

Review of the paper “Improved cloud detection for the Aura Microwave Limb Sounder: Training an artificial neural network on colocated MLS and Aqua-MODIS data” by Frank Werner et al.

General comments

This paper nicely illustrates that the implementation of machine learning to MLS cloud classification leads to an impressive improvement in MLS cloud detection, compared to current operational techniques.

The paper is concise, well written, and discusses well-selected calculations. The discussion of both global statistics, and individual cases, is very appealing. The Summary and Conclusions section is very well written.

The discussion of the machine learning methodology is very concise, but could benefit by briefly defining some of the machine learning terms which may not be familiar to the atmospheric science research community.

The paper should be published following very minor suggested changes.

Specific comments

The use of machine learning techniques and terminology is likely unfamiliar to many in the atmospheric sciences. There are several places in the text in which a few additional words / sentences could help the reader understand better what is being done by the authors. There are some terms which need to be defined. Please discuss, for example, what is meant by “feedforward” on line 121. Other terms that should be defined (briefly discussed) are “imbalanced classes”, “learning rate”, “Nesterov momentum value”, and “weight decay”.

Technical comments

Line 21 the phrase “cloud amount” is vague. Please be more specific.

Line 46, add commas, revising to e.g. “radiances, from lower in the atmosphere, and smaller downwelling radiances from above, into the MLS raypath” to improve readability. In my first reading of the sentence I had a hard time making sense of the sentence.

Line 55, what is meant by “discount them” ?

Line 89, please specify Figures in Waters et al 2006 or other papers that illustrate the spectral sampling details of the AURA MLS experiment, so the reader can obtain a fuller understanding of the MLS experiment.

Line 130. It would be helpful to point out that Figure 1 is presented for illustrative purposes, since line 253 later points out that each hidden layer has 851 neurons (instead of 2 neurons).

“Figure 1 illustrates the general setup of a simplified multilayer perceptron that contains four layers, and is instructional. The full model setup is discussed in Section 3.4”

Line 168. Is the MLS aggregation at $1^\circ \times 1^\circ$ because the MLS data sampling is (line 100) near 165 km?

Line 173 are the 5,000 samples MODIS, MLS, or MODIS-MLS samples?

Line 262. Approximately how many epochs are calculated?

Line 318 clarify what is meant by “classification going forward”.

Line 549. If the current MLS data version is V5, why not include the new ANN capability in the V5 product instead of “future versions of the v4.2x” product?

Standard Criteria

1. Does the paper address relevant scientific questions within the scope of AMT? Yes
2. Does the paper present novel concepts, ideas, tools, or data? Yes. The discussion of machine learning techniques is relatively new to the atmospheric sciences, and very appropriate.
3. Are substantial conclusions reached? Yes, the machine learning technique makes an impressive improvement in MLS cloud detection.
4. Are the scientific methods and assumptions valid and clearly outlined? Yes
5. Are the results sufficient to support the interpretations and conclusions? Yes
6. Is the description of experiments and calculations sufficiently complete and precise to allow their reproduction by fellow scientists (traceability of results)? Yes
7. Do the authors give proper credit to related work and clearly indicate their own new/original contribution? Yes
8. Does the title clearly reflect the contents of the paper? Yes
9. Does the abstract provide a concise and complete summary? Yes
10. Is the overall presentation well structured and clear? Yes
11. Is the language fluent and precise? Yes (with only one or two exceptions, noted above)
12. Are mathematical formulae, symbols, abbreviations, and units correctly defined and used? Yes
13. Should any parts of the paper (text, formulae, figures, tables) be clarified, reduced, combined, or eliminated? A few clarifications, noted above, are suggested.
14. Are the number and quality of references appropriate? Yes (with perhaps one place to add specific reference to figures that illustrate the AURA MLS spectral territory).
15. Is the amount and quality of supplementary material appropriate? Supplementary material is not included in this paper