

Interactive comment on “Statistical downscaling of water vapour satellite measurements from profiles of tropical ice clouds” by Giulia Carella et al.

Anonymous Referee #1

Received and published: 20 February 2019

This review is about the article “Statistical downscaling of water vapor satellite measurements from profiles of tropical ice clouds” by G. Carella.

General comment: In the beginning, I was impressed by the approach. Downscaling water vapor from sounder observations could really bring a change to the sounder game. And using microwave sounder information for this is a really promising approach. The article was fluent to read and easy to understand in most parts. The mathematical approaches and the reasoning for modifications are usually well described. However, I also got a little disturbed after I worked through all the fancy mathematics. My main problem: Physically, I do not see a correlation between the SR values from

C1

CALIOP and the layered RH in the atmosphere. So, for me, to a first degree you are applying “magic”, covered in – I admit very interesting – mathematical approaches to derive a RH-profile. Ok, do not get me wrong. It is an interesting downscaling and I am still impressed by your approach. I just want to understand some things. I will go through the text chapter by chapter to explain.

1. Introduction: Could you please add, why did you choose CALIOP? For me, there is no physical reason to connect these two pieces of information. CALIOP measures SC at 532 and 1064 nm, which has no connection to water vapor. All information in SC is from cloud particles – mainly ice, in your case. You mention a few articles to infer correlations, but you do not really point out or cite the physical reasoning from there. Please explain, why you think CALIOP is a good choice for water vapor information. I know, your regression models do not need a relationship. But you had your reasons to connect a microwave sounder and a lidar, didn't you?! Because for me, this is not an obvious choice. So please reiterate, why you think this would lead to a physically correct downscaling.

2. Data

Please explain, if the product by Brogniez an official product. Is there an official webpage, source ...? Same with CALIOP. Did you use the official product? It seems like, but I want to make sure.

3. Methods

I consider the last sentence in 3.1 is crucial for justification of your technique. You should explain this a little but more, perhaps with more citations from Schroeder's or other papers. It indicates a connection between RH and SR, something which is very important for your approach. But I am not quite sure, how to understand chapter 3.2. You have already a cloud classification from CALIOP (right plot), so why did you prefer your own k-mean method? Both tell you, that clouds above 10km are ice, which is not really big news. You could have just used that value or the CALIOP classification, so

C2

why do you insist to do your analysis based on this extensive k-mean clustering?

Formula 1 in chapter 3.3 is my biggest problem: it assumes RH_I is connected to (SR₁, SR₂, ... SR_p) via a function. That seems to be the foundation of your idea. But for me, there is none. At least, no physical connection. Is it enough for this approach to find a correlation without reason? I am ok with that, but the results would be of limited use for research (see conclusion). Please re-iterate more here – or at least in the introduction, perhaps based on the articles by Brogniez or Udelhofen-and-Hartmann.

Choice of regression model:

I miss some important information: I understand your limitation to ice clouds. How much data do you use for the regression training with respect to the mission time frame? Do you use specific dates? Do you use the same amount for all regression models (RF,QRF,GMR1,2,3). Did you make tests with different amounts/dates? Where the results always the same?

Also: Do you deal with the error of the RH-retrieval in RF and QRF? If you look at Figure 2, there are lots of layers with uncertainties > 30%, especially below 500hPa. (Remark: You might want to choose a different color scale there, it is really hard to understand. Everything above 30% is the same color ...). Retrieval tend to get worse closer to the ground. Do you deal with it differently?

I try to understand, why you would chose so many variations of the GAM approach. There is no reason to assume a linear connection between RH and SP, so RF and QRF are quite reasonable to me. But here you suddenly force a linear connection. Is it just for comparison? Because it seems to do bad anyway, when I look at later results. Please re-iterate the reasons for this selection and the two derivatives (GMR1 and geoadditive). Are there other options?

Chapter 4: Actually, p.10, line 18-21 is another short talk about a possible physical connection between RH and SR. If you could extend this a little bit more, especially in

C3

the Introduction, then the approach would be much easier to understand. I consider 3 GMF approaches, which have all bad skills, a little bit redundant. I would rather see a third different approach than 3 similar fails. But ok, if you want to keep them, that is fine too.

I also have problems to understand Figure 5. Is the predicted from RF? And is the observed RH the one from 10x10km SAPHIR? The description in the text is very short and confusing. Please explain more here: source of predicted, source of observed. Why would you then have such a bad correlation for L6? Please explain this plot in more detail, it seems it is your only source of verification for your approach. Most people would prefer an independent source (radiosonde, airplane observation, ...), but I guess you don't have enough data for this in the Indian ocean. So, you have to convince the reader about the "success" of your approach with this plot. Honestly, I didn't get convinced, you didn't write enough.

Chapter 5:

At the moment, I am questioning some bullet points in your conclusion. I am not quite convinced that your data can help "study ... small scale water processes" or "evaluate ... water vapor interactions". You need to convince me, that you have a physical foundation, not just correlated sorting. On the other side, I agree that you can always "evaluate small scale inhomogeneities" in reanalysis or "guide parameterizations". Models need to know the behavior of parameterizations on smaller scales, so you might be very helpful to find out scale breaks on scales around 100 m. I also think, you should talk a little bit more about the extension to other clouds. It sounds interesting, but based on your requirements (homogeneity, strong SP signal), you might be in trouble. If you could talk more about future possibilities and obstacles, it would be a better selling point for this article. But that is more my opinion...

Minor comments, found during reading: p. 2, line 3: Should be "state-of-the-art" p. 2, line 31: I am not quite sure, what "space clouds" are ... p. 5, line 10: should be "nadir"

C4

p. 6, line 1 : is the "1" necessary here? p. 10, line 2-4: this sentence is hard to read with all the comma and brackets. I would propose to redo it a little bit. p. 12, line 16: should be "CALIPSO"

Still, I consider it a very interesting article. My Regards.

Interactive comment on Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2018-138>, 2019.