

Interactive comment on “Estimating fractional snow cover from passive microwave brightness temperature data using MODIS snow cover product over North America” by Xiongxin Xiao et al.

Anonymous Referee #2

Received and published: 5 April 2020

Overview and General Comments

This manuscript describes a new approach of estimating fractional snow fraction from satellite-based passive microwave (PM) sensors and higher resolution MODIS snow cover estimates. The authors present different regression and machine learning type algorithms, including multi-regression, artificial neural networks (ANN), and a random forest regression technique, for estimating the PM-based snow cover fraction using the MODIS snow cover as a reference input to the algorithms along with accounting for different PM retrieval and ancillary datasets, like vegetation types. The methods

C1

are demonstrated and validated against independent in situ measurements across the region of interest (Canada and the US).

Overall, the paper includes comprehensive descriptions of the data and methods used, and detailed background and justification for the work presented. It also is within the scope and appropriate for the journal, The Cryosphere. The supplementary material does help support the overall findings in the paper. However, some of the methods and conclusions may require some revision and may not be conclusive enough as there is a limitation on the years evaluated and the wintertime period focused on. A few major and minor comments are noted in this review that hopefully help to strengthen the paper and the organization of the methods and results presented. There are a few sections that were difficult to follow and some of the English grammar and syntax was unclear.

One downside to this study is that the authors only focused on seven years of available passive microwave and optically based snow cover observations and then just the peak snow months of January and February. Though it seems to make sense to focus only on when the snowpack is at the peak months and more spatially continuous, however, it is also worthwhile to capture the temporal and spatial heterogeneity in the accumulation and ablation seasons and more fully test the algorithms described and applied in this study. Otherwise, the algorithms are only somewhat effective for peak wintertime in US and Canada and not applicable for studies, like prescribing observational snow cover conditions in climate projection or snow-land-atmosphere climate interaction studies, which are pointed out as one primary reason to perform this present study.

Also, in relation to the timeframe of the training and validation data years, only having one year to perform the validation seems quite limiting, as a given year can be hard to note overall performance given snow cover can vary greatly from year to year (e.g., snow drought conditions). This is somewhat reflected in Figure 7 (right column panels), which show how highly variable and not as predictable in the validation year (2017).

C2

Please explain why a longer period of record is not used, e.g., 2002-2019 (Terra+Aqua MODIS combined) and the passive microwave combined product by Brodzik et al. (2018), to perform the training and validation period. Perhaps, use Water Years (WY) 2002-2013 for training and WY 2014-2018 for validation?

Using only one year for testing and a second year for validation is very limiting for this study, and it is highly recommended for additional years to be included. Also, for the four different approaches of estimating the fractional snow cover from passive microwave should have longer evaluations performed in this context as the summary of the results would be inconclusive for one year of validation.

Some of the methods sections are hard to follow, though the authors provide many details there and in the Supplemental material. For example, Section 3.3.1 of “Selecting input variables” was at times hard to follow and why each scenario was selected. Improving the organization of the sections to flow better in terms of their logic and why different experiments were performed would be helpful for the overall background and discussions of this study.

The English grammar and syntax used require additional review and editing by editorial services to help correct these issues before resubmitting. A few suggested corrections are offered below in the technical corrections section.

Specific Comments

Abstract: The authors introduce “Grody’s snow cover mapping algorithm” towards the end of the abstract without any other background. Perhaps they could provide one introductory phrase on this algorithm within the abstract to give more context.

Page 2, Lines 9-10: The authors mention that snow cover data from station measurements are “time-consuming, [and] cumbersome,”. What do the authors mean by these adjectives? Please clarify here. Any dataset, including satellite, requires time and careful derivation of the final product. However, in situ snow cover data are spatially

C3

discontinuous and require more time to maintain.

Page 6, lines 11-12: Would like to point out here that North America includes Mexico as well. The authors should specify that their study domain spans the continental U.S. and Canada only.

Page 7, lines 10-11: Authors state here that “to the best of our knowledge, there are no researchers have developed fractional snow cover . . . using passive microwave data.” Please take a look at the following references and cite appropriately:

Foster, J.L., D. K. Hall, J. B. Eylander, G. A. Riggs, S. V. Nghiem, M. Tedesco, E. Kim, P.M. Montesano, R. E. J. Kelly, K. A. Casey and B. Choudhury (2011): A blended global snow product using visible, passive microwave and scatterometer satellite data, *International Journal of Remote Sensing*, 32:5, 1371-1395, DOI: 10.1080/01431160903548013

Page 8, lines 24-27: It would be helpful here to provide a lead in sentence to introduce your first two equations.

Page 8, last line: “Calculation areas should be in a larger feet . . .” What is meant here by “feet”? It does not seem to make sense to use this word here, but perhaps “footprint area” makes more sense? Please correct.

Page 11, lines 2-3: MODIS Collection 5 products are considered older and not “current”, as they have been replaced by Collection 6. Recommend removing “current” here.

Subsection 3.4.1: The authors discuss both the linear and multi-linear regression methods here, which makes the discussion confusing to follow. They then have the reader refer to the Supplementary material for more information. It is recommended that the authors better describe in this subsection how the “linear regression” is applied. Was it based on the equations in Salomonson and Appel (2004) or new linear equations and parameters derived for the four different vegetation categories? Please try to better

C4

organize and explain this linear method in this subsection.

Page 14, lines 10-11: Please provide citations and references where possible for the metrics, especially Cohen's kappa coefficient and the F1 score.

Page 15, Lines 11-12: Authors indicate here that their "Scenario-6" variable sensitivity case "generated the worse performance, with the low R, the great MAE and RMSE". When looking at Table 4 results, Scenario-6 appears to perform rather well overall. Perhaps it would help if the authors specify here that of the Scenarios of 1, 4-5 and 6, Scenario-6 performs the "worst". It is also recommended to change the last part of that sentence to: "this scenario's setting had the third worst performance with lower R values and higher MAE and RSME values."

Page 15, line 31 to top of Page 16: Make "Figure" plural and change the last part of this sentence to something like: "show that this finding was not coincidental." This sentence is a bit hard to understand in what is meant by "not coincidental". Please elaborate or better explain the meaning here.

Page 16, line 29: Please clarify here what is meant by "neglected to assess the rationality of estimated value . . .". Are you referring to the out-of-bounds events that occur in the other methods, other than the random forest approach and that that "rational" was not well checked?

Page 21, line 1: Authors state that only a few studies validate the accuracy of MODIS snow cover products in forested areas. Actually, there are several in addition, including:

Arsenault, K.R., P.R. Houser and G. J.M. De Lannoy, 2014: Evaluation of the MODIS snow cover fraction product, *Hydro. Proc.*, 30, 3, pps. 980-998. <https://onlinelibrary.wiley.com/doi/full/10.1002/hyp.9636>

Kostadinov, T. S., and T. R. Lookingbill, 2015: Snow cover variability in a forest ecotone of the Oregon Cascades via MODIS Terra products, *Rem. Sens. Env.*, 164, pps. 155-169. <https://www.sciencedirect.com/science/article/pii/S0034425715001303>

C5

Page 22, lines 14-16: The first statement here about the "strong limitations in the understanding of physical mechanism" is a bit hard to understand. Are the authors referring to the underlying physics and characteristics that relate the fractional snow to the signature of the passive microwave bright temperature responses? Perhaps, it might be better to frame these concluding statements more in that way vs. "mechanisms".

Table 1: In the row of references, does the Xiao et al. (2018) paper cover both Scenario-4 and -5 columns in the table? If so, it might be helpful to specify this in the body of the paper.

Figure 8: In panel A, more binary MODIS snow cover present (e.g., large green pixelated areas in Canada), but that does not seem to get translated over to panel B for the fractional MODIS snow cover (mostly filled in with no fractional values). Please explain why most of the derived MODIS snow cover fraction is removed here, especially over Canada? Also, for the MODIS snow fractional product, there is no fractional snow representation between 0.3 and 0.8, the other two categories shown in panel B. What is happening here in that regard – no fractional snow within 0.3 and 0.8 at any noticeable gridcells? Please provide an explanation in the text as well.

Finally for Figure 8, it would be helpful to assign a different color and category for the non-snow pixels (at fractional value of 0.) in panels B and C to better discriminate the non-snow areas from the snow-based areas. Currently, snow-free pixels are lumped in with the low snow fraction category of 0 to 0.3.

Fig. 11: This is a nice figure that summarize and present these results well.

Technical corrections

Page 2, line 25: Please specify what "FY" stands for in "FY series sensors".

Page 3, line 25: Awkward phrasing here: "To unite resolution, . . ." Perhaps try: "To be at a common resolution, . . ."

Page 5, line 7: Recommend here to separate the two phrases here with either a semi-

C6

colon (between “collected” and “all available”) or place the conjunction “and” after the comma.

Page 6, line 6: Please specify what “ETOPO1” stands for.

Page 6, line 11: Add citation and reference for “ArcGIS 10.5” software.

Page 6, line 17: Replace “heterogeneous” with the noun, “heterogeneity”.

Page 7, line 7: MODIS misspelled here as “MODSI”.

Page 7, line 31: Remove “with” before “accurate”.

Page 9, line 21: Either replace the semicolon with a period, or make the word, “Thereby”, lower-case.

Page 9, line 27: Change the “not” in this line to “cannot”. Also on that same line, the word use of “Correspondingly” here does not seem to make sense.

Page 10, line 5: Make “variable” plural here in “an optimal combination of input variables”.

Page 13, line 6: “researches” should be changed to “researchers”.

Page 18, line 5: Remove “be” before “misclassified” and change “into” to “as”. Also, please remove the phrase, “As we all know”, and change the start of the second sentence there to: “Permafrost is known to be widely distributed in the northern part of ...”

Page 20, line 4: Change “researches” to “studies”.

Page 21, lines 23-24: Change “were” to “was” in relation to “The accuracy of the proposed algorithm was further ...”.

Table 2 caption: “unite” should be “unit”, and “clod desert” should be “cold desert”.

Figure 7: The use of the capitalized and lower-case plot labels is fine but not conven-

C7

tional. Would it make more sense to simply use, “A, B” then “C, D”, etc., for the paired columns?

Interactive comment on The Cryosphere Discuss., <https://doi.org/10.5194/tc-2019-280>, 2020.

C8