Use of thermal signal for the investigation of near-surface turbulence

Matthias Zeeman[1]

[1]Karlsruhe Institute of Technology, Atmospheric Environmental Research Institute of Meteorology and Climate Research,Garmisch-Partenkirchen, Germany

Correspondence: Matthias Zeeman (matthias.zeeman@kit.edu)

15 Oct 2021

Dear Editor,

I thank the two reviewers for their helpful comments and suggestions. Detailed information on the changes made to the manuscript can be found below. As some sections were moved and restructured, also the figure order changed accordingly. In addition, I found errors in Table 1 that are now corrected.

I look forward to learn how we can proceed with the revised manuscript.

Kind regards,

Matthias Zeeman

Final Response

Reviewer #1

General Comment

The author presents a measurement technique which combines a thermal imaging instrument and distributed temperature sensing system in order to monitor spatial and temporal fluctuations of the temperature. Those measurements are supplemented by point observations of the temporal wind fluctuations, which are used in order to provide a more thorough insight of the local ambient conditions. The article is in general well written and the measuring techniques and the data analysis are presented in detail.

However, I think that the results presented in the submitted manuscript are lacking an assessment of how accurate and precise the estimation of the measured temperature fluctuations are. This is discussed briefly in the text and a qualitive comparison could be performed visually from the results in Figure 3. However, it will contribute to the

assessment of the measuring capability of this setup if a direct comparison with the reference sensors is performed.

Response: Assessments of measuring capabilities were the subject of earlier studies and I chose not the repeat those here to avoid redundance. As the reviewer points out, the reader may be inclined to use Figure 3 for a qualitative comparison of temperature observations, but the presented detail is not well-suited for a performance assessment. I agree that the addition of a comparison between the temperature methods would be helpful. I further suggest to add text to the interpretation of the Alan variance results (Figure A7), which had not been included in previous studies and reveals insights for possible improvement of the calibration of DTS data.

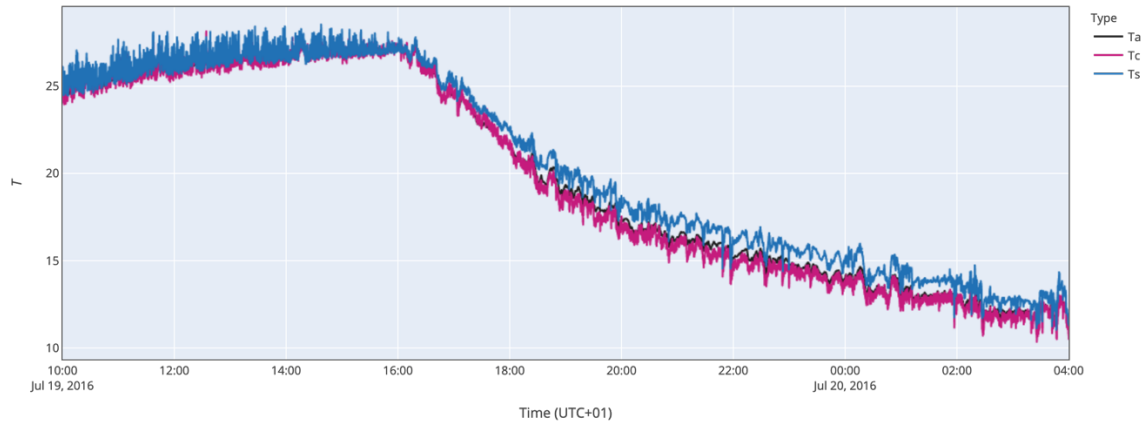

Figure R1: Timeline of the temperature comparison between Ta (aspirated), Tc (DTS cable) and Ts (sonic anemometer).

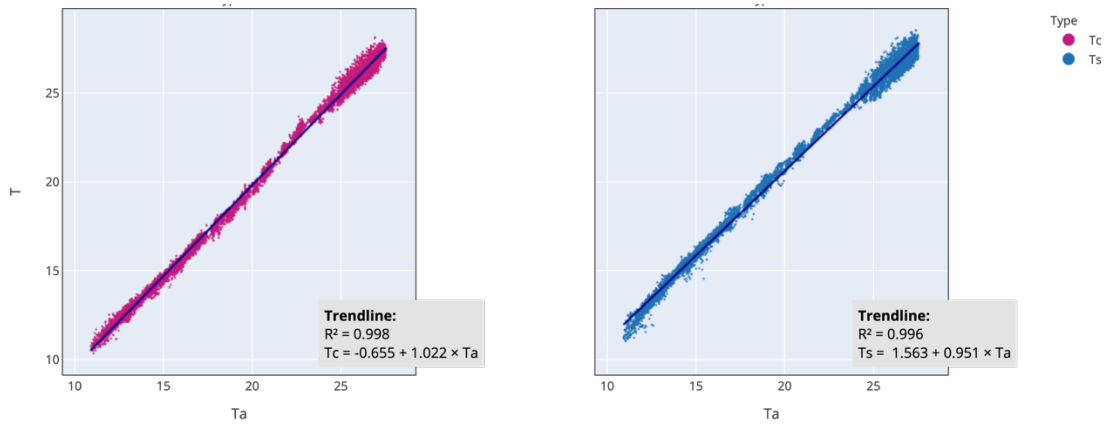

Figure R2: Comparison between Ta and Tc (left) and Ta and Ts (right).

A comparison between temperature measurements by the different systems was made, using sensors that were closely located in space. For the cross-comparison, sensors at 3 m height were selected nearest to the Ta observations (aspirated sensor profile) and a period with approximately equal number of daytime and nighttime hours

(12:00 19 Jul 2016 to 04:00 20 Jul 2016). The results show that both Tc and Ts are linearly correlated with Ta, but did not exactly match the aspirated temperature in absolute value and high frequency variance (Figure R1 and R2). The Ta sensors exhibit a slower thermal constant due to the size of the metal sensor rod, leading to a dampened high-frequency response compared to Ts and Tc. Furthermore, Ts was cross-calibrated among the EC sensors prior to this experiment, but not calibrated to a temperature reference scale. This shows a need to include further calibration or Ts in future experiments. However, differences between aspirated and non-aspirated sensing methods are inevitable and explain part of the deviation between sensing systems seen here.

A comparison between Tc and Ts shows linearity between the signals, which appeared to improve in absence of direct sunlight during the night (Figure R3). However, the computed coherence between Tc and Ts deteriorates at high frequency (f > 0.1 Hz, dt < 10 s; Figure R4). This can partly be explained by noise in one or both of the sensors, as well as small differences between the location and path length of the sensors being compared.

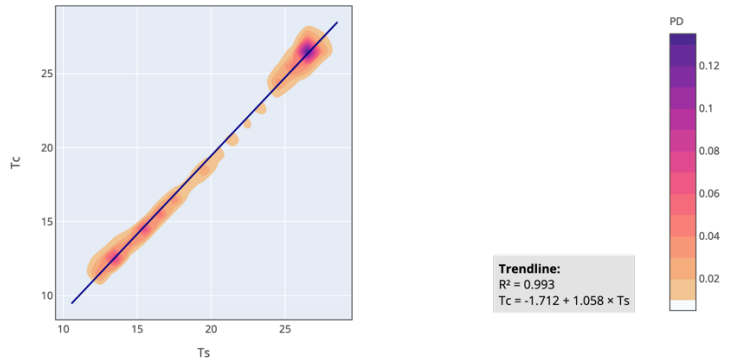

Figure R3: Comparison between Tc and Ts, shown as Probability Density (PD) with the overall trend (darkblue line).

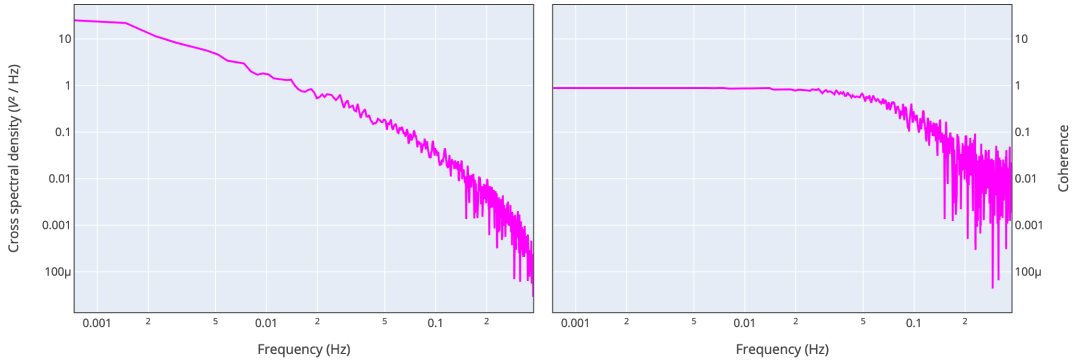

*Figure R4: Shown are against frequency (**left**) the Cross spectral density and (**right**) the computed coherence between Tc and Ts.*

An additional subsection and figure describing the comparison between different temperature observations was added to the Appendix (A3.4 Cross-comparison).

Moreover, I think that it is not explained clearly the reason for selecting the specific shape and size for the experimental setup. This information is going to be useful for understanding and interpreting the results of this study.

Response: Thanks for the suggestion.

In a nutshell, the design was intended to support studies on coherent structures, advection processes and conditional sampling methodology. Both as empirical study and in combination with fluid dynamics models. Placement of the wind sensors on the corners and at the center was intended for the determination of a representative wind vector at the walls of the box. It was thought that placing tripods at the center of a wall would result in more uncertainty. The dimensions and shape were primarily limited by the maximum range of the DTS instrument, which can support up to 1.8 km of optical fibre per channel. The DTS profile height was extended to reach sufficiently above the sonic anemometers at 3 m. In an initial design, the guyed mast would be taller and placed in the center. A compromise had to be made during deployment and the mast was moved to a corner. This also had obvious implications for the field of view of the TIR system.

The setup was not a stand-alone experiment. Other experiments were conceptualized to use the setup, as part of cooperative research during the experimental campaign (ScaleX; https://scalex.imk-ifu.kit.edu/; line 70-76). To name a few:

- Adjacent to the setup, a horizontal gradient of Tc was measured over a distance across the shallow valley, and in parallel to a transect of sonic anemometers during a different period (Mauder and Zeeman 2018)
- The combined DTS setup was included in an area of TIR mapping by UAV (Brenner et al. 2018) to evaluate an approach similar to the aquatic study cited in the review comments below (Dzara et al).
- The setup was located inside a larger valley area observed by a network of multi-Doppler lidar, SODAR/RASS and meteorological stations.
- Adjacent to the setup, a team observed advection of trace gasses and energy in similarly oriented 20x20m area (Zhao et al.).
- Adjacent to the setup, trace gas and water vapour fluxes were observed by a network of automated chamber systems (e.g., Zhao et al 2018).
- Adjacent to the setup, a team investigated patterns in air pressure perturbations (Mohr et al.)

A subsection was added to the methodology section to include such details in the main text (section '2.5 Design considerations').

Specific Comments

Line 1. Organized motions of what? Please specify.

Response: Thank you for the comment. The following change was made:

'Organised motions **of air** in the roughness sub-layer **of the atmosphere** …'.

Line 6. What is meant with the term "Variance Events"?

Response: Thanks for pointing out the use of jargon. The term variance event is used to distinguish signal in the time series without (obvious) periodic pattern, but representing significant excursions from a mean or trend. I suggest to rephrase 'variance events' to help the reader.

'**Events in the temperature signal**…'

Line 9. I suggest replacing the "with the naked eye" with visually.

Response: Thanks for the comment, I agree. **The change was made as suggested.**

Line 13 – 14. The author states that "the available methods to determine energy and scalar fluxes from terrestrial land surface are relatively imprecise due to a multiscale of irregularities in the land surface and the turbulent transport mechanisms". I think that this is statement is not very clear. This imprecision originates from limitations in the precision of the measuring methodologies or does the imprecision refer to the need for spatially distributed measurements?

Response: The line refers to limitations in the measuring methodologies, which in part relate to spatiotemporal irregularities in surface-atmosphere interactions.

The methods are precise, just not applicable everywhere and all the time. As a consequence, there is a pattern, perhaps bias, as to where and when energy and scalar fluxes are observed today using micrometeorological techniques. In general, complex terrain is avoided, and periods with stable atmospheric conditions (e.g., night) or hydrometeorological events (e.g., rain, fog) are excluded. This introduces a systematic uncertainty.

Many researchers have asked the question: 'what [are we] missing outside the applicable range of the methodology and assumptions?'. Spatial distributed

measurements may not be the final answer, but I think they could help with the assessment of processes and the development of empirical methods.

Line 24. An abbreviation should be added after the “roughness sublayer”. Later in the document (line 26) is referred to as RSL.

Response: I agree, thank you for noting the omission. **The definition of the abbreviation was made as suggested by the reviewer.**

Line 45. What it is meant by the following statement: “the quantification of Tb outside a controlled laboratory environment is a challenge, in and of itself”?

Response: There are many possible sources of interference in thermal imaging, which in a controlled laboratory environment can be observed and corrected for. In field studies this is much more difficult.

The sentence was modified:

‘The quantification of \Tb\ outside a controlled laboratory environment is a challenge, in and of itself; **there are many possible sources of interference in thermal imaging, which in a controlled laboratory environment can be anticipated and corrected for.’**

Line 57. I recommend that this statement about the goal of this study is also mentioned in the abstract. It will give a clearer idea to a reader about the objective of this study.

Response: Thank you for the comment. I agree. **A statement about the goal was added to the abstract.**

Lines 58-61. I think that references to previous studies that have used the DTS and TIR measuring techniques should be mentioned. An example is:

Dzara, J. R., Neilson, B. T., & Null, S. E. (2019). Quantifying thermal refugia connectivity by combining temperature modeling, distributed temperature sensing, and thermal infrared imaging. Hydrol. Earth Syst. Sci., 23(7), 2965–2982. https://doi.org/10.5194/hess-23-2965-2019

Please note that I am not neither the author or any of the co-authors of the aforementioned study.

Response: Thank you, I was not aware of the study. The suggested study used the two techniques in an aquatic environment. Please note that TIR in the study by Dzara et al was used to generate thermal maps (one per season), which were then matched and compared to statistics of a DTS time series. In contrast, this manuscript presents an approach where DTS and TIR are both used as high-resolution spatial time series.

Lines 65: What does ICOS stand for?

Response: Thanks, the definition is indeed missing. **The acronym was added to the text**:

'The study was conducted at the DE-Fen station, Fendt--Peissenberg, Germany, which is a TERerstrial ENvironment Observatories (TERENO) and **Integrated Carbon Observation System (ICOS)** core site …'

Lines 68 – 69: In these lines the author gives details about elements of the landscape surrounding the experimental area. It is not clear how this information is relevant to the study. I suggest that the author explain briefly the impact of the landscape to the experiment presented in the manuscript or remove that part.

Response: I fully agree with the reviewer's suggestion to remove this part.

Line 70: What are the ScaleX campaigns?

Response: I agree that this concept is not properly introduced and suggest the following.

'During **intensive field** campaigns **at DE-Fen**, additional experiments were conducted for the investigation of scale interactions between the atmospheric boundary layer and the surface, as well as validation of measurement techniques (**ScaleX**; …).

Line 74: Why is the period between 18 – 22 Jul 2016 considered as a reference period?

Response: This is indeed not a necessary qualification. **The word 'reference' was removed from the sentence.**

Line 75: What was the purpose of the UAV use? And how could they have an impact on this study?

Response: UAVs were used for mapping surface brightness temperature at a larger spatial scale and for in situ measurement of wind field and air quality properties. Those studies were also part of ScaleX and referenced in the paragraph text. In some

instances, horizontal transects were flown above and upwind of the setup. Particularly the heavy airborne platforms, e.g., carrying a gas analyzer, generated a downwash jet that could be sensed at some distance. Also, some UAV operations required teams of people moving in and out of the field, e.g., for hourly off-site charging of batteries. The flight tracks were recorded in detail, the movement of people and vehicles not.

Added to the text: …'perturbing the air **by down wash and frequent transit through the area**.'

Line 78: What does EC stand for?

Response: The definition used is 'Ultrasonic anemometer (EC)' (line 77). EC refers to the Eddy Covariance technique in which these instruments are used. No changes were made to the text.

Line 78: I think that it would be very helpful for a reader if the author specify that is the figure 1c and table 1.

Response: I fully agree. **The references were updated as suggested.**

Line: 82: What was the reasoning for the number of sonic anemometers used, the selection of the locations of the tripods and the heights of the sonic anemometers?

Response: This was limited by available hardware at the time. Ideally, only 3-axis sonic anemometers would be used. A compromise had to be made for the number of 3-axis sonic anemometers deployed here and elsewhere during the campaign. The height of the 3-axis type instruments on the tripods was chosen to be similar to the ICOS station and other studies using the Eddy Covariance technique on permanent grassland (including DE-Fen; see also the study by Mauder and Zeeman cited on line 80). There are sensitivity limitations for working with current model ultrasonic anemometers close to the surface. A level of 2-axis sonic anemometers closer to the ground (0.25 m) was planned but the appropriate mounting hardware could not be arranged in time for deployment.

Details were added to section '2.5 Design considerations'.

Also, from the Figures 1 and 2 it is visible that the sonic anemometers were located between the supporting poles of the DTS mast. Could there be any interference to the sonic anemometers measurements acquired during the period selected in this study from wakes generated from the supporting poles?

Small-scale interference in the wake is possible. The distance from the DTS masts to the EC profiles (on tripods) was 3 m. DTS masts had a diameter of 0.1 m. Increasing the distance would have required for the suspension cable to be mounted higher and with larger tension force to keep the steel cable straight. This could not (safely) be realized during the deployment.

Details were added to section '2.5 Design considerations'.

Line 97: Where was the TIR system pointed to?

Response: The TIR system was pointed to the ground at a slanted angle to include as much surface within the DTS box as well as static objects for georeferencing. The guyed lattice mast was planned to be taller, but had to be kept below 10 m for safety of nearby glider planes. This limited our options for the camera viewpoint during the deployment.

Details were added to section '2.5 Design considerations'.

Line 101: What is meant that the location was determined in post-processing?

Response: Thank you for the comment. This means that it required a georeferencing step as described in the text below the line and Appendix A. **Change the text now make this clear.**

'Each EC, DTS and TIR record was stored with an accurate time stamp and locations were **georeferenced** in post-processing. The calibration and **georeference** details are provided in Appendix A.'

Line 116. The air temperature (Ta) is mentioned here, but it is only discussed how it is measured in Appendix A3.3. I would suggest a brief statement about those measurements also in section 2.4.

Response: I agree. This is an issue and I agree with the suggested solution.

Added sentence:

'Reference air temperature measurements were made using resistance temperature devices in fan-aspirated enclosures (Table 1; Appendix A)'

Additionally, regarding Figure 3. What is the sampling frequency of the time series presented in Figure 3?

Response: Those are 1 min averages in all panels. I suggest to change the caption accordingly to better inform the reader.

Changes were made to include the frequency of the time series in the captions of Fig 3 and Fig 4.

How is the Tc estimated at the presented heights? Is it the average over all the four sides of the box?

Response: Yes, it is the average over all Tc profiles.

Which sonic anemometer's data is being used in the Figure 3 c?

Response: Only the 3-axis sonic anemometer models provide a measurement for Tv, hence these are data for 3.0, 6.0 and 9.0 m height.

Line 127. The author states that "some turbulence statistics were rarely acceptable … "

What is it meant by the words "some" and "acceptable"?

Response: For the application of the eddy covariance technique, it is currently recommended to perform a number of (self-)validation tests, e.g., based on stability, stationarity and friction velocity. The test results can be simplified as a quality classification for the averaging period. During the nights it was rare to find the quality of averaging period results classified as 'acceptable'.

I suggest to rephrase this to 'eddy-covariance flux computations **rarely produced acceptable results**…'

Line 129-130. It is not clear how what are the assessment criteria used here to assess the quality of the flux computations.

Response: The assessment criteria may be the same, the computation to derive a stability classification is different.

Line 132. How were the temperature gradients calculated?

Response: Thank for noting the omission. This is indeed not mentioned clearly. **The definition for temperature gradient was added to the list in Appendix B.**

Line 136. Figures 3a-c allow a visual comparison of the time series. However, there is a lack of a statistical comparison of the different methods (e.g. correlation, mean absolute error). I suggest that the author elaborate more on this part.

Response: I agree. Please see the response to the General Comments.

Line 140. In Figures 4b-c time series of the normalized by the Obukhov length scale height and the friction velocity are presented. Measurements from which sonic anemometer were used for those calculations. Which criterion has by used to assess the atmospheric stability is stable or unstable?

Response: Only results from the 3-axis sonic anemometers can be used for the computation.

Details about height were added to the caption of the figure 4.

Line 145. Is it the air temperature or the cable temperature presented in Figure 6?

Response: These panels are derived from cable temperature. I think this is sufficiently clear from the caption and legend.

Line 189. The DTS measurement set-up has a rectangular shape. What is meant here the mean wind was mostly aligned to the set-up?

Response: Thanks for the comment. I now recognize the wording can lead to confusion for the reader and this aspect should be rephrased. What was meant is that for a period of time the mean wind was either perpendicular or parallel to the walls of the box.

The sentence was modified for clarity:

'During the daytime animated period, the mean wind is mostly aligned with the DTS measurement setup**, i.e., either perpendicular or parallel to the walls of the box-shaped array,** with mean flow from a SSE direction.

Line 191: How does the animations reveal scale interactions? And why they are note easily identified in the statistical analysis?

Response: Thanks for the comments. Scale interactions are revealed in statistical analysis, as the results show. Nevertheless, it can be helpful, educational, to review those results visually.

Line 196: I do not understand what is meant with this statement. Can the author elaborate explain this a bit more?

Response: Thanks for pointing this out.

The surface was not homogenous in terms of TIR signature. Some signal in the TIR image time series were revealed when and where the background (the surface) and foreground (air that had interacted with the surface upstream) show a different heat signature. Therefore, motion was revealed from hot air advecting away from relatively hot areas in the plant canopy, against a background of cooler surfaces. I agree with the suggestion to rephrase line 196.

The following line was added to the paragraph:

'The surface was not homogeneous in terms of TIR signature. Some signal in the TIR image time series were revealed when and where the background (the canopy surface) and foreground (air that had interacted with surface upstream) show a different heat signature.'

Line 213. Why is there a sudden jump in the TKE in Figure 11 between 00:00 and 12:00 in 21 jul 2016?

Response: The jump is correlated to the passing of a short storm with brief precipitation (See panels Figure 3f and Figure 4a, and the text at line 120).

A sentence was added to clarify this:

'A sudden jump in TKE shortly after sunrise on 21 jul 2016 is correlated to the passing of brief rain storm (Figure 11c; Figure 3f; Figure 4b).'

Line 228. The author gives a very thorough list of the limitations of the current measurement technique. It would be very constructive if the author could provide a short recommendation regarding in which applications this setup should or shouldn't be used.

Response: Thank you for the comment. A short recommendation can be added here (see the response to the General Comments).

Line 235. Can the author elaborate more on why three-dimensional sonic anemometers at lower heights would be advantageous in this study?

Response: Thanks for the comment. It would have been helpful to show TKE, w' and stability information, as derived from 3-axis sonic anemometers, at a lower level.

The following sentence was added after the line in question:

'It would have been informative to compare vertical wind speed, TKE and stability information to height levels of interest below 3 m.'

Line 249. How is this precision calculated?

Response: This conclusion refers to results discusses in line 241 and shown in Appendix A). As suggested in the general comments section above, more detail on the performance would benefit the presentation.

Line 250. What is the reference scale for the recommendation for the size spatial domain and what is meant with the "2.5 dimensional or better"?

Response: A setup with a combination of 2-dimensional planes as cross-sections provides more than 2-dimensional information, but less than a 3-dimensional grid. Hence this is referred to as 2.5-dimensional. A fully 3-dimensional setup would, for example, resolve locations in a regular grid similar to many fluid dynamics models. In principle such a 3-dimensional setup could be achieved in the field, with some effort, using DTS.

Line 260. Does "turbulence" refer here to wind speed or temperature?

Response: Both.

Line 261. Both here and in the abstract, it is mentioned the development of physics-aware machine learning techniques. The current topic is not discussed in the introduction, so it is difficult to understand what a physics-aware machine learning technique is, assess how this study contributes to their development and understand their potential value. I think that it would increase the comprehension of the manuscript if the author could briefly explain this.

Response: I agree, a brief explanation will be helpful. Please note that the subject is discussed in the results section (lines 244-247).

Line 271. How accurate was the time keeping?

Response: The hardware clock data sheet specifies a 2ppm accuracy. There was no measurable drift on any of the systems.

Line 285. How did the author recognize the period with winds from the north?

Response: Both the sonic anemometer (EC) network and the wind observations from the DE-Fen station indicate wind direction. This wind sector is frequently observed. The situation is maintained for several half-hour periods during the day, due to the proximity of the Alps to the south. Assumed was that wind from this sector would have limited wake effects on any of the sonic anemometers by mast structures or topography.

Lines 369 – 380. Why is this paragraph in the appendix? Isn't this part of the results?

Response: Thanks for pointing this out. **The paragraphs were moved to the methods section.**

What is the physical meaning of the grouping of the clusters presented in Figure A7? Also, what is the impact of variations of atmospheric stability in the results presented in Figure A7?

Response: The original study on the TED method shows the shape of the temperature variance events for each cluster, as extracted from idealized data. Different shapes suggest a different physical meaning.

Table 2. What is the reason for mentioning the different ways of parameterizing the atmospheric stability? How is this used in this study?

Response: The different parameterizations of atmospheric stability are used as background information for the reader. I am not sure at this point if or how any of the parameterizations can help improve the classification of turbulence events. Personally, I found the differences between a low (1.0 m) and high (3.0 m) location in the gradient intriguing and indicative, without exploring possible explanations in more detail here.

Figure 1. Units are missing from the x and y axis in all three plots, as well as from the color plot in figure 1a.

Response: Thanks for the comment. I suggest to add text to the caption of Figure 1 to indicate the use of UTM coordinates on both axes and add a unit to the color scale.

Figure 4. Over what time scales the friction velocity has been calculated?

Response: Over a 30 min time scale. **The averaging time has been added to the caption.**

Figures 3,4 5. I suggest changing the color scale in Figure 3a-c, Figure 4a, Figure 5 a-b, the colors are going to be very difficult distinguished from color blind people.

Response: Thank you for the suggestion. The colors were picked using recommendations for color blindness safe color scales (see, e.g., Colorbrewer by Cynthia Brewer). An online Daltonism simulator reveals a diverging gradient with distinguishable colors between blue/purple and yellow. I suggest to leave this aspect of the figures unchanged.

Technical Corrections
Line 150: “or were more” should be changed to “or more were”

Response: Thank you for noting this, I agree with the suggested correction.

The correction was made as suggested by the reviewer

Line 177: The acronym TED is explained only in line 370. That explanation should be moved here.

Response: Thanks, I agree that the description for TED should be moved from the appendix to the methods.

The correction was made as suggested by the reviewer

Line 179: change “though-out” with “through-out”

Response: Thank you for noting this, I agree with the suggestion.

The correction was made as suggested by the reviewer

Line 201: I suggest change the “it is assumable” with “it is assumed”

Response: Thank you for noting this, I agree with the suggestion.

The correction was made as suggested by the reviewer

Line 242: “eight” -> “eighth”

Response: I agree this should be corrected.

Line 244: “nine” -> “ninth”

Response: I agree this should be corrected.

The correction was made as suggested by the reviewer

Line 332: “data are” -> “data is”

Response: Thanks for the suggestion. As `data` is the plural form of `datum`, it is used accordingly here. No change was made.

Line 425: There is one extra dot before “1”

Response: This should be corrected. Thanks for the comment!

Corrections were made to the citation.

Line 467: “ 17, 0”-> “17: 1-17 180060”

Response: This should be corrected as suggested. Thanks for the comment!

Corrections were made to the citation.

Lines 503, 516, 521, Please check that all the details in the references Petrides 2011 Sayde 2015, Selker 2006 are written correctly.

Response: Thanks for the comment! I’ve made note to check the convention for Dutch names and the page number formats.

Corrections were made to the listed citations.

Reviewer #2

Response: Thank you for your effort reviewing the manuscript and providing helpful discussion. A response to the questions and comments can be found below.

The manuscript presents a novel approach to combine distributed temperature sensing and thermal imaging instruments to study near-surface turbulence. The new technique enables detailed spatio-temporal analysis of both scale and shape of temperature structures and opens new opportunities to advance micrometeorological research. The manuscript is well written and provides a detailed overview of various aspects related to the application of the new technique. In my opinion, it might make the manuscript more accessible if the data science techniques would be briefly explained in the main text (and not only in the appendix).

Response: I fully agree with the reviewer that a restructuring of the text regarding the data sciences techniques would be helpful to the reader. **The relevant sections were updated and restructured.**

It would be also helpful to explain the meaning of variance events since they are at the centre of a part of the analysis. What is their physical meaning?

Response: It is valid to ask what the physical meaning of a variance event is; if there is momentum involved, what processes drive the observed patterns. In this manuscript, a variance event is a significant deviation from a temperature background signal (and noise floor) of less than a few minutes. Their physical meaning can often be explained from context, but not always. For instance, temperature ramp series during unstable conditions have been reported and interpreted from experimental data and fluid dynamics model simulations, often linked to a rolling mode of coherent propagation near the surface. However, single events during unstable conditions can be more elusive in nature. Events coinciding with substantial local destruction of stable stratification may be driven by a non-local process, e.g., waves generated remotely. I refer to AC2 (https://doi.org/10.5194/amt-2020-500-AC2) for an overview of multi-scale observations from instruments that were operated to provide such additional context. Explaining the physical meaning of the events and their impact would be more meaningful when those observations are included in analysis, i.e., following up on this study.

I would also find it helpful if the potential of this technique for long-term monitoring would be discussed. It appears if the instrumentation was only deployed during an intensive measurement campaign. How realistic is it to deploy these instruments year-round?

Response: I agree with the reviewer that a discussion on long-term monitoring would be helpful, but it would be somewhat speculative as it was not the aim of this experiment.

Long term monitoring would be possible. I think this primarily depends on the stability of the support structure used to suspend the fibre-optic cable. In addition, it would be important to prevent accidental damage by animals, particularly wildlife. Precautions can be as simple as increasing the visibility of the set-up during the night with a floodlight and marking the area with bright warning tape. Fibre-optic cable can deteriorate under mechanical stress, but I have not seen evidence thereof based on the used set-up. However, optical cable can be repaired in the field in case of damage or, if the budget allows, replaced. A custom reel of fibre-optic cable cost approximately 500 USD/EUR in 2014. The instruments, particularly the TIR and DTS models used in this study, are designed for long-term (industrial) operation.

Details were added to section '2.5 Design considerations'.

Please see below some comments:

Line 17: Please clarify "about additional details contained in such data". What were the specific research questions that were addressed in these studies? It would also be informative to further elaborate to which research questions the presented new measurement techniques could contribute.

Response: Thank you for the comment. I agree with the reviewer that the formulation is vague.

There is a tangential connection between this manuscript and the cited studies. These studies aim at separating or computing component scalar fluxes, and hinge on (cross-)correlation structure between scalars. Conditional sampling approaches for heat fluxes (e.g., Klosterhafen et al and the works cited therein) rely on separating 'high' and 'low' frequency variance at the height of typical eddy covariance observation, assuming that the isolated scalar signal (water vapour, carbon dioxide, temperature) observed at some distance away from the surface are a result of coherent exchange, which in turn is driven by coherent shapes and scales that preserve the signature of processes at the surface. Here, an approach is presented to observe the spatio-temporal evolution of such coherent structures (temperature). An outlook could be to use that detailed information to improve said alternatives to the eddy covariance technique. That outlook is expressed later in the text, specifically line 257: 'The ability to trace coherent motion in space and time may proof useful for the development of conditional sampling methods that complement the eddy-covariance technique.'

I think it might go beyond the scope of the manuscript to review the various approaches that are in development to quantify vertical scalar fluxes or derivative variables, in much more detail than the first two paragraphs (line 17 to 34).

My suggestion is to replace 'about additional details contained in such data' (line 17) with 'about additional details contained **within the scalar (co-)variance data also used by the eddy covariance technique**'.

Line 128: Which criteria were applied to determine if eddy covariance flux computations were "acceptable"?

Response: Thank you for the comment. I agree that this needs to be clarified. The eddy covariance computations followed a (self-)validation conundrum, e.g., based on tests for stability and stationarity. The tests results are simplified in a classification as Acceptable, Ambiguous and Incorrect. The quality classification is shown in the top three bars of Figure panel 4f. Here, 'acceptable' indicated that data classified as Incorrect were excluded from the presentation, such as Figure panel 4e. Please note that all computed eddy covariance flux data, correct and incorrect, are included in the support material data sets in case these are needed.

Following up on RC1, I suggest to rephrase the sentence to 'eddy-covariance flux computations **rarely produced acceptable (classified as Acceptable or Ambiguous) results**...'

Line 156: Please elaborate how these findings "suggest an interaction between scales".

Response: I agree that this could be elaborated. The simultaneous occurrence of multiple scales at the same place (height) and time suggests interaction, perhaps cascading energy from one scale to another.