Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2017-218-RC2, 2017

Atmospheric Measurement Techniques Discussions

Interactive comment on "Characterisation of the artificial neural network CiPS for cirrus cloud remote sensing with MSG/SEVIRI" *by* Johan Strandgren et al.

Anonymous Referee #2

Received and published: 22 August 2017

The authors have done a very good job at characterising the performance of CiPS. I am impressed by the depth. The only section where I note some problems, is the final section before the conclusions, on uncertainty estimates and propagation. That one may be harder to get right than it seems and is not essential for the paper, so perhaps the best option here is to skip it. Regardless of that section, a list of suggestions for a minor revision follows below.

page 1, line 3: replace "implemented" by "modelled". Only nature can implement physics!

page 1, line 20: what do you mean by "physical implementation"? The implementation

of the physics in the model?

page 2, lines 11-12: the formulation here is a bit sloppy. The radiometer itself does not have a spatial resolution or a vertical component. The spatial resolution comes from either a scanning mechanism or a detector array. The ability to resolve vertical information from passive sensors depends on the availability of multiple channels with different weighting functions, but not in clouds. I would rephrase this sentence as: "Imaging radiometers typically view a large area (by scanning or otherwise) to observe complete cloud systems, but a passive infrared sensor cannot resolve cloud features vertically and has a limited sensitivity to thin and sub-visual (visible optical thickness < 0.03) cirrus clouds." or similar.

page 2, line 16: the active nature does not intrinsically lead to a poor spatial coverage, but is a consequence of other design considerations (scanning radars exist in space and certainly on the ground). I would replace "leads to ... between orbits" by "those sensors have a small footprint and observe only at nadir, which leads to a poor spatial coverage."

page 2, lines 19-31: ANNs can indeed exploit collocations to train a retrieval, but that property is by no means unique to ANNs. Any type of machine learning can do, whether it is a basic linear regression, a neural network, a support vector machine, or others. You need to add a couple of lines here, pointing out that a collocation database can be constructed with most pairs of orbits and that this can be used to train a retrieval. For cloud retrievals, ANNs have proven to work quite well to perform this training in practice.

page 3, line 22: replace "brightness temperature" by "brightness temperatures" — you're using multiple channels

page 4, lines 6–7: you could add a small figure showing a map of the SEVIRI disc.

page 4, line 12: those radiometric noise levels are design specifications and not ac-

tual measured noise levels. Actual noise levels can be derived from on-board level-1 measurements by looking at the dark corners of the disc and then propagating the uncertainty in counts through to radiance, reflectance, and brightness temperature units. If you are going to use those noise levels for uncertainty estimates you need to be aware they can be quite far from actual noise. If you are not using them I would not report them, as it just propagates a number that many people are using incorrectly.

page 5, line 3: IWP is not a layer product but a column-integrated product. Should this perhaps be IWC or pIWP (partial column IWP)?

page 9, Figure 1: although I think the figure looks beautifully crafted I'm not sure if it's the most effective way to visualise the weights. At least, the colour for OPF is too similar to the colour for IOT&IWP. But in general, I believe the information would be easier to read in a tabular format, with 4 columns (for the products) and 17 rows (for the input variables), writing the weight as a number in each table cell, and perhaps colour-coding by the value. The downside of the figure is that the connecting line appears to give an implied meaning to an essentially meaningless ordering; it may be hard to read values when they are very close to each other; and the relatively long tick labels necessitate alternating them between the bottom and the top, which adds to the confusion (if the authors rotate the labels they could have them all at the bottom). A table might work better. Yet another way might be a flow diagram, where the seventeen inputs would be written below each other and the thickness of the line connecting to each input would be proportional to the weight (those can again be labelled), although I'm not sure how cluttered the visualisation might become when four output "flows" are visualised in the same diagram.

page 9, Figure 1: why are all weights positive? Is this a standard property of the ANN, did the authors impose it, or is it a coincidence? In particular for the DOY_SIN and DOY_COS, which after all can take either positive or negative value, it is not obvious to me why it should be.

page 10, Figure 2 / page 11, lines 9-10: is there any permanent ice & snow in the SEVIRI disc? The line for "permanent ice and snow" looks rather noisy. Is that just the strip of Greenland barely in the field of view? How many pixels are those? The POD appears remarkably good over permanent ice and snow but I wonder how significant the results are.

page 10, line 7: Are all barren surfaces in the field of view bright, hot deserts? I wonder if there are any dark barren surface types to test against? Maybe Iceland? Might that show a better performance?

page 11, lines 14-16: I understand why mixed phase or supercooled liquid clouds complicate the cirrus detection, but why is this (independently) the case for temperature inversions, that should be a boundary layer phenomenon anyway?

page 15, line 19: I'm confused by the reference to Strandgren et al. 2017 here. Surely those are the classes that have been discussed in the previous paragraph of the present paper? Why the reference?

page 15, line 28: 97%... is this shown? If yes, where? If not, please state that it is not shown so the reader does not needlessly look for a figure showing this.

page 17, line 9-10: this is not accurate. CALIOP flies in a sun-synchronous orbit and thus only samples a very limited part of the diurnal cycle. The statement that it represents the "natural" distribution is too strong.

page 17, line 15: replace "error" by "uncertainty". You cannot retrieve an error; the error is the difference between truth and retrieval. The uncertainty is a statistical estimate of the distribution of errors. The uncertainty can be estimated, the error is only known in an artificial scenario.

page 19, line 25-27: again, the authors should be aware that those are most likely not actually measured noise levels, but rather design specifications. If the authors are sure that the former is the case this should be clearly stated, because the design design

specification for radiometers is often misinterpreted as an actual noise level.

page 19, line 32: personal communication with a large institute seems somewhat unusual, is there a name attached to this person?

page 20, line 1-2: I think the authors need to add a caveat that although they obtain a number that varies per-pixel, this is in fact not a metrologically traceable per-pixel uncertainty. Determining the latter is possible (EUMETSAT and others are part of an effort to do this for MVIRI, see www.fiduceo.eu) but a lot of work. As currently phrased, this paragraph is at risk of misleading readers into thinking there are true per-pixel uncertainty estimates.

page 20, line 17-18: the ANN is essentially a set of equations, all of which are differentiable. Therefore, it should be possible to directly apply the Law of Propagation of Uncertainties. Have the authors considered calculating this propagation directly, instead of via an ensemble?

page 21, line 10: what is the resolution / digitisation level of SEVIRI at those brightness temperatures?

page 22, line 17: although I believe this to be accurate, I am not convinced the evidence presented by the authors is sufficient to show this. I suspect it may actually be smaller.

page 30, figure 5: I opened this figure with two different software packages but in both cases it seems something is wrong. I see only panels 1, 2, 3, and 5 labelled, at the bottom I see "N 13", "N 1 03 2", "N 1", "N 135", "N 2", with varying amounts of whitespace. It appears some labels have gone missing.

page 30, figure 6: I find it slightly confusing to see "transparent cirrus" with IOT up to (and possibly exceeding) 4. Perhaps a different word would be more suitable (but I don't feel too strongly about this).

page 33, figure 9: There is some evidence that 2D-histograms with hexagonal binning are superior to those with rectangular binning; the authors may wish to search

the internet for the keyword "hexplot", try to show the data or a hexagonal grid, and judge for themselves. An article illustrating why it may be superior can be found at http://www.meccanismocomplesso.org/en/hexagonal-binning/

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2017-218, 2017.