Interactive comment on “Lidar-based modelling approaches for estimating solar insolation in heavily forested streams” by Jeffrey J. Richardson et al.

Anonymous Referee #2

Received and published: 21 March 2019

The study presents an interesting approach to derive solar insolation estimates at and near forested riverbanks from Lidar data. Other than point-based measurements with pyranometers or hemispherical photographs, this allows for a spatially continuous mapping taking topography and vegetation geometry into account. The authors compare two Lidar-based estimates to measured references at 19 points in four transects along the Panther Creek NW Oregon, USA. They conclude both methods to be feasible for heavily forested stands, although areas with high insolation due to more open vegetation have not been covered by the study.

While the study is generally well-written and will likely become an interesting contribution to HESS, I find some of its structure and arguments to require and deserve revisiting. In the following I will outline my major concerns and suggestions. I hope the authors will find these as constructive as I intend them to be.

1 General comments:

I appreciate very much that the authors provide their data and analysis (as is HESS standard now). While I could easily follow the general setup of the study, I found it difficult to grasp the information residing in the Lidar data set and how it has been used. Since the latter is not included in the repository: Did I understand correctly that the Lidar data was commercially acquired and preprocessed to 1m pixels? So each pixel has values about all point returns, the number of highest hits (canopy) and the number of lowest hits (ground)?

Please be more specific about the calculation methods than naming the Software ArcGIS. I suppose this is an array operation which could be done in R (or any other math software) too. Which approaches did you employ? What can be understood about the "10m Buffer around the field points" (L187) and how does it differ to the "shifted square buffer" (L188f.)? Did you average within this area for comparison? What are the effects on the performance of the estimates. Especially with regards to the issue of "registration errors" L277ff. would this mean that a higher resolution could be more accurate or in other words that the hemispherical photographs suffer from minor shading effects to become representative at stand scale?

For a validation of the Lidar-derived solar insolation there is basically the correlation plot in Fig. 8 comparing it to pyranometer measurements. To me this does not appear very convincing to support the conclusion. By not allowing for an intercept
in your linear regression model, you define the bias-term to be zero. While this is an understandable desire in comparing two measurements which should give the same results, I do not understand your statement in L298f. The 16 points appear to overestimate the pyranometer references in most cases. High insolation references are underestimated. With an R² of 0.63, I find it rather problematic to speak of accurate: L329f. "a synthetic hemispherical photograph approach accurately predicts solar insolation and light transmittance".

In this respect, I moreover have difficulties to relate this back to the presented indices which leaves me with a couple of questions about the reason of their introduction in the first place. This confusion might partially stem from the manifold usage of the term "model" in the manuscript. I would suggest to allow for a more precise terminology to differentiate regression analyses from conversion models, from indices and from spatial map models. From the title I was expecting several modelling approaches using the Lidar data, which I did not find in the manuscript.

Coming back to the indices (Fig. 6, Tab. 3) I do not find the focus of the study specifically suitable to address these correlations. Contrastingly, the comparison of synthetic and actual hemispherical photograph (Fig. 7) is very compelling but falls in my view a little short in its analysis and evaluation (e.g. applying this for all 16 locations).

Since the validation of the "Lidar-based modelling" is rather difficult using the 16 measurements alone, maybe some further reference could be derived from remote sensing products? This could also provide the link to some of the addressed indices?

2 Minor comments:

L28f.: why only ecological applications?
L29: do trees really interact (so having feedbacks) with solar radiation?
C3

L36: can (solar) energy intercept with something? maybe irradiate a stream?
L37: how does solar irradiation limit options for forest management? I do not understand.
L48ff.: is it really necessary to describe the function of a pyranometer (at this broad level of detail)?
L53: I do not see the difference between the time references of a direct state measurement and the photograph
L56: Depending on the type of pyranometer, diffuse radiation is directly measured too.
L67: Start new paragraph with "Airborne lidar..."?
L113ff.: very confusing. please rephrase.
Fig 1: I would prefer all four Lidar models/maps instead of the grey box, which I assume to be the total Lidar dataset footprint. If you find my suggestion feasible, maybe a map of a satellite RS derived index could also be a reference here. A colourbar would be nice.
L200ff.: What happened to the longitudinal profiles? Were they processed?
L215: See general comment. Which exactly are THE models? do you refer to the different indices? the calculus to derive them? a model to generate the synthetic hemispherical what are the assumptions behind the comparison approach? What is the observation reference deemed as closest to the true value?
L257: model performance? in reference to what? Is a R² to each other really a good measure?
I do not understand why this should not be desirable... actually, I find the results in fig 7 quite convincing and the sensitivity might be quite an interesting feature. Please see my general comment on this, too.