Biogeosciences Discuss.,
https://doi.org/10.5194/bg-2019-194-AC1, 2019

Biogeosciences Discussions

Interactive comment on "Comparing Stability in Random Forest Models to Map Northern Great Plains Plant Communities Using 2015 and 2016 Pleiades Imagery" *by* Jameson Brennan et al.

Jameson Brennan et al.

jameson.brennan@sdstate.edu

Received and published: 15 October 2019

Overview This manuscript describes how vegetation groups switch between years based on random forest classification models utilizing remote sensing imagery. The authors demonstrate how to produce highly accurate images with purely spectrally based predictors, and also quantify the variability in their vegetation groups between years. Understanding these shifts is an important undertaking for ecology and remote sensing, however there are several factors that lead to confusion in the interpretation of the results. In addition, the application of Random Forests is limited compared to what is set out in the intro, and when coupled with what seem like arbitrary (or unex-

plained) decisions, I feel their objectives have not been fully met. Specifically the confusion between community types and functional groups needs to be addressed in the manuscript through defining terms, and clarifying the difference between community changes and production differences of functional groups between years. For the assessment of random forests, there is opportunity to develop the analysis much deeper. There are unanswered questions that could be explored with RF models. For example, what number of trees are needed for the model to stabilize, how does the months of imagery (what if I have two instead of five each year) change the classification, why separate models for on and off prairie dog towns (when transitions between these and the three off town types may be important), etc. Specific Comments

-There are several comments by this reviewer regarding definition of plant community as used in this paper. Plant communities can be described at many scales; at a very large scale, we could simply distinguish between grassland and shrubland communities, knowing full well that there is tremendous potential for variability within those two categories. We could also describe plant communities at a very fine scale, combining land units only when the percent compositions (either by cover or biomass) of the communities are nearly identical. For the purpose of the overall study within which this study was a part, we determined that an intermediate scale provided reasonable distinctions between areas on the landscape without being too broad or too detailed to be useful. Thus we combined the plant communities into categories defined by several functional groups. As indicated, above, we are very cognizant of the fact that plant communities defined by a dominant functional group will not only have other functional groups within those communities, but the composition of those other functional groups may vary widely. We will add an explanation early in the manuscript that details the scale at which we are dealing with plant communities to clarify.

-Regarding the issue of number of trees to attain stability, prior remote sensing research has focused on differing number of trees on model predictions, as well as seasonality of imagery collection on classification error within the same year. While our

data could be used to explore that question, that is not the focus of this paper.

Line 44 – Awkward sentence, and if I understand this correctly, then I disagree. I actually am not surprised by changes in species dominance between years. Composition may stay the same, but representation can change depending on growing season conditions.

-Sentence can be re-written for clarification. Grasslands within the northern great plains are primarily composed of perennial species, thus it is unlikely that shifts in dominance will occur from one year to the next. Substantive changes to the relative dominance of species in plant communities tend to occur over longer time spans. It is more likely that, when plant communities dominated by one functional group (e.g. C3 grasses) also have a fairly large percentage composition of plants from a different functional group (e.g. C4 grasses), specific growing conditions may result in the appearance of a dominance shift.

Line 48 – Vegetation classification can be done at many scales in multiple vegetation hierarchies, you need to be much more specific here (and throughout) about what you are looking at and where in a vegetation hierarchy your results are relevant.

-Sentence can be re-written to specify high resolution satellite imagery for northern great plains plant communities.

Line 55 – Very broad and general and probably needs a citation. Take a look at (Browning, D. M., A. Rango, J. W. Karl, C. M. Laney, E. R. Vivoni, and C. E. Tweedie. 2015. Emerging technological and cultural shifts advancing drylands research and management. Frontiers in Ecology and the Environment 13:52-60) to think about how remote sensing fits into monitoring and assessment for rangelands.

-Citation can be added.

Line 90 – You are not exploring “plant and animal interactions” in this paper, although this is the aim of your larger project. This is out of place and confusing. In addition,

“plant and animal interactions” is vague, I started thinking about a wealth of LIDAR and similar studies used to create thematic vegetation maps for animal habitat studies. Do you mean there is limited studies on animal space use across vegetation communities?

-The aim of the larger project is to explore plant-animal interactions, specifically livestock use of plant communities on the landscape. However, prairie dogs have a large impact on plant communities and can drive differences in species composition, production, etc. There are certainly studies on animal selection of vegetation communities; the use of thematic maps developed from high resolution satellite imagery is, however, much less studied.

Line 101 –Three examples don’t prove that RF is better in all situations. And as written, it seems like you cite one study that actually compares RF with other techniques, and this used Landsat, very different than your study. Nothing majorly wrong here, you just need to introduce applications of RF to vegetation classification problems as one useful technique.

-RF may not be the best algorithm for all situations. The studies in this paragraph highlight some of the uses of RF to classify plant communities. Additional studies can be added to demonstrate RF outperforming other methods to demonstrate the high degree of accuracy for these models. We can make a change from “proven” to “demonstrated” in the sentence to alleviate the concern.

Line 114 – Probably this illustrates a limitation in predictor variables, rather than Random Forests. The tool can work at broad scales if the data and processing power is available. See Jones, M. O., B. W. Allred, D. E. Naugle, J. D. Maestas, P. Donnelly, L. J. Metz, J. Karl, R. Smith, B. Bestelmeyer, C. Boyd, J. D. Kerby, and J. D. McIver. 2018. Innovation in rangeland monitoring: annual, 30 m, plant functional type percent cover maps for U.S. rangelands, 1984-2017. Ecosphere 9.

-The Juel et al study mentioned deals with high resolution imagery, compared to 30m Landsat data as mentioned in the suggested study. Perhaps the issue of spatial trans-

ferability of models becomes a greater issue at fine scale mapping.

Line 120 – Plant community classification (used here and throughout) can be conducted on many different levels of vegetation hierarchies (like the USNVC). Or you can ask other questions, like changes in productivity between years. Community type generally shifts when a system crosses a threshold from disturbance/stressors, through succession, etc. You need to define much better in this manuscript what you mean, and what you are looking at in regards to, for example, plant community classification vs. plant community species or productivity. A single Landsat image may work very well for some purposes, but for the more detailed questions like yours, multiple images may be required (although you did not actually test accuracy differences between the number of images). You seem to be focused more on classification of functional groups rather than a community.

-See initial response, above, to the overview comments. Though several of our plant communities are dominated by specific functional groups (shrubs vs grass or grass vs forb), this does not mean, nor did we intend it to mean, these are the only functional groups within those communities. Communities are composed of a variety of species representing multiple functional groups. For example, forb dominated sites will also contain grass species and vice versa. We are classifying plant communities in this study based on the dominant functional groups within a plant community.

Line 121 – The three references here are specific studies (two over 10 years old), not reviews of plant classification studies. Maybe in the past it was more common to use a single time period, but seems now with increased computing power and the availability of the entire Landsat achieve, etc., it is very common for much more robust and multiple acquisition studies. Maybe phrase this more to acknowledge this evolution.

-Sentenced can be changed to acknowledge this.

Line 135 – Again, what is a plant community. Have you really found or are looking that the community changes? Or are you looking at species representation within a

community, i.e. functional group dominance and shifts in this between years.

-See initial response, above, to the overview comments. A plant community is a collection of species within an area of a relatively uniform composition different from neighboring patches. In this study, differences in neighboring patches are sometime evident in differences in dominant functional group (forb vs grass) or differences in photosynthetic pathways (C3 vs C4 grasses). The aim of this study is not to measure community changes, as shifts in plant communities occur over longer time scales within perennial northern great plains plant communities.

Line 143 – NGP probably too broad for the implications of your study. Be more specific with the MLRA, or mixed grass prairie systems, etc. that you are testing. Intro – Lots of general and vague statements in here and limited citations supporting broad brush statements. I suggest going through the intro to make it more specific. For example, landscape, local, various etc. scales will mean different things to every reader. Define these or what they are for the studies you cite. Also make sure your statements are supported by citations or explained. For example, line 61 – 65 is not a summation or conclusion of the paragraph, so these new statements should be supported. Finally, you should mention this is part of a larger study looking at cattle use compared to prairie dog prevalence and impact to pastures, but the paragraph starting on line 130 had me confused between what this study was going to do, and what the larger study did.

-Intro can be changed to include mixed grass prairie specifically, though much of the Northern Great Plains is comprised of mixed grass prairie. Intro can specify the differences in scales between citations such as Landsat imagery versus high resolution imagery. Line 130 states that this is part of a larger study linking livestock use of plant communities within pastures occupied by prairie dogs. Line 130 highlights the larger study, then shifts to the specific goals of this paper.

Line 174 – These sound more like plant functional groups than communities. Noth-

ing wrong with mapping those, but the terminology issues are prevalent and I believe confuse your conclusions. Changes in representation are common between years, changes in community are a different boat.

-See initial response, above, to the overview comments. In this study, plant communities are organized by dominant functional groups, with the knowledge that other functional groups are also present within these communities.

Line 199 – Why not compare them all together? You need to add rationale for why separating these out beforehand is appropriate. If you want to scale up your study, how will you separate out prairie dog towns at the "landscape" scale (watershed, county, etc.). As another option, I would find this much more compelling if the comparisons and RF models were tested to separate out all five groups. This would be a much more thorough test of RF.

-Separate models can be created to test the ability of RF to separate all five groups. The study was setup to map plant communities of interest on- and off-town for the purpose of relating this data to livestock use. Though plant communities are defined by dominant functional group, functional groups are not mutually exclusive within those communities, however, a site can only occur on- or off-town and not both thus is mutually exclusive.

Line 217 – Do you think the wider spectral bands (compared to Landsat or UAV options) played into your results at all?

-No

Line 227 – Again, this needs more justification then saying they are mutually exclusive. You either decided to map prairie dog towns separately from the rest of the study area (which you need to justify why) or you could test what the implications are of not having mapped towns in the first place (which also can vary between years).

-Prairie dog towns were mapped separately from the rest of the study area as part

of the larger project objectives. An additional test can be whether plant communities on prairie dog towns are different enough to be remotely sensed from off-town areas. Though the scope of this study was to test the ability to map different plant communities on and off town (two separate areas separated by the presence/absence of prairie dogs) and assess differences in RF model predictions between years.

Line 239 – Why only 100, when the default is higher (which is used for the number of nodes)? You may be ok here, but in many cases, at this point the model error is just beginning to stabilize. You could examine the impact of the number of trees on your model by looking over a range of "number of tree" values.

-100 trees were enough to allow the model error to stabilize. Adding additional trees (default=500) was computationally prohibitive.

Line 241 – Why just spectral bands as input into the models? You don't explicitly say your objective is to use just satellite imagery (and prior to the RF algorithm you used other data, e.g. to differentiated prairie dog towns and off site)

-It is explicitly mentioned in the goals of the project to 'assess the utility of using a RF model with high resolution satellite imagery to classify plant communities'. Delineating prairie dog town boundaries with GIS would be akin to outlining the boundary of any other study area of interest. Line 246 – How did you apply your models to produce predictions for prairie dog towns vs. off town locations? I think you run the predictions on two separate parts of the study area (be explicit). The 'predict' function in Program R was used to apply the models. For example 2015 on town model was used to predict class belonging to on-town areas.

Line 262 – The way your "communities" were picked seem to almost guarantee this? You picked areas "dominated" by three (or two) very different functional groups. Is this overlap more than you expected, and what is the overlap? This very much may help explain the differences between years.

-We would expect a large separation in ordination space based on how plant communities were selected. I think it is of value for plant classification studies to demonstrate that the plant communities one is classifying are actually distinct. The amount of overlap between plant communities may also factor into error rates or help explain differences between years. As mentioned prior, plant communities may be dominated by a specific functional groups, but other functional groups and species exist within these areas.

Line 287 – Are the models unstable, or does this indicate the models are accurate within years, but species representation (as seen through your methods) changes between years in heterogenous areas?

-Unsure what is meant by species representation. Given that these are perennial plant communities, it would be unlikely for major shifts in dominance to occur between successive years without a major disturbance event (i.e. fire).

Line 303 – These peaks seem like they very much may affect the production of warm vs. cool season grasses between the years as well.

-Agreed, it could affect production to some extent.

Line 304 – Was there a temperature difference between years as well? These curves seem farther apart then I would expect just based on precip.

-Temperature was similar between years, additional weather data could be added. Prior research has shown annual above ground primary production is related to current as well as previous two years precipitation. The above average rainfall at the study site in 2015 could have added to the increase in average NDVI in 2016 when compared to 2015 through an increase in cumulative biomass or production at the site.

Line 328 – Is there a transition zone at the edge of the prairie dog towns too?

-Though there are transition zones at the edge of prairie dog towns, they tend to be much sharper boundaries and occur in off-town sites. This results in much more distinct

boundaries and improves the ease of mapping colonies.

Line 340 – Based on your discussion so far, what is a more accurate thematic map? Which year is the truth, if the heterogeneous transition zones may switch categories depending on which group dominates in a given year? How about comparing this map to the two yearly maps?

-The map which includes both 2015 and 2016 data is likely the most accurate map, as demonstrated in the lower error rates. More information (spectral values across seasons and years) would produce a more accurate thematic map. As mentioned prior, switching in dominance, especially functional group dominance, between consecutive years is unlikely to occur in perennial mixed-grass prairie ecosystems without a major disturbance occurring.

Line 351 – Any limitations in the approach though? How about the lack of coefficients for your variables? I.e. good for prediction, not as good for understanding relationships

-The goal of creating predictive models is to generate good predictions. The aim of this study was prediction, not inference.

Line 353 – Why not include the variable importance for the combined model?

-This can be included.

Line 562 – Break this out to be more specific on the changes per year (what was it in 2015 and what is it now in 2016) rather than lumping the switches between types that switch both ways in the two years. If there is a dominant pattern of switch that would be useful for your conclusions.

-The table can be broken out to include specific changes between the years.

Line 588 – How are the draws mapped? These are not one of your groups, need to talk about this in the methods. Technical Corrections Line 3 – Consider replacing stability, I think this could be confused with other definitions and is not quite what you mean

-Draws were delineated visually via ArcGIS and imagery. This can be included in the methods.

Line 32 – Replace highly with high Line 46 – Replace instability with disagreement Line 66 – Remove colonization and replace dog with dogs

-Can be included.

Line 86 – First time you use the acronym NDVI. Write out fully.

-Can be included.

Line 98 – Replace several with many (or similar idea) -Can be included.

Line 101 – Replace proven with demonstrated -Can be included.

Line 109 – There are a lot more RF packages and implementation options now, compared to 2013. Standard software like R, ERDAS Imagine, QGIS, and ArcGIS have RF, as well as more specialized options like Ecognition (and even Google Earth Engine). I don't think you need this sentence, not relevant to the paper. -Can be deleted.

Line 173 – Need year you accessed the Mesonet data Line 181 – About how big are these (median, range, etc.) -Can be included.

Line 192 – How were they randomly located? -Plot frames were tossed into the area of interest to determine sampling area.

Line 236 – Did you consider other potential predictors that you could derive from these inputs?

-We didn't include additional vegetation indices from the spectral data. Additional metrics could include elevation, ecological site, etc. though given the low training error from the models this would not likely change the predictions.

Line 239 – What is the default number of nodes. Define this.

-500

Line 256 – A table of the species for each of the five groups would really help. Would also help understand what "dominated" means for your training sites. -Could be included.

Line 267 – Mishra and Crews should be outside parentheses -Can be changed.

Line 310 – What was the 2014 precip then? Dry? -Wet

Line 355 – For the town or off-site model? -Off town, can also be included.

Interactive comment on Biogeosciences Discuss., https://doi.org/10.5194/bg-2019-194, 2019.