Review of « VELOX - A new thermal infrared imager for airborne remote sensing of cloud and surface properties », by Michael Schäfer et al.

General comments

This paper presents a new thermal infrared (TIR) imager called VELOX aimed at measuring cloud and surface properties, primarily developed to be used onboard the High Altitude and Long Range Research Aircraft (HALO). VELOX is adapted from a commercial TIR imager and measures brightness temperature in 6 channels with 10 m spatial resolution at the surface when flying at 10 km. First the instrument hardware and technical characteristics are described, then the calibration procedure, which includes several successive steps, is detailed. The performances of the instrument in terms of noise-equivalent temperature difference (NETD) and absolute accuracy are assessed using a temperature-controlled blackbody reference. Finally, measurements acquired during the EUREC^{4}A campaign in 2020 are used to illustrate the capabilities of the instrument, with first hints to the derivation of a cloud mask and cloud top altitude from measured brightness temperatures.

The paper is well organized and well written, and perfectly fits with the scope of AMT. However, given that the focus is on the instrumentation and on the data processing, more details would be appreciated to clarify several technical points that so far remain somehow unclear. Only once these important clarifications have been made the paper could be considered for publication.

Specific comments

1) As very similar sensors have been (or are still) used in airborne configuration, it would be useful to highlight the specificities of VELOX. In particular, does the configuration respond to specific requests that no existing instrument would match? Do the performances enable improved retrievals?

2) As explained in the text, the successive images acquired by VELOX largely overlap. It is not clear whether this massively redundant information is useful or whether using larger integration times could advantageously improve the accuracy of the measurement. In any case the chosen acquisition configuration would deserve more justification.

3) In general the description of the calibration procedure is very qualitative, making hard for the reader to really guess what is practically done. Replication of the procedure would probably be quite difficult. More details (in particular for the correction of the window impact) would be helpful. Adding equations to more explicitly describe the successive steps would certainly help as well. Several suggestions are made in the technical corrections.

Technical corrections

l.10: analysis

l.22: is estimated wrong → estimation is wrong

l.23: not clear what *polar orbiting satellites* refers to. Is the sentence valid only for such orbits?

l.28: *comparable higher* sounds contradictory

l.31: information on overpasses should be merged with that at l.23

l.33: can you detail why MSI will be better than current sensors? More channels, higher spatial resolution?

l.35: *kilometer range* sounds similar to MODIS. Is it actually coarser?

l.51: *dominated by* is unclear. Radar are indeed not sensitive to small particles, but lidar are sensitive to large ones (although few large particles may reflect much less radiation than many smaller particles). Radar are mostly useful where lidar signal saturates

l.53: *different vertical weightings* is unclear. Are you talking about the profiling capability?

l.55: dominated → driven, governed?

l.63: "3D effects" also take place in the TIR. Maybe clarify what effects typical to the SW are avoided

l.65-68: this paragraph is not very clear. What is challenging? To measure a temperature difference, to know perfectly the reference (background signal)? Knowing the detector temperature is not enough, the whole instrument contributes to the measured signal. What is *not always given*?

l.69: when presenting a new instrument, it's useful to point how it differs from existing available (sometimes commercially) instruments, here or later on in the manuscript

l.73: repetition of reference for EarthCARE

l.93: what's the size of the filter wheel? Of the filters? As they do not appear in Fig. 1 I assume they're quite small

l.94: 100 Hz is for a complete rotation or to go from one filter position to the next? Clarify the link with the 100 Hz acquisition. A single frame on each filter or integration of multiple frames on the same filter? State here that the full measurement on all filters takes 0.06 s, this is a major information.

l.95: *partly adapted* is unclear. Are the filters meant to match MODIS filters characteristics or only the central wavelength?

Table 1: what is the temperature reference for the NETD? Could you provide more details in the text about the cooling of the sensor (temperature, stability etc.)

l.102: is the spectral response of the detector really zero outside of the range 7.7 – 12 microns?

l.102: maybe state here the reasons (if any) for duplicating this broadband channel

l.144: I'm not sure to get what the first issue is. Is it to project the pixels at the Earth surface? Is the aircraft movement used to tackle this, or just the position (including attitude)? Does it work when the emitter is not the surface, but a cloud?

l.169: can the difference in acquisition time for different filters be an issue as well? To be related with the distance traveled by the aircraft between two successive filters

l.175: this scene identification/matching deserves more details. What kind of algorithm is used?

l.189: In a system without on-board calibration, this calibration procedure is crucial. Can you provide more details on the way the corrections were obtained (lab experiments to isolate the impact of temperature changes?). At least consider referring to a paper detailing how this is achieved.

l.205: it is not clear what the link between *gain/offset* and *variable ground potentials* is. Once the non-uniformity of the pixels is identified, how is the amplitude of the correction determined? Why are stripes removed with this procedure? Do these stripes come from pixels with different gains or different offsets (due to straylight for instance)?

l.207: does the scene need to be homogeneous to apply this calibration, as stated above?

l.214: I don't understand why this specific calibration is not implemented directly at the step 3.2.1. Practically, is the correction pixel-dependent? Is it static or does it depend on environmental conditions?

Eq. 1: how where the different parameters of this equation determined? Was the method validated by cross-calibration against a black-body?

l.268: what is *accuracy* here? Absolute accuracy detailed just below?

l.297: how can you know that no cloud-free ocean was observed?

l.300: are the differences between simulated and estimated cloud-free BT due to differences in atmospheric state, or could they be solely explained by measurement uncertainty? The differences should be compared to measurement uncertainty on the one hand, and to simulation variability on the other hand. Are the points away from the 1:1 line actually those acquired far from a dropsonde?

l.317: what is the interest of such a comparison with pushbroom configuration, if the obtained differences are not better described?

l.347: for this, is the average of 10x10 pixels used or it is performed for individual pixels? Does the maximum envelope comes from the time series of individual pixels or from a single image?

l.368: here it is somehow assumed that the cloud is optically thick and that the emission comes from the top of the cloud. Can you discuss a bit these assumptions and their limits? Would changes in LWC or r_{eff} make a difference on the emissivity fixed to 0.99?

l.376: 470 m offset seems huge for a cloud mostly ranging from 600 to 1400 m. Can it really be explained by errors in actual atmospheric profile? How does an error in BT translate into an error in cloud top altitude, roughly (for the atmospheric profiles observed)?

l.378: can you detail this correction procedure since it may be critical (when errors are larger than the measured range of variations)

l.396: references to EUREC^{4}A not needed here

l.408: typo: *Oceanc*

l.433: how is set ocean emissivity?