Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2020-324-AC1, 2021


Geoscientific Model Development Discussions
Open Access
EGU



# *Interactive comment on* "Cutting out the middleman: Calibrating and validating a dynamic vegetation model (ED2-PROSPECT5) using remotely sensed surface reflectance" *by* Alexey N. Shiklomanov et al.

**Alexey N. Shiklomanov et al.**

alexey.shiklomanov@nasa.gov

Received and published: 5 February 2021

General Comments (overall quality): This manuscript uses remotely-sensed surface reflectance observations to calibrate/constrain a series of ecosystem parameters from the ED2-PROSPECT (EDR) model characterizing the land surface at 54 forested sites related to leaf biochemistry, canopy radiative transfer, and soil characteristics. An important innovation is the introduction of the radiative transfer model PROSPECT to the biosphere model ED2 to provide an improved spectrally resolved simulation



Printer-friendly version

Discussion paper

of surface reflectance. This is done, in part, to bring the observed surface reflectance closer to what the model actually predicts, helping to reduce the impact of observational uncertainty as well as more effectively constraining multiple components of the model. The authors find that through the assimilation of surface reflectance EDR can provide better simulations of the surface reflectance spectra and leaf area. The findings suggest that this approach could be used to better constrain surface energy balance as well as overall ecosystem functioning. Given many other ecosystem models include a two-stream radiative approach they contend their results should be widely applicable.

Scientific Questions/Issues: The authors bring up the issue of equifinality in the introduction, which presents a challenge for this application in that the surface reflectance can be a function of leaf biochemical properties, leaf structure, and canopy and soil radiative transfer characteristics. To some degree, equifinality was reduced in that the prior parameter distributions for biochemical leaf properties were tightly restricted, and limited to the extent the surface reflectance observations could influence them. In contrast the canopy radiative transfer parameters and LAI (through SLA) were simultaneously being adjusted by the optimization. Was hoping the authors could comment more about how equifinality of the surface reflectance influenced their results.

We agree that the topic of equifinality in the relationship between surface reflectance and canopy radiative transfer is worth additional discussion. We agree with the reviewer's assessment that equifinality was reduced by using informative priors for the leaf parameters. In addition, although not demonstrated here, another way to reduce equifinality is through the ecophysiological mechanisms embedded in the model itself — i.e., some combinations of model parameters and states that are consistent with a given surface reflectance may be excluded because they imply ecologically unrealistic

states in previous time steps or result in ecologically unrealistic outcomes in future time steps. Incorporating this dynamic aspect is an important future direction of this study.

In our revision, we added the following text to the discussion:

Nevertheless, our analysis echoed some known challenges in canopy radiative transfer modeling. One challenge is equifinality in the contributions of leaf biochemistry, leaf morphology, and different aspects of canopy structure to canopy albedo, which means that multiple variable and parameter combinations can produce very similar canopy albedo responses. We mitigated the equifinality between leaf traits and canopy structure by using informative priors on leaf traits from an independent data source (Shiklomanov 2018). However, there is additional equifinality in the effects of the EDR canopy structure parameters. For example, because the effective LAI used in EDR's actual radiative transfer calculations is defined as the product of "true" LAI and clumping factor, and because LAI is, in turn, derived from multiple parameters (leaf biomass allometry, specific leaf area), these parameters collectively cannot be independently determined from reflectance data alone. At the same time, increasing the leaf orientation factor (more horizontal, or "planophile", leaf orientation) has a similar (although not identical) effect to increasing LAI and clumping factor—namely, increasing canopy reflectance, especially in the near-infrared. Collectively, these issues may help explain some of the edge-hitting behavior (parameter distributions clustered at the ends of the distribution) observed in our posterior estimates, and some of the bias in our LAI estimates.

That being said, one major advantage of the Bayesian calibration approach is that its output is a joint posterior distribution that includes not only fully quantified uncertainties for each parameter but also the variance-covariance matrix of each parameter. Equifinality in parameters would manifest as strong pairwise correlation between parameters in the posterior distribution. Examining this correlation matrix (attached Figure 1) shows that there are some parameter pairs with strong correlations, such as the positive correlations between leaf and wood allometries for all PFTs except northern pine, and the hypothesized negative correlation between the leaf allometry (LAI) and

clumping factor, which was only observed for the early-successional hardwoods and northern pines. However, these correlations mostly do not occur in the parameters that exhibited strong edge-hitting behavior—namely, clumping and orientation factors for mid- and late-successional hardwood PFTs. Strong correlations also occurred among some of the PROSPECT parameters, and between PROSPECT and structural parameters, but contributed little to equifinality because the strong constraints on PROSPECT led to overall small covariance terms. Finally, because our calibration captured all of these covariances the presence of moderate equifinality did not preclude ecologically meaningful parameter constraints or accurate predictions because these covariances are being propagated into predictions. This is directly analogous to how a linear regression can have a tight confidence interval, despite high correlations between the slope and intercept, with that equifinality driving the characteristic hourglass shape of a regression confidence interval.

> As a follow up question to the equifinality question above – ED2 is a dynamic vegetation model with the ability to simulate competition amongst cohorts thus providing a simulation of co-existing dominant PFTs. It wasn't clear how well the simulation of cohort competition influenced the final distribution of PFTs and to what extent this matched the site level observed vegetation state. Given that the parameter optimization was PFT specific, the precise vegetation PFT distribution could have a large impact. Was the PFT distribution prescribed?

In this study, the vegetation composition at each site (including the PFT distribution and size-age structure) was prescribed in detail based on data from the NASA Forest Forest Functional Types (FFT) field campaign. In our revision, we have clarified this point in the first paragraph of "Site and data description", and briefly revisited it in the results in the context of future directions involving dynamic model simulations where the PFT distribution is predicted with some uncertainty.





Printer-friendly version

Discussion paper

> Was there any attempt to withhold some site reflectance data and apply the calibrated model parameters at those sites? It seems the optimized surface reflectance simulations were performed at sites that were already calibrated. The fact that the authors performed an across-site joint assimilation may in part account for this, but was interested how the calibrated parameters would perform at sites outside the calibration sites.

We acknowledge that cross-validation or out-of-sample validation are useful tests of model performance, and in our revision, we recommend these activities as future directions for this work. However as the reviewer points out, because our calibration was joint across all sites, we did not feel that a separate validation at other sites not used in the calibration was necessary. With 54 sites in our calibration, any single site represents <2% of the data, and for a joint calibration without site random effects, we have every reason to believe that the calibration is not overfitting to any individual site; trying to fit any one site well would cause others to do worse (especially given the large observed variability in forest structure) unless the EDR model structure was reasonable and the parameters chosen were genuinely good choices. We have added this information to our revised discussion.

> It is known that radiative transfer models are challenged in simulating evergreen species in part because of the irregular and open space canopy structure. Many of the figures in the supplement demonstrate stronger biases in simulated surface reflectance exist for evergreen sites as compared to deciduous. Was hoping the authors could comment on this, and recommendations for getting around this.

The reviewer brings up a great point about conifer canopies historically being harder to capture. Although some conifer-dominated sites did demonstrate significant biases in reflectance predictions, our analysis of reflectance bias by composition, structure,

and spectral region (Figures A3 and A4) shows that these biases are not systematic (though they may drive greater predictive variance). In our revision, we describe this analysis in more detail in the Methods and Results.

> Whereas leaf level biochemistry related parameters were well constrained by the priors, many of the radiative transfer posterior parameters seemed to be edge hitting parameters. To what extent do the authors believe this behavior was caused from structural error in canopy radiation transfer and/or mismatch caused in part from radiation directionality differences between the simulated and observed canopy reflectance? The authors devote a considerable amount of general discussion regarding this topic, but don't directly address how this might have effected their own results.

In our revision, we added as appendices a parameter sensitivity analysis of EDR (attached Figure 2) and an analysis comparing EDR to PRO4SAIL, a similar 1D two-stream canopy radiative transfer model popular in the remote sensing community. These analyses point to two likely explanations for the edge-hitting behavior of some parameters. The first—described above—is equifinality in the effects of several structural parameters on LAI and canopy albedo.

The second explanation is a structural issue in the EDR model that leads it to systematically underestimate albedo, particularly for low solar zenith angles (sun directly overhead). A detailed description of the issue is provided in Yuan et al. (2017). Briefly, EDR (and the Sellers, 1985, model from which EDR is derived) defines direct radiation backscatter as a function of the single-scattering albedo, which in turn is an integral involving the leaf scattering phase function and leaf projected area function. The Sellers (1985) analytical solution to this integral assumes a uniform scattering phase function, which is appropriate for point scatterers but less so for horizontal surfaces like leaves. The practical consequence of this assumption is a lower value of the direct radiation backscatter and therefore a tendency to underestimate albedo, which is consistent with

the results of our comparison of EDR and SAIL. Our EDR calibration is likely to be compensating for this behavior via a preference for higher effective LAI values (i.e., higher values of leaf biomass allometry and clumping factor) and more horizontal leaf distributions (i.e., higher leaf orientation factor), both of which increase albedo. We have added this information to our discussion.

> The manuscript begins by justifying the inclusion of the PROSPECT model in to ED2 to bring the model closer to the observations, to, in part help reduce the uncertainty of the observation that are assimilated into EDR. More explanation on how the surface reflectance observational uncertainty was quantified here, and what it represents, and to what extent overconfidence in uncertainty may have led to the posterior edge hitting parameters.

As shown in equation 26, observation error in the reflectance data was not estimated a priori based on the instrument itself, but was modeled as the residual error between the model and the data, analogous to what is done for any linear or nonlinear regression model. A key difference, however, is that the error model accounts for the known heteroskedasticity in spectral data (i.e., the size of the variance increases with the magnitude of the reflectance). In terms of random spectral errors, there is no reason to expect this variance to be overconfident for inferences made on these landscapes, especially as the study sites were not all imaged on the same day under the same atmospheric conditions, though we'd agree that one might not want to apply this variance to entirely different ecoregions. Furthermore, because the variance slope and intercept are fit parameters, whose parametric uncertainty is being quantified and propagated, this makes it even less likely that our uncertainty estimate is overconfident. That said, the current approach does not formally account for any possible systematic errors in the observations, which could have a more serious impact on inferences. However, we would note that we are unaware of any derived data products that account for these systematic errors either. Furthermore, in addition to the same uncertainties about

reflectance that we face, those products additionally contain numerous uncertainties about model structure, parameters, and covariate data whose uncertainties are rarely fully propagated, meaning that the alternative approach (calibration to derived data products) is more likely to result in overconfidence in uncertainties than the approach taken here. Along these lines we added a line to the Discussion about how accounting for systematic data errors would be a useful future direction.

That said, we feel it is highly unlikely that overconfidence in surface reflectance estimates contributed to the edge-hitting behavior of some of our posteriors. Random errors would not result in a systematic parameter bias and, given the long history of the instrument and maturity of atmospheric correction approaches, any systematic errors in AVIRIS are likely to be quite small relative to the structural uncertainty in ED2's RTM (i.e., biologically implausible parameters should not be necessary to capture observational data biases of the magnitude likely to be present).

> Detailed Comments: Abstract and manuscript in general: Need more discussion on what we hope to gain by this. We don't really care about surface reflectance (although energy balance is important), but we do care about how LAI, chlorophyll, pigments and water status influence ecosystem functioning through carbon and water exchange. I think this needs to be emphasized more, and provide evidence that this sort of setup can accomplish this.

We agree that additional emphasis on the implications of this work is warranted. In our revisions, we emphasize several important implications in the abstract, introduction, and discussion.

First, as you say, energy balance is important, and the contribution of vegetation to land surface albedo is a critical mechanism by which vegetation influences regional and global climate (Bonan 2008). Therefore, ensuring the accuracy of model simulations of

albedo, including its sensitivity to vegetation structure and composition, is essential to accurately projecting the effects of climate- and land use-driven changes to terrestrial ecosystems on future climate.

Second, canopy radiative transfer directly affects many physiological, ecological, and physical processes included in complex demographically-enabled vegetation models like ED2 (Viskari et al. 2019). Light availability and absorption is a first-order control on photosynthesis, and ability to survive under different light levels is an essential component of a tree species' position in forest succession. Meanwhile, temperature—which is strongly influenced by albedo—directly affects the rates of both enzyme-kinetic physiological processes and evaporation.

Finally, ED2 and similar models are highly sensitive to many of the leaf traits constrained by this analysis (e.g., Dietze et al. 2014; Raczka et al. 2018; Shiklomanov et al. 2020). Given the large variability of these traits through space and time, remote sensing is an essential data source for model parameterization, and our work provides a useful approach for doing so.

> Line 1: Remove 'derived'. The fact that they are 'data products' and not 'observations' gets across the point.

While we agree that "data products" should imply a difference from observations, treatment of remote sensing data products as true observations without accounting for uncertainties or biases is widespread in the Earth science community. Therefore, we think it is important to emphasize that these products are derived.

> Line 5: 'compared against airborne and satellite data' Technically, this is still data and not observations in that even reflectance data requires RTM models, I believe. But it is more direct relationship

We agree that additional nuance is required here (especially when also considering reviewer T. Quaife's comments). Therefore, we have revised this section to capture the fact that although these data are still derived (processing steps include atmospheric correction and orthorectification), they rely on fewer assumptions (especially about the land surface), have fewer processing steps, and therefore are closer to observations.

> Line 23: add 'to' calibrate or constrain

We have revised this accordingly.

> Line 25: I know exactly what you mean by 'constrain', but could you use 'calibrate' or 'inform' in this context?

We have replaced "constrain" with "inform".

> Line 32: "More sophisticated approaches for estimating vegetation properties based on physically-based radiative transfer models face issues of equifinality, whereby many different combinations of vegetation and soil properties can ultimately produce the same modeled surface reflectance (Combal et al., 2003; Lewis and 35 Disney, 2007)." This is important I think – and raises a key point for this analysis – is there not equifinality when trying to constrain leaf structure vs. leaf status? I would think equifinality could be a problem here, and I think you need to acknowledge this and how you might address this – Strong priors? Demonstration that surface reflectance can tease apart these two things. . .

Please see our above response about equifinality.

Line 52: Awkward topic sentence. Simplify “Some land surface models already include there own . . . that allowd for a more direct comparison to remotely sensed surface reflectance.”

We have revised this sentence according to the reviewer’s suggestion.

Line 57:”Canopy radiative transfer plays a particularly important role in the current generation of demographically-enabled dynamic vegetation models, where differences in canopy radiative transfer representations and parametrizations have major impacts on predicted community composition and biogeochemistry (Loew et al., 2014; Fisher et al., 2018; Viskari et al., 2019).” Seems weird to word it this way. Isn’t it the other way around, community composition and biogeochemistry impact the RTM?

It is true that composition and structure impacts the RTM, but all of these studies show that the opposite is true as well! These studies illustrate that because the RTM determines the overall energy balance of the ecosystem and the distribution of light within the canopy, RTM formulation and parameters profoundly impact the predicted biogeochemical fluxes and vegetation dynamics. For example, Viskari et al. (2019) show that uncertainties in canopy RT can cascade and impact a number of associated processes, including photosynthesis, energy balance, internal competition, and demography. We have revised the sentences here to more clearly and explicitly convey this idea.

Line 72: “. . . will significantly constrain model parameters related to canopy structure.” So the goal all along was to constrain canopy structure with surface reflectance, not necessarily foliar biochemistry? Maybe talk a bit more about the differences in sensitivity of surface reflectance to canopy structure vs foliar biochemistry.

Our objective was to evaluate which parameters could be constrained. The list of candidate parameters included parameters related to both structure and biochemistry. However, we hypothesized that, because of the informative priors on foliar biochemistry, the constraint would be relatively greater for canopy structural parameters (as stated here).

Per our earlier responses, we have included an additional parameter sensitivity analysis of EDR and discussion thereof. We have also added more text on the issue of equifinality to the introduction and discussion.

> Section 2.3: Can you provide a sense of scale? For example for the 54 sites, what spatial range was the inventory data taken, and what spatial resolution did AVIRIS cover? Trying to get a feel for spatial mismatch, etc. Were sites chosen because they were rather homogeneous for certain PFTs?

In response to this and Reviewer T. Quaife's comments, we have elaborated on the methods behind the data used in this analysis. Specifically, each of the 54 sites here consisted of a 60 x 60 m transect within which forest inventory data were collected. AVIRIS-Classic data were extracted as the average of a 3x3 pixel array (each pixel is 15-20m, depending on aircraft altitude) centered on the site transect center, resulting in a single composite spectrum for the 60 x 60 m area. Additional details on the sampling methodology are described in Singh et al. (2015) and references therein.

> Figure 1: Was a little surprised to see many sites so close to Lake Superior. No issues with interference from nearby water reflectance?

Although sites do appear very close to Lake Superior in the map we provide, all sites are sufficiently inland (several kilometers) that contributions from water reflectance of

nearby pixels can safely be assumed to be negligible. In the revision, this figure has been broken up into two figures

> Line 200: Could you provide a bit more explanation of what including the EDR predicted LAI term within your probability function accomplishes? EDR response becomes saturated to LAI so is this an artificial way to account for increased reflectance?

Yes, the LAI penalty in the likelihood function accounts for the saturating effect of increasing LAI on reflectance. This effect is not unique to EDR—rather, it is a well-known consequence of the exponential extinction of light through a medium, following Beer's Law. Therefore, a canopy with an unrealistically high LAI like 15 has virtually the same reflectance as a canopy with a high but more feasible LAI like 6 (all else being equal), and therefore a likelihood calculation that does not penalize excessively high LAI values would consider both outcomes equally likely.

We have added additional text to this effect to the methods, and have added the attached figure to the supplement demonstrating the saturating effect of LAI on reflectance.

> Line 227: So, to evaluate the model you compared the EDR-spectra against the AVIRIS observations at sites that were used to calibrate the model? Was there any attempt to withhold some site data and apply the calibrated model at those sites?

As we stated in our main response, because our calibration was joint across all sites, we did not feel that a separate validation at other sites not used in the calibration was necessary.

> Line 232: Can you quantify what 'informative' means

We have clarified this sentence to say these are “leaf parameters whose prior distributions were already independently constrained by an earlier analysis”. These prior distributions are shown in Figure 2.

> Line 233: I cannot see in figure where PROSPECT N parameter is?

In our experience, labelling PROSPECT’s “N” parameter as such is confusing to readers unfamiliar with PROSPECT because it suggests leaf nitrogen content. Therefore, we prefer the more explicit name “# mesophyll layers”. We have revised the methods and results text in a few places to clarify this.

> Section 3 Results: Although Figure 2 was very informative, I found the Results section in general, relatively vague, perhaps some sense of % reductions in 95% credible interval.

We have revised the results to include more precise statistics, including the suggested relative reductions in the width of the 95% credible interval. For example, relative to the prior, posterior credible intervals were 19% to 58% as wide for clumping factor, 14% to 78% as wide for leaf orientation factor, and 1% to 10% as wide for leaf biomass allometry.

> Figure 2: Hopefully, the authors comment on some of the apparent edge-hitting parameters specifically related to canopy RTM parameters such as leaf orientation, canopy clumping and water. I worry that the information from the relatively strong and defensible leaf biochemistry prior parameters leading to relatively self contained and PFT differentiated posteriors for the leaf biochemistry parameters is lost or made irrelevant due to biases between model simulated and observed surface reflectance that are corrected

by ‘fitting’ the RTM parameters. The context of this manuscript doesn’t indicate how sensitive the surface reflectance is for the suite of parameters calibrated here. . . perhaps included in one of cited manuscripts.

We added text to the results section highlighting the edge-hitting behavior for wood biomass allometry, canopy clumping, and leaf orientation. We also added figures showing the sensitivity of EDR to the relevant parameters in this figure to the supplementary information.

For wood biomass allometry, the edge-hitting behavior approaching zero is consistent with the typically small-to-negligible contribution of aboveground woody elements (stems and branches) to reflectance of dense canopies during the growing season (Banskota et al 2015). However, these woody elements are a large biomass sink, so constraining wood allometry parameters is important. We have added text to the discussion about the importance of constraining wood allometry, the limitations of doing so using canopy reflectance alone, and additional analytical (e.g., via known covariance with other parameters) or observational (e.g., leaf-off optical, LiDAR, radar, in situ) constraints that should be explored.

For canopy clumping, we observed the edge-hitting behavior primarily for mid- and late-hardwood PFTs in the direction of no clumping. In EDR, clumping factor appears only as a scaling factor on LAI—namely, EDR defines the effective LAI, eLAI, as the product of LAI and clumping factor, and the main radiative transfer calculations use eLAI to quantify the depth of vegetation. This makes the clumping factor parameter highly confounded with LAI (i.e., a two-fold increase in LAI is perfectly compensated by a two-fold decrease in clumping factor). The edge hitting behavior of clumping factor at its maximum—1.0—suggests the tendency of the calibration for these PFTs to prefer larger leaf area indices, which we also see in the leaf biomass allometry. As mentioned in our earlier response, this may be compensating for a tendency of EDR to underestimate albedo (increasing LAI tends to increase albedo). Moreover,

it is expected that dense, mature hardwood PFTs would trend toward a less-clumped canopy compared with a more open, clumped evergreen needle-leaf canopy.

For leaf orientation, we primarily observe edge-hitting behavior approaching the most horizontal leaf orientation for the late conifers and mid- and late-hardwoods. As mentioned in the earlier response, and as with LAI above, this may be compensation for EDR's tendency to under-predict albedo.

Finally, the leaf water parameters are not truly edge-hitting but do show a consistent directional shift with the AVIRIS data suggesting a higher leaf water content than the prior leaf-level data. This is perhaps not surprising as retrieval of canopy water content with hyperspectral data, and specifically AVIRIS, has been one of the most widely-utilized methods in the literature (e.g. Gao and Goetz, 1995; Clevers et al., 2010). In this case, our parameter distributions for EWT suggest that given the time of year when the imagery was collected the vegetation tended toward higher canopy moisture conditions, likely given this was during the peak of the growing season, and peak greenness and neither year had any indication of lower than normal precipitation or drought.

Banskota, A., S. P. Serbin, R. H. Wynne, V. A. Thomas, M. J. Falkowski, N. Kayastha, J. P. Gastellu-Etchegorry, and P. A. Townsend. 2015. An LUT-Based Inversion of DART Model to Estimate Forest LAI from Hyperspectral Data. Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of 8:3147-3160.

Gao, B.-C., & Goetz, A. F. H. (1995). Retrieval of equivalent water thickness and information related to biochemical components of vegetation canopies from AVIRIS data. Remote Sensing of Environment, 52(3), 155–162. https://doi.org/10.1016/0034-4257(95)00039-4

Clevers, J. G. P. W., Kooistra, L., & Schaepman, M. E. (2010). Estimating canopy water content using hyperspectral remote sensing data. International Journal of Applied Earth Observation and Geoinformation, 12(2), 119–125.

https://doi.org/10.1016/j.jag.2010.01.007

> Figure 3: A map would be helpful with site codes provided. Perhaps a zoomed in map that demonstrates where these sites are spatially? Also is the stem diameter plot on the right simulated or observed? In fact, doesn't that have a large impact on the assimilation, the PFT distribution and stem diameter distribution?âĂŤ has it been demonstrated that ED2 can properly simulate the competition of PFTs at the site providing the correct vegetation state, such that the parameter optimization reflects the observed vegetation state ?? Or has the vegetation state been prescribed in this case?

We have split up the original Figure 1 into two separate figures: one showing a larger map with sites labelled (attached Figure 3) and one showing the density vs. diameter plot (updated to include a self-thinning curve based on T. Andrews, unpublished; attached Figure 4).

As stated in the main response, the vegetation composition at each site (including the PFT distribution and size-age structure) was prescribed in detail based on NASA Forest Functional Types (FFT) campaign field data. In this figure's caption, we have clarified that these are the observed stand structures.

Zeide, B. (2010). Comparison of self-thinning models: an exercise in reasoning. Trees, 24, 1117–1126. https://doi.org/10.1007/s00468-010-0484-z

> Figure 5: "The observed vs. predicted line had a slope of 0.37 and an intercept of 2.80, indicating that EDR calibration underpredicted LAI on average but overexagerrated across-site LAI variability." What do you attribute this clear structure in residuals between observed-simulated LAI?

We agree the clear mismatch between predicted and observed LAI was not given sufficient attention in our original draft. We have incorporated the following analysis into our revision:

EDR tended to underpredict LAI at sites with lower mean DBH and overpredict LAI at sites with higher mean DBH (attached figure). However, there was no relationship between LAI bias (predicted - observed) and stand density, proportion of conifer PFTs, or bias in predicted reflectance (figures attached). Breaking the LAI bias down based on PFT site dominance shows that the bias was most pronounced in sites dominated by Early Hardwood and Northern Pine trees (attached figure). Based on these results, we conclude that the LAI bias can be attributed to an underestimation of the leaf biomass allometry parameter, as both Early Hardwood and Northern Pine PFTs had relatively low estimates of the leaf biomass allometry parameter. In part, this may be a consequence of site sampling: both Early Hardwood- and Pine-dominated sites skewed towards stands with low DBH, where sensitivity to changes in the allometry parameter are relatively low. For pine sites, the challenges with modeling conifer radiative transfer (see earlier responses) likely also played a part. Overall, this analysis reiterates the importance of evaluating models against multiple distinct variables—after all, none of these biases would have been apparent from looking at the reflectance simulations alone.

> Line 258-270: I like the overview explanation of bringing observations closer to models or alternatively bring models closers to the observations. In the end, it's a bit of semantics, especially for using information from satellites we will always need some sort of transfer function or forward operator to convert from what a satellite observes and what a model predicts. I don't think one way or the other should take precident. The advantage in your approach, however, is the potential for the observed surface reflectance to constrain multiple model components, whether that be leaf structure or biochemistry or water status. I think that is potentially extremely valuable

although I am not sure it has been demonstrated, yet, that this is the case. I feel more could be 'learned' about what information surface reflectance could provide if you could prescribe the LAI and PFT-distribution in ED2, then you could really get a grasp on what it can inform, leaf biochemistry? Within-canopy RTM parameters? Etc.

We disagree that the difference between bringing models closer to observations and vice-versa is a purely semantic one – the fact that all observations need some transformation does not mean that all transformations are the same. Instead, we argue that comparing model output to a highly derived data product (e.g., MODIS GPP) is closer to a model intercomparison than a model validation, since generating those data products invariably requires their own models (whether statistical or process-based) that make specific assumptions about different processes, often very different assumptions than the model the products are compared with. Thus it's often really an "apples to oranges" comparison. In other words, in addition to the "shared" observational uncertainties (e.g. atmospheric correction, instrument calibration) that are present in both approaches, derived data products include numerous additional assumptions and uncertainties that are avoided in a spectra-to-spectra comparison. Meanwhile, as you point out, the approach of bringing models closer to observations is advantageous precisely because of the model's emergent covariance across many disparate variables rooted in specific assumptions about biophysical and ecological processes. In practice, for an observation operator that is almost entirely disconnected from the model (for example, if we just took the total LAI from ED2 and used a standalone RTM like SAIL to simulate the reflectance), we agree that there's relatively little advantage relative to a derived product, but the more tightly an observation operator is integrated into a model (as in our analysis, where canopy RTM parameters and outputs are used elsewhere in ED2), the greater its value.

To address this, we have significantly revised the discussion to (1) more explicitly highlight the kinds of assumptions typically made in remote sensing data product genera-

tion; and (2) to discuss the relative advantage of building models with remote sensing-friendly observation operators, especially when such operators are tightly integrated with other processes in the model. For the latter, we now discuss some future directions for ED2 and optical remote sensing specifically, such as using the PROSPECT leaf model to couple the model's representations of plant hydraulics (for leaf water content) and shade-driven trait plasticity (for leaf structure and pigmentation) to canopy radiative transfer. Similar approaches that leverage the model's internal RTM can also be applied to other remotely-sensed data, such as lidar (canopy structure), radar/microwave (structure and water content), thermal (canopy traits and water use), and fluorescence (photosynthetic traits). Indeed, this paper is an important first step towards the ultimate goal of leveraging internally-consistent process-based models to make joint inferences across multiple remotely-sensed data constraints simultaneously.

> But, as you brought up in the introduction, this brings up equifinality issues. Not so much in this case for your leaf biochemistry parameters because of strong priors, but it does seem to be the case for canopy RTM parameters and predicted LAI (SLA). I think you need to caveat or address this concern.

We agree that equifinality between canopy structure and leaf biochemistry in general, and the specific ways that canopy structure is represented in EDR, are concerns that warrant further discussion. As we stated in the introduction, and as you point out in your review, an effective way to address equifinality is by incorporating prior information that can constrain the parameters. In this study, we showed that external informative priors on leaf biochemistry parameters are effective. Other data—for instance, observations of canopy structure from active remote sensing (LiDAR, radar) or in situ measurements—could help alleviate some of the other issues with our results. Moreover, if our approach was applied in dynamic model simulations, the internal logic of the model's dynamics of leaf biochemistry and canopy structure would provide additional constraint on the possible parameter space, which is an explicit future goal.

We have elaborated on all of these points in the revised discussion.

> Lines 286-310: It seems like you are pointing out sources of structural error within the radiative transfer of EDR or, to the extent, mismatch between what EDR simulates vs. what AVIRIS-Classic observes. Could this help explain why the calibration caused some posterior RTM parameters to be edge-hitting against the bounds of the priors? Also, this work was in part motivated by being better able to quantify the uncertainty in the surface reflection observations, but I am not sure I saw a clear explanation of the uncertainty that was used or provided for the AVIRIS and how was this quantified. It seems parameters have the potential to be overfit, if the observation uncertainty is not realistically quantified. May have missed this.

See earlier responses.

> Line 320: Really it's the power to upscale that remote sensing products provide. However, this comes at a cost, they 'observe' reflected radiation from the land surface which indirectly characterizes things that we care about like, like leaf biochemistry, albedo etc.

We agree the reflectance data collected by passive optical remote sensing are affected by many different surface features and processes, often in confounding ways, which makes them challenging to use in isolation. Fully leveraging the power of remote sensing requires additional sources of information from both in situ data and the understanding of biophysical and ecological processes embedded in our vegetation models. We have revised this text and expanded this text accordingly.

> Line 326: 'accurately reproduce surface reflectance and leaf area index' I think this is a bit of an overstatement, especially because the fig-

ures demonstrate systematic mismatch between the optimized surface reflectance and observed reflectance (figure 3), and strong residual error structure in LAI (Figure 5). Perhaps this approach provided 'improved' reflectance and LAI is better wording.

We agree that we overstated our LAI prediction accuracy here, and have revised this text to temper our conclusions accordingly. However, as we pointed out in earlier comments, although our reflectance predictions were not perfect everywhere, the quantitative analysis of surface reflectance revealed no systematic biases by PFT or stem density.

Line 330: I think you also need to say where this work can lead âĂŤ this is of interest for those that are concerned with ecosystem functioning and that this could provide improved estimates of both biomass and land-atmosphere carbon and water exchange. Also some discussion of the differences between evergreen and deciduous forests would be helpful. Generally RTM's have more difficulty with more open canopy evergreen species and not really discussed in this manuscript.

Our significantly revised discussion (see earlier comments) includes much more text on ways that ED2 (and other vegetation models) can be further enhanced to better take advantage of passive optical and other remote sensing techniques. In addition, to address this comment specifically, we have added several sentences to the conclusions about the role of better-constrained vegetation models in improved estimates and understanding of biomass and vegetation-atmosphere interactions. Finally, earlier in the discussion, we have added a few sentences about known the challenges with canopy radiative transfer modeling in conifer species, though again, our results did not show any systematic biases in reflectance predictions by plant functional type.





Printer-friendly version

Discussion paper

Appendix:

Figure A2: What would be really helpful is to use the same color-coding in Figure A1 to show deciduous vs evergreen sites. Also this brings about the question – why did you choose the sites that you did to put in the manuscript itself?

Figures A5-A13 provide the requested breakdown of plots by PFT composition. For the main figure, we selected sites that, collectively, span the geographic, stand-structure, and PFT composition space of our study. We have added this information to the Figure 3 caption, highlighted these plots in our site map, and mention this in the revised Results section.

Figure A3: This sort of gets at the hardwood vs conifer performance as well. I think it would be helpful to comment on this distinction in performance within the results/discussion.

As mentioned above, we have added additional text about our analysis of performance by PFT to the methods, results, and discussion. As this figure clearly shows, there is no systematic bias in reflectance simulations for conifer species.

Figure A13: This is also a very compelling figure that gets at the increased bias in spectra for conifer. Worth discussing in main manuscript.

See earlier comments about this. We note that even this figure shows no systematic bias—three sites have underpredicted reflectance (AK06, AK60, OFO1), two sites have overpredicted reflectance (MN02, MN04), and the remaining 3 sites have relatively accurate reflectance predictions (BH02, SF01, SF04).





Printer-friendly version

Discussion paper

Please also note the supplement to this comment:
https://gmd.copernicus.org/preprints/gmd-2020-324/gmd-2020-324-AC1-supplement.pdf

---

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2020-324, 2020.



Printer-friendly version

Discussion paper