

## ***Interactive comment on “Subgrid snow depth coefficient of variation within complex mountainous terrain” by Graham A. Sexstone et al.***

### **Anonymous Referee #1**

Received and published: 23 October 2016

Major comments: 1. One major issue with this paper is that the content addressing the 3rd objective stated in the introduction is far from enough. If they would like to develop a general methodology for parameterizing CVds in complex mountainous terrain, the author should test their methods on other mountainous areas, such as the Sierra Nevada (Opentopography does have a snow-on snow-off lidar dataset, and also Airborne Snow Observatory (NASA JPL) collected three years of time-series lidar data in two basins of the Sierra). But the only testing the author addressed was cross-validation, which is quite site-specific and did not provide enough evidence this method could be applied on other regions. The authors should shrink down this broader goal in the current manuscript.

2. The second issue is that I expected model selection to be a key point being addressed for a statistical modeling paper, but it is not there. The author should elaborate

C1

a lot more in detail about model selection so the reader will know how your final model was selected. And are model-selection results site-specific? Your process of selecting the right model could be great information to other researchers who are working on similar topics but in different geographic areas.

Minor comments:

Page 4, line 3-8: Although 10-m cell have been shown to influence snow-depth distributions, were spatial resolutions in previous studies the same as what you are looking at (1 m)? And how do you justify the ds statistics calculated from the 1-m resolution data could be parameterized by 10-m resolution topographic and vegetation data using statistical models?

Page 4, line 24: Why use these three window sizes and what are their effects on the TPI image?

Page 4, line 25-30: Why not use the canopy data from the lidar? And how do you account for the forest changes over time.

Page 5, line 13-14: I am really confused about this sentence. Rephrase for clarity.

Page 5, line 28-30: If the influence of ds on CVds was removed using a best-fit power function, why does Table 1 still have Ånds's Pearson's r coefficient? Are these results in Table 1 before detrending or after detrending?

Page 5, line 34-35: The reason that mean z was excluded in model testing is not convincing. As it is known that ds is highly dependent on z, why not exclude ds as well? And also why is VH not commonly available? If you have snow-on and snow-off lidar for the area you are looking at, why you do not have VH? And if you are trying to create a parameterization of CVds globally for all mountainous regions, then failing to test your models in other regions is a serious shortcoming. So from my understanding, your model is already site specific so why not explore as many variables as you could?

Page 6, line 2-4: I think that under Gaussian linear regression, Mallows' Cp is the

C2

same as AIC. And also why you are using AIC over likelihood-ratio test if your models are nested? Is it necessary to penalize the number of parameters?

Page 6, line 10-13: The way you presented k-fold cross-validation is incorrect. You should shuffle your dataset first and then iterate over the k-folds. Page 6, section 3.1: Although snow conditions are close for the two lidar campaigns, you should not blend them together for modeling, especially then there is a precipitation event between the two. Because the spatial snow conditions are not completely independent of time. Therefore, temporal effects on the snowpack will be a confounding factor for modeling the CVds when you mixing the data from two dates together. And all the causal inferences you made and regression model parameters you estimated will be confounded by temporal effects.

Page 7, line 15: What is your U-statistic for the Mann-Whitney test?

Page 8, line 3-4: What transformations were applied on  $\sigma_{sx}$  and CD?

Page 8, section 3.4: What is your justification of the variables used in the multiple linear regression model? And where is the result of your model selection using AIC?

Page 16, Could you also show the footprint of snow-off lidar data you used?

Page 17, Similar comment as before. You should not say the spatial distribution of the two dates' lidar flight are the same.

---

Interactive comment on The Cryosphere Discuss., doi:10.5194/tc-2016-188, 2016.