions are often missed by the model from May to August, while the T (see Fig. 5a). The explanation is now better described: "From November to January, KENDA-1/MEE detect most of the near-ground T inversions, which last all day long in winter. Their amplitude is, however, always underestimated by 1-2°C (Fig. 5.b) by KENDA-1/MEE. From February to August, the presence of T inversion at the end of the night and in the first hours after sunrise is often underestimated by KENDA-1/MEE, which can impact the onset time of up valley winds (section Along_Valley_winds). The missed T inversions by KENDA-1/MEE lead to both its important *overestimation of the T at ground level (Fig. 4) and its slight T underestimation between ~850-1200 m (Fig. S5 for detailed examples)."*

We didn't analyze the later forecast hours.

28. l. 288-289: Wind speed values at which height, time, and station are used to distinguish between days dominated by thermally driven and synoptically driven flows? As mentioned in my general comment, it would be more physical to inspect days for the typical reversal of wind direction twice a day instead of looking at thresholds.

The 20 km/h threshold is defined for each height and time from the DWL/MEE time series and then applied both the observed and modeled data. The same amount of data is then used for the MWR/MEE, KENDA-1/MEE and KENDA-1/MER. This applied threshold is arbitrary as explained as an answer to the general comment.

As specified as answer to the general comments, Fig. 6 was modified, and the arbitrary threshold of 20 km/h is no longer used. A new figure with a selection of days with fear weather was also added to the supplement. We refer the readers to the answer to the second general comment.

29. l. 295-297: Do the authors have an explanation why the upvalley wind shifts with height? Could this be a result of averaging over a variety of conditions?

As further explained in section 3.3, the shift in the wind direction above 1000 m is probably a direct influence of the NE wind from the Brünigpass at 1000 m, namely 400 m above the valley floor. This NE wind is observed by the DWL/MEE for all weather conditions (new Fig. 6a= Fig. 3 of this document) and for fair-weather days (Fig. 4 of this document), modeled by KENDA-1/MEE but below 1000 m since the Brünigpass is only 200 m above the valley floor in the model terrain. Surface measurements show that this NE wind can even suppress the up-valley wind at Brienz (Fig. 10 a). We then do not think that it is an effect of averaging over a variety of conditions.

30. l. 297ff: What is the purpose of providing this amount of detail on synoptic flows? Consider removing or shortening to streamline the manuscript.

Since the representation with a wind speed threshold was abandoned, this sentence was removed.

31. Fig. 6: Please add ridge height to the plots. It is really hard to see anything in the panels. Consider limiting y-axis to 2500 m (like for T plots) and rearranging the figures (maybe remove synoptically driven days). What is the temporal resolution of the DWL plots? The same as for the model? Were the winds averaged to 1 hour before computing the monthly composites? What is the point of showing panel c? It is not discussed here.

Fig. 6 was modified so that it comprises now only 3 rows. The ridges' height was inserted. Fig. 6c with KENDA-1/MER is discussed in § 3.3 (heterogeneity of wind pattern in the Haslital) but is needed here to allow the comparison with KENDA-1/MEE.

The temporal resolution of DWL is 1 hour corresponding the average (sect 2 line 105) of raw data each 5 minutes (Sect 2.3.3, line 176). KENDA-1 produces only one data per hour.

32. l. 307ff: I don't see a good agreement in November. Also there are quite large difference between the observations and the model in summer/spring. The features described here (e.g. N flows from Br¨unigpass) are hardly visible in Fig. 6b in its current presentation. The conclusion 'This feature is mostly caused by the KENDA-1/MEE cell overlapping the slope towards the Br¨unig Pass so that winds at the junction between Haslital and Sarneraatal can influence the median modeled wind compounds.' is not valid in my opinion without a detailed analysis of the model terrain.

> *Fig. 6 and his description were modified. Considering all wind speeds, the agreement between KENDA-1/MEE and DWL/MEE is poor as described now in section 3.2.1. Since Fig. 6 has now only 3 rows, the N flows in red is visible in Fig. 6b.*
>
> *The reviewer is right, the conclusion if not well formulated and we modified it: "This feature is caused by the lower altitude difference between the topography (400 m) and the model terrain (200 m) and a smaller horizontal distance due to the 1.1 km cells (see Sect. 2.2)."*

33. l. 324-325: Typically, downvalley winds gain in strength throughout the night since the driving horizontal temperature and pressure gradients strengthen. Do the authors have any explanations why this is different here? Could this be a result of clouds forming during the night or a sampling issue? Please discuss this contradiction to theory.

> *The Haslital is a medium size valley and I do not know the usual used terrain (valley width, depth, length and slope, bending, tributaries, etc) involved in theoretical studies. Anyhow the occurrence of a maximum in down-valley speed 2-3 hours after sunset in not only measured by DWL/MEE, but also modeled by KENDA-1/MEE. It is also interesting to note that, at MER, a constant down-valley wind speed during night is both observed by SMN/MER and modeled by KENDA-1/MER at the ground. Fig. 6 (this document) shows that this constant down-valley wind speed at MER is modeled up to the ridge height.*
>
> *The pressure difference along the valley was not analyzed. It has however to be noted that constant down-valley wind speed during night was also measured and modeled in the Rhône valley at Sion, Visp and in the narrow Magia Valley at Cevio (Fig.4 and 6 in Schmidli and Fig. 11 in Quimbayo, 2021, Schmid et al., 2000) without a discussion about the difference with theory. It seems however that the valley size is not the main explaining factor.*

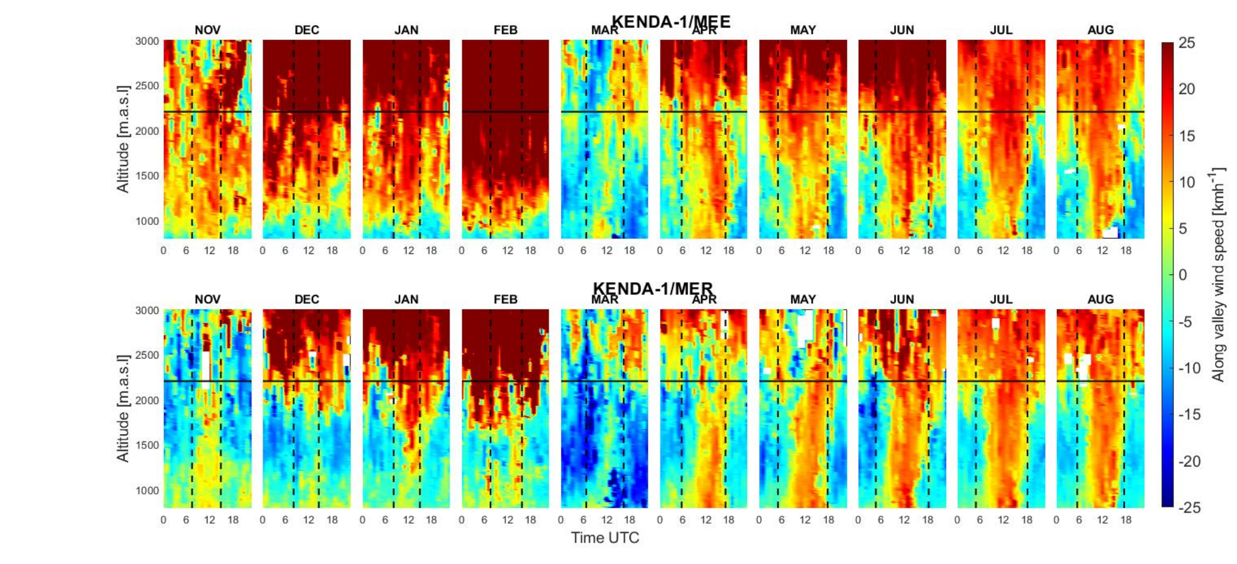


Fig. 6: Along valley wind speed modeled by a) KENDA-1/MEE and b) KENDA-1/MER.

34. l. 328-334: The described upvalley wind characteristics are not typical. For example, why should the upvalley wind near the surface be stronger and more regular than at 200 m above the ground. This should be discussed. Could this be related to sampling or an artifact of the projection to along valley direction? Would the valley wind system show more typical characteristics when synoptically dominated winds were excluded from the composites? The difference in onset time of downvalley wind at MER and MEE could be related to the Br¨unigpass. However, this needs to be formulated as hypothesis without sound observational evidence (which is not given in Sect. 3.3).

The first column of Fig. 7 (a and c) presents the SMN/MER observations and the modeled KENDA-1/MER at MER whereas the second column (b and d) presents the DWL/MEE observations and the modeled KENDA-1/MEE at MEE. The SMN/MER observations are done at 10 m a.g.l. whereas the other plots correspond to data from 775 m a.s.l., the first DWL level. KENDA-1 data were taken at the same altitude as the DWL to allow a comparison with DWL/MEE, the first KENDA-1/MER level (739 m) being anyhow much higher than the SMN station. Fig. 7 clearly shows that the main difference in both up-valley and down-valley wind speeds is found between both stations and not between different altitudes. MER presents stronger maxima in up-valley wind speed and MEE stronger maxima in down-valley wind speed for both the observation and the model. Fig. 8 confirms that the valley wind speed is usually rather constant along the profiles, at least in the first 1000 m above ground.

We try to select different angles around the valley axis (from +5° to +-15°). These tests as well as a real projection on the valley axis leads to the same result, with stronger up-valley wind at SMN/MER than at MWR/MEE and KENDA-1/MEE. We do then believe that the described up valley wind characteristics are not measurement artifacts.

Further analysis (see answer to the first general comment) explains now the mechanism of flows from the Brünig Pass. Anyhow this last sentence is now formulated as a hypothesis.

35. l. 335-336: The statement that turbulence is leading to daytime varying wind direction is not obvious and needs to be supported by observational evidence or removed.

Yes, the mention of turbulence was removed.

36. Fig. 7: Please enlarge axis labels. Replace MWR/MEE in caption with DWR/MEE. What values are plotted? Are these hourly values or monthly values? If monthly why not plot hourly values to show day to day variability?

The axis labels are enlarged and the figure caption corrected. Monthly values are plotted since the noise induced by the daily data masked the main features.

37. l. 338: 7b instead of 7c.

Done

38. l. 340-343: This atypical behavior of the valley wind (upvalley wind during the night, upvalley wind during all day in winter) in KENDA-1 is possibly a result of averaging over all types of conditions. It is hence not clear if KENDA-1 struggles with thermally driven flows or channeling events or on clear or cloudy days. Filtering and focusing on specific conditions would be beneficial to learn more about model deficiencies.

This comment is right, and the present analysis does not allow us to find the causes of the model deficiencies. We hope that colleagues from the modeling community will use our dataset to further test the models, but this is beyond the scole of this paper.

39. l. 344-354: In my opinion, it is not valid to draw conclusions for up- and downvalley winds and on impacts of synoptic flows from the monthly composites, since the composites are most likely strongly affected by synoptic winds and clouds. To draw meaningful conclusions on up- and downvalley winds, the days need to be filtered for conditions favorable for thermally driven winds. Also, some of the aspects (decrease of downvalley wind speed with height) could be a result of projecting on the valley axis and should be investigated. Without investigating other factors (clouds, synoptics, sampling size, etc) it is not valid to attribute varying wind direction during daytime to turbulence.

As answered to the first general comment, no projection but a selection of wind around the valley axis are applied to compute the along valley wind compound. There is consequently no artifact from the projection.

Fig. 3 (this document) and further analysis with cloud cover <3 oktas at FRU showed that the described features are similar with a selection of clear and very clear days. Moreover, we think that thermal wind occurs not only during fair weather days but also, with a lower intensity, under less good conditions. A selection of only very good weather situations also restricts the analysis of thermal valley wind.

The mention of turbulence to explain varying wind direction during daytime was removed.

40. Fig. 8: How are the composites computed? Are KENDA-1 data only used when observations are available and valid? If not, the comparison is not fair. Are white gaps in panel a) due to small wind speed or missing data? A different color should be chosen for missing data (e.g. grey).

Fig. 8 was corrected to consider only data present in both time series. White gaps correspond to missing data and low wind to light green. The text was then revised according to the new Figure.

41. l. 366: What does 'data are scarce' mean? Shouldn't the sample size for across valley wind be the same as for along valley wind?

No, the lower amount of data in Fig. 9 is due to the absence of wind in the cross valley direction. As described in the experimental section, too weak winds (speed < 2km/h) were discarded for wind direction analysis.

42. l. 366-367: The northerly flows in January and February are not clearly visible.

Northerly flows come from the south-facing slope, namely from the Sarneraatal and are depected in red. We think that they are visible on the figure as indicated by the red square on the next Figure:

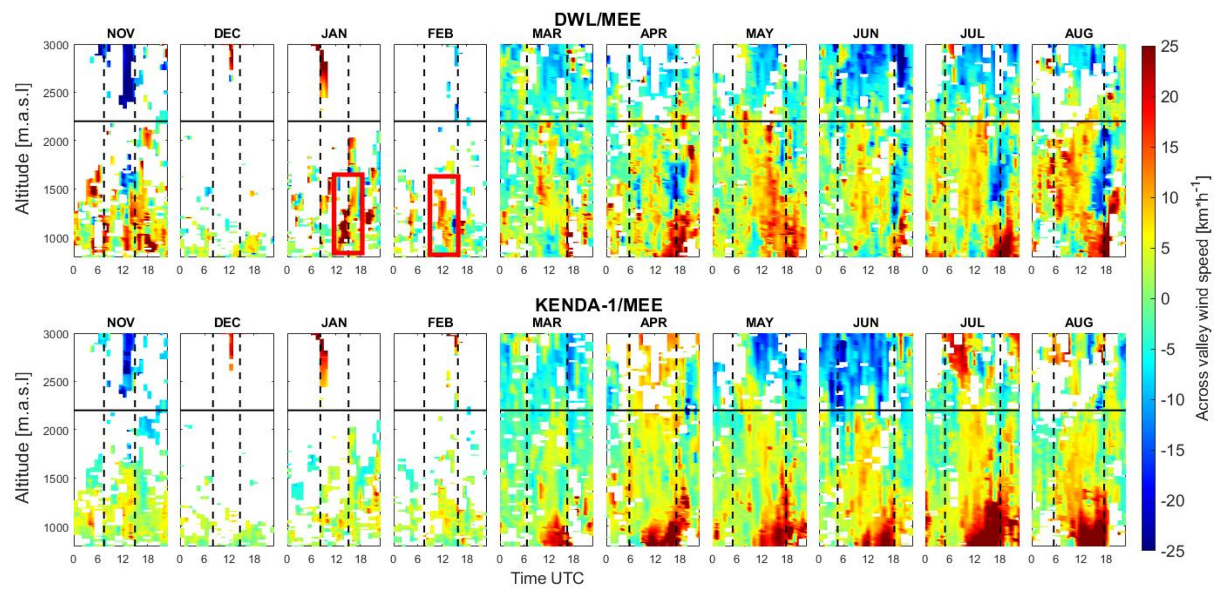


Fig. 7 Evolution of the diurnal cycle of the cross-valley wind component [km/h] as a function of altitude for a) the DWL/MEE measurement and b) the KENDA-1/MEE. Winds coming from the south-facing slopes take a positive value (red), for the north-facing slope wind speeds values are negative (blue). Sunrise and sunset at ground level are given by dotted lines.

43. l. 370: What are north-facing slope winds? Do the authors mean downslope winds from the slopes north of the valley (they are south-facing)? For clarification, it would be helpful to repeat the colors used in Fig. 1 to distinguish southerly and northerly cross valley flow.

As explained by the reviewer, north-facing slopes lie south of the valley. The sentence is perhaps misleading and was modified: "Intense winds from north-facing slope (> 25 km/h) are also observed between 1400 and 2000 m during some hours around sunset with a much lower intensity in May." The colors correspond to the colors used in Fig. 1 and are described in the figure caption of Fig. 9 presenting the cross-valley winds.

44. l. 371-374: In my opinion there is not enough evidence for a cross valley vortex from the example RHI plot in Fig. S9. I can see downslope and upslope components, but no closed circulation.

The instrumental set-up does not allow us to obtain a 3D image of the wind compound (see answer to the first general comment). The radial wind depicted on Fig S9 presents however a clear example of a cross valley circulation and we do agree that the vortex nomenclature is misused so that it was not used in the manuscript. We then replaced it with cross valley circulation in the supplement.

45. l. 375: What color are the winds that descend from Br¨unigpass? Are they red? Please specify.

As specified in the caption of Fig. 9, winds from the south facing slope are taken as positive, namely positive. The Brünig is situated on the ridge of the south facing slope (see Fig. 1) so that winds from the Brünig are depicted in red. This is now more precisely specified in the figure caption:

"Winds coming from the south-facing slopes, namely from the Brünig Pass, take a positive value (red), for the north-facing slope wind speeds values are negative (blue)."

46. l. 380-383: Why is the along valley wind mentioned here? This section is about cross valley. It needs appropriate context and reference.

This sentences and the next ones are not relevant here and were deleted.

47. Fig. 9: How are the composites computed? Are KENDA-1 data only used when observations are available and valid? If not the comparison is not fair. Are white gaps in panel a) due to small wind speed or missing data? A different color should be chosen for missing data (e.g. grey). What is the sample size at each point (can some of the noise be explained by varying sample size)?

As explained concerning Fig. 8, Fig. 9 was also modified so that only data present in both time series are plotted. White dots correspond also to missing data of no across valley wind. The text was also adapted to the new figure.

48. l. 394: On this fair weather day, wind speeds of 25-30km/h are reported. How does this fit to the filter of 20km/h to distinguish thermally from synoptically driven days?

The threshold of 20 km/h is no longer used in this revised manuscript.

49. l. 399-401: The outflow from Br¨unigpass cannot be thermally driven. Why should warm air during the day descend from the pass to MEE? If there is upvalley wind in Sarneraatal that reaches the pass and descends on the south side of Br¨unigpass there must be a dynamic effect driving this (wave, etc). Please explain.

As explained to general comment 1, a new analysis showed that the air masses in the Sarneraatal are often colder than in the Haslital over MEE. The outflow from Brünig pass is then thermally driven and descends to MEE because it is colder than air masses above at the altitude of Brünig pass. The valley volume effect can explain the T differences between both valleys. This is now explained in the manuscript.

50. l. 409-418: This whole paragraph is based on Fig. S8. If this figure is so important, it needs to be included in the manuscript.

The new analysis of the pressure difference between GIH and MER and the potential T difference between BRU and MWR/MEE at 1010 m deserves a new figure. The inclusion of Fig. S8 would too much lengthen the manuscript.

51. l. 419ff: It is essential to include the model terrain in this analysis to understand how the model sees the Sarneraatal and the pass.

The height of the ridge and the altitude of the pass are given in the experimental section. In the model, the Brünig Pass is only 200 higher than the valley floor, enhancing the potential impact of the Sarneraatal on the wind system in the Haslital.

52. Fig. 10: Both panels should have the same y-axis range.

In fact, both panels have the same y-axis range. It is not obvious since the wind direction stripes are pasted on the figure and that Fig. 10a provides 3 wind direction stripes and Fig. 10b only two.

53. l. 439-441: Specify where clear weather conditions can be expected. Describe the foehn characteristics at MER (direction, over which ridge it is coming). What stations are used to compute the foehn index?

Clear weather conditions are expected during the whole foehn event on the northern side of the Alps' ridges. The foehn at MER comes from the Grimsel Pass and follows then the Haslital. This is now specified in the manuscript.

The main prerequisite for the occurrence of foehn on the northern slopes of the Alps is a southerly wind on the main Alpine ridge, which is measured at the Gütsch station, Andermatt, (GUE). Conversely, on the southern slope of the Alps, the wind on the main Alpine ridge must come from the north for foehn to occur. The other parameters (average speed, wind gust, wind direction, relative humidity and potential T are taken from SMN/MER as described in sect. 2.3.1.

54. l. 442: No June episode shown.

The foehn event in June is considered in the analysis of the T, while only three events (10-16 March 2022/19-22 March 2022/26-24 April 2022) are described in the analysis of the wind. The sentence was modified.

55. l. 445 and Fig. 11: Unless foehn starts always at the same time of the day, the composites should be shown relative to foehn onset and not for hour of the day.

This figure intends to show that the diurnal cycle of the T difference between KENDA and the measurements is not found during foehn events and that the T is always overestimated by KENDA-1.

56. l. 453ff: This whole paragraph is again based on Fig. S11 and S12 from the supplement. If this is discussed in so much detail, it needs to be included. However, related to my comment on adding more focus to the paper, I think the whole discussion on foehn should be much shortened or even removed. I also cannot see a T gradient in the types of plots in the supplement.

We consider that the modeling of foehn events is still challenging so that the description of foehn events should not be removed but the description of the wind during foehn event was shortened. Figure S11 and S12 are of clear interest and are not really discussed in detail, so that they must remain in the supplement not to lengthen the paper.

57. l. 465-466: I don't see the point in comparing different heights at MEE and MER.

The sentence was modified. As it was intended, the point is now focused on the comparison of the maximum wind speed: "The maximal wind speed (60-75 km/h) of DWL/MEE is observed at 800 m and is much higher than at the SMN/MER (45 km/h), especially for the event of March 11."

58. l. 467-484: I think this description of wind is way too detailed and distracting and could be much shortened or removed.

Sect. 3.4.2 was largely shortened and consists now of one § to describe the SNM/MER and DWL/MEE observations and one § to describe the differences between the observations and the modeled data.

59. l. 485-494: Given that KENDA-1 provides 3D output, the foehn cases could instead be investigated in KENDA-1 to understand the spatial differences and the model errors during foehn events.

This is out of the scope of this paper. The aim of the paper is to compare REM measurements with KENDA-1 and not to study the foehn event in the Haslital. Since a poor agreement is found between KENDA and the measurements, we cannot rely on the 3D KENDA data to study to have a realistic picture of the foehn event in the Haslital.

60. Discussion: The discussion section is too long in my opinion and should be more focused. For example, a lengthy comparison to inversions and thermally driven flows in other studies is shown, but given that the composites in the present study are not filtered for thermally driven flows this comparison is not very meaningful. The evaluation of KENDA-1 was done visually based on time-height sections and not by computing model skills. This would have been more meaningful, instead of the descriptive comparison of the model to the observations. I think saying that 'KENDA-1 proposes good monthly median values' (l. 662) and 'Despite the complex topography around MER and the induced elevation bias, the modeled climatology of ground T is comparable to standard verification results' (l. 595-597) is hence not sufficiently supported by the analysis.

As already mentioned before, thermally driven wind occurs not only during fair-weather days. We think that a comparison with the three cited studies is necessary. The discussion was a little shortened.

The evaluation of KENDA-1 with computed skills would have been strengthened. The visual evaluation is still valuable. The mention of standard verification results was removed.

61. l. 574-584: The vortex in the Inn Valley was caused by the valley curvature (Babi´c et al., 2021). This means that the mechanisms here are likely not comparable and caution is advised.

Yes, the vortex described by Babic et al, 2021, is caused by the valley curvature. The valley curvature between MEE and MER is situated between both stations and not exactly at the DWL site. Secondly, the *observed cross valley circulation is certainly influenced by the winds from the Brünig Pass. The § was modified to enhance the differences between both environments: "However, contrary to the CROSSIN campaign's results, valley winds from the Sarneraatal are probably the main drivers of this cross valley circulation in MEE."*

62. l. 585-587: No skills are computed for KENDA-1 and an objective verification of model skills was not done. This conclusion is hence not supported by the presented results.

Since no skills were computed for the used locations, we omit this sentence.

63. l. 595-597: The reference to standard verification results is confusing and needs to be explained and justified by results.

This sentence has been rephrased and a reference has been added.

64. l. 599-600: Since monthly composites are shown this statement is not supported by results.

This statement comes from an in-deep analysis of the cold event at the end of March that is not reported in the manuscript. "Result not shown" is now added to the sentence.

65. l. 605-617: RH depends on T. Thus, a warm bias leads to a dry RH bias. The statement on RH does not provide additional information. Bias in terms of specific humidity would be more meaningful.

What does 'artifacts from the NWP can be expected under conditions favorable to surface T-inversion' mean? The statement 'Finally, the differences with observations can also originate from a modeled ongoing turbulent mixing whereas in reality a cold pool with a full or partial decoupling from the above flow is present in the valley.' is not supported by results.

The sentence has been reformulated as a hypothesis.

66. l. 618ff: MWR liquid nitrogen calibration plays a role in MWR profiles biases and should be discussed. Average differences are discussed, but what about individual profiles?

In our experience, after a successful calibration with liquid nitrogen, the biases become negligible.

Some individual profiles are shown in the supplement. However, the goal of the paper is to perform statistics, not to discuss individual profiles.

67. Conclusions: Since many readers only read the Summary some basic information on sites and data should be repeated. I don't think that all conclusions are sufficiently supported by the results (e.g. l. 688-690, l. 693-695). I furthermore do not think it is fair to say that the study 'deepens our consensual knowledge about atmospheric phenomena in complex topography'. It is mostly a description of conditions without any in depth investigation of processes.

The conclusion was modified following the recommendations. Basic informations on sites and instruments were added. The sentence on cross valley circulation was modified and the words "vortex" and "closed circulation" were removed. We consider however that the measurements showed a cross valley circulation. The last sentence was also modified.