Projecting Responses of Major North American Vegetation Types to Climate Change

Stephen Huysman, Connor Nelle

Authorship contribution
Stephen Huysman - Conceptualization, data curation, analysis code, data review, presentation (slides and delivery), figure creation, report writeup
Connor Nelle - Conceptualization, analysis code, data review, presentation (slides), report writeup

Introduction:

Climate projections almost universally show increase in temperature and changes in precipitation due to increased radiative forcing from anthropogenic emissions (IPCC 2023). Climate is a force that fundamentally shapes the distributions of plants (Stephenson 1990, 1998). Understanding how plant distributions may shift in response to climate change is an important question with implications across human society. Accurate predictions of changes in plant cover due to climate change are relevant to policy makers, land managers, recreationists, and the general public as shifts in plant distributions could affect food security, industry, and land availability.

Climate change modeling uses different models of varying severity to forecast potential changes. Our analysis will focus on two Representative Concentration Pathways (RCPs) outlined by the 2023 Intergovernmental Panel on Climate Change: RCP 4.5 and RCP 8.5. RCP 4.5 is an intermediate scenario where emissions decrease by around 2045. This means that the radiative forcing stabilizes at 4.5 watts per meter squared. RCP 8.5 is that worst case scenario: “business as usual”, where emissions do not decrease. Radiative forcing in this scenario stabilizes at 8.5 watts per meter squared.

Our analysis will identify the climatic drivers of the current vegetation cover types across the Contiguous United States (CONUS). Then, using projections of climate to the end of the century, projections of cover change across CONUS will be developed.

The objectives of the analysis are:

1. What are the climatic drivers of distributions of major vegetation types across CONUS?
2. How are the distributions of vegetation types likely to shift under future climates?

Methods:

Analysis was completed with R (R Core Team 2024) and the tidyverse (Wickham 2019). The terra package was used for spatial data manipulation (Hijmans 2024). Maps were created with the assistance of the tidyterra R package (Hernangómez 2023). Random forest modeling was performed using the ranger package (Wright 2017).

Computational efforts were performed on the Tempest High Performance Computing System, operated and supported by University Information Technology Research Cyberinfrastructure at Montana State University.

Climate data:

Historical and projected water balance data was sourced from the NPS 1 km daily Gridded Water Balance Dataset (Tercek 2021). Water balance variables such as Actual Evapotranspiration (AET) and Climatic

Water Deficit (CWD) have been found to be well correlated with vegetation distributions across spatial scales and are thought to better represent the biophysical environment of plants than simple temperature or precipitation alone (Stephenson 1998, Figure 1). We used the "summary layers" provided by the dataset as predictors for our random forest model. A total of 25 water balance variables representing either annual or seasonal averages of 2000-2019 climate were included as predictors (Table 1), as well as the soil water holding capacity layer that was used as input for the NPS Gridded Water Balance Dataset model runs.

The historical (1980-2019) water balance data from the NPS Gridded Water Balance Dataset was generated using the gridMET climate dataset as input. The projected (2000-2099) water balance data used the MACA climate dataset as input and provided projections for two emissions scenarios, RCP4.5 and RCP8.5. The MACA dataset downscaled GCM data using gridMET, therefore the historical and future time periods of the water balance dataset can be compared directly without bias correction.

Our projections were made for two future time periods: Mid-century (2040-2069) and End-century (2070-2099). Ensemble average conditions from the NPS Gridded Water Balance model were used for these projections.

Land cover data:

2019 land cover data from the NLCD dataset was used to label corresponding pixels of our climate data by cover type. Developed and water cover land-types were removed from the analysis so that only natural cover types were modeled, i.e., those cover types most likely to be highly influenced by climatic conditions. The 30m NLCD raster was upscaled to the 1km NPS Gridded Water Balance dataset cell size by resampling using "mode" pixel selection, thus our cover type data represents the most common cover types within each 1km climate grid cell.

Random forest model:

A random forest was fit to the data using the R ranger package (Wright 2017). The land cover data was used as a response, and the climate data used as predictors (Table 1, p=26). A total of 4,851,134 pixels of natural vegetation cover types were included as observations. Our random forest model was fit using 500 trees. 5 candidate variables were sampled at each split. Variable importance was scored using the impurity metric. Gini was used as the splitting rule.

Results:

Random forest model:

The random forest model had an out-of bag (OOB) error rate of 25.85%. Summer growing conditions (AET, rain, and CWD) had the highest importance, while soil water characteristics and accumulated snow water equivalents were the least important variables (Table 3). Predictive accuracy varied between classes of cover types (Table 2). The model appears to have decent performance predicting evergreen forest, deciduous forest, barren land, shrub, and grassland cover type, but does not perform as well with mixed forest or wetland cover types. Performance predicting perennial ice/snow also seems poor, however, the number of observations of this class is much lower than the other cover types included.

Projections of cover change:

Figure 5 illustrates projected changes of the land cover types for the RCP 4.5 and 8.5 scenarios from current area to mid-century and end-of-century. Shrub-scrub cover is projected to increase relatively linearly across both scenarios. However, under RCP 8.5 projections shrub-scrub is projected to increase in land cover by roughly 500,000 square kilometers. This is possibly correlated with the reductions in evergreen forest, which has decreased by at least 200,000 kilometers under both scenarios, and grassland cover, which shows a linear negative prediction with an extra 250,000 square kilometers lost under RCP 8.5 . Mixed forest is also predicted to decrease substantially by mid-century, retaining a similar level until the end of the century. Woody wetland is the most variable wetland type, increasing by 500,000 kilometers squared by mid-century under RCP 8.5, with the RCP 4.5 scenario catching up by the end of the century. Emergent herbaceous wetlands remain mostly unchanged. Deciduous forest cover under both scenarios is predicted to increase by mid-century, then decrease slightly by the end of the century. Barren land is projected to decrease by mid-century and remain constant until the end of the century. There is no important or interpretable change in the perennial ice/snow land cover.

Discussion:

Climatic predictors of cover class:

Stephenson 1998 showed a clear separation of several North American major vegetation types using only annual AET and CWD (Figure 1). We did not find such a clear separation between vegetation types in our analysis, with all vegetation types being present across nearly all observed values of annual AET and CWD (Figure 2). However, the regions of highest pixel density in annual AET by annual CWD space did in many cases correspond closely with the distributions reported by Stephenson. For example, the relative positions of Evergreen and Deciduous forests match those reported by Stephenson, with Deciduous forests occupying regions of higher AET. However, given the large amounts of overlap between cover types in bivariate AET x CWD space, we decided to incorporate more dimensions of climate in order to obtain better classification performance. In the end, we included all 24 seasonal climate variables provided by the NPS Gridded Water Balance Dataset, as well as annual accumulated snow water equivalents and soil water holding capacity (Table 1).

The variable importance table from our random forest model shows summer growing conditions (AET, Rain, CWD) being the most important predictors of cover class, and soil characteristics (soil water, soil WHC) and accumulated snow water equivalents being the least important (Table 3). The water balance model accounts for soil dynamics implicitly through the calculation of AET and CWD, which move water to and from the soil water storage as it becomes available to plants and is used. Given the higher performance of seasonal AET and CWD variables compared with the soil water characteristics and ACCUMSWE, it seems that the water balance model is accurately accounting for soil moisture and snow storage dynamics as they relate to the biophysical environment of plants.

Interestingly, the model found runoff during spring, fall, and winter to also be an important classifier of cover class. Runoff should not directly impact plants aside from erosion resulting in mortality or habitat loss, which is a dynamic not accounted for in our model. Runoff could represent periods of time that the soil is fully saturated, which may be important for predicting some cover classes, such as wetlands. Runoff from the water balance model used here is also well correlated with streamflow (Thoma 2020), and perhaps could indicate potential for riparian or hydric (wetland) vegetation. Whatever dynamics the runoff variables represent, it seems to be predicting some aspect of the impacts of climate on cover class that the bivariate relationship of AET and CWD (Stephenson 1998) cannot account for alone.

Model performance:

Overall, our model had decent performance classifying the 2019 NLCD cover types with an out-of-bag prediction error of 25.6%. Model performance could be improved by specifying cover types with a higher taxonomic resolution, for example modeling bioclimatic niches for regional vegetation types or individual species instead of the broad classes here that are not true taxonomic groups of plants. Performance could also be improved by using finer-scale climate variables, such as the temperature or water balance during specific months instead of across seasons. Chang et al. 2014 used a random forest to model the bioclimatic niche space of whitebark pine using climate and water balance means during key months in the species growing cycle with an OOB error of 16.1%. Our model performance compared favorably to this given the inherent noise in ecological data and the broader taxonomic and spatial scales at which we are predicting cover classes.

Our random forest model had varying performance predicting different cover types (Table 2). While performance was relatively good with cover types such as deciduous and evergreen forest, performance was especially poor with mixed forest and wetland cover types. At the continental scale studied here, the annual AET and CWD values for mixed forest types, which consist of mixed deciduous and evergreen trees, have a large amount of overlap with deciduous and evergreen forest cover types (Figure 2). While our random forest model was fit to a higher dimensional than just annual AET and CWD, it is likely that the climate space of mixed forest overlaps much of the climate space of the individual forest types, and prediction at this scale may be difficult.

Our model also had difficulty predicting wetland cover types (Emergent Herbaceous Wetlands and Woody Wetlands). Annual AET and CWD values for these cover types were not provided by Stephenson 1998, so we did not have *a priori* expectations for the bioclimatic envelope these cover classes occupy. Wetland formation likely relies on many factors outside of the climatic water balance alone, such as soil characteristics and broader drainage patterns that are simply not accounted for by the water balance model used here. Our projections show a great deal of the Southeast United States changing to woody wetland or emergent herbaceous wetlands in the future projections, with much of the land originally occupied by Coniferous or Mixed forest (Figures 3 & 4). It seems unlikely that the entire state of Alabama could transition from mostly mesic evergreen forest cover to hydric woody wetlands. The NPS gridded water balance projections show an increase in AET across the Southeast United States (Tercek *et al.* 2023), where our model is predicting these broad shifts to woody wetland cover types. Wetlands are inherently characterized by high AET, as they have persistently saturated soils and experience little climatic water deficit. This projected increase in AET is likely resulting in our model classifying areas as wetlands which may not have the hydric potential to support wetland cover types regardless of the potential of the climate to do so.

Projections of cover change interpretation:

Our projections of vegetation change appear to show minor change at the continental scale, with little to no change in total area covered predicted for most cover classes with the exceptions of shrub/scrub and woody wetland cover types (Figure 5). However, at more local scales the projected changes can be dramatic. By the end of the century, the Everglades in Florida (25.7459° N, 80.5550° W), defined by expansive emergent herbaceous wetlands, are projected to either shrink greatly (RCP4.5) to vanish entirely (RCP8.5) (Figure 4). The Black Hills in South Dakota (43.9939° N, 103.7718° W) are projected to see an expansion of grassland into the existing coniferous forest zone, with almost all of the existing coniferous forest cover being replaced in our worst-case RCP8.5 scenario (Figure 4). The Sky Islands in Arizona are a location of note as they are almost completely taken over by surrounding shrub-scrub land cover. Reductions of a cover class in one locality are counterbalanced by expansions of that cover class in another region, which leads to patterns such as the near constant area under coniferous forest in our projections (Figure 5). There is an increase in evergreen patchiness in the northern rocky range, which is consistent with projected evergreen habitat shrinkage to higher elevation (Halofsky 2015). Therefore, our model shows highly unequal effects of climate change on the land cover of CONUS by the end of the century.

Conclusion:

Overall, our projections of cover change across CONUS by the end of the century can be made only with much uncertainty. While the distributional shifts in shrub-scrub and evergreen forest are plausible and conservative, the behavior of our model when predicting wetland and mixed vegetation types brings into question the validity of all projections. Future analyses could improve our model by identifying the factors necessary for predicting wetland and mixed vegetation cover types and incorporate them into our model.

Additionally, our model does not account for many important processes that can alter distributions of plants. Wildfire and plant communities interact in complex ways to determine the cover of the landscape (Hill and Field 2021), and altered wildfire occurrence with climate change could result in land cover different from what we are able to predict with climate alone.

Perhaps more importantly, our work removes anthropogenic cover types from our historical analysis and projections, with the assumption that current anthropogenic cover types will be constant to the end of the century. This assumption is almost certainly false, as increasing human population will necessitate the use of more land for agricultural or urban uses. Human influence across the terrestrial biosphere is pervasive, and even the cover classes we designate as natural are to some degree influenced by human use. Intensive alteration of the terrestrial biosphere by humans has led to the creation of the concept of an "anthrome" - which is an ecosystem with novel anthropogenic ecological processes (Ellis 2011). In light of this recognition of the near universal influence of humans across the globe, it is unlikely that land cover types can be predicted with complete accuracy without accounting for patterns in land use by humans.

It is important to consider that this analysis models a small portion of an incredibly complex system and these predictions should be taken as potential limited shifts which inform future modeling efforts. However, with future tweaks and refinement, this modeling strategy could have significant implications for land management. Incorporating additional factors such as fire regimes, invasive species, and human land-use could improve predictive accuracy. This study serves as a first step in understanding the complex relationships of vegetation types in the CONUS under climate change.

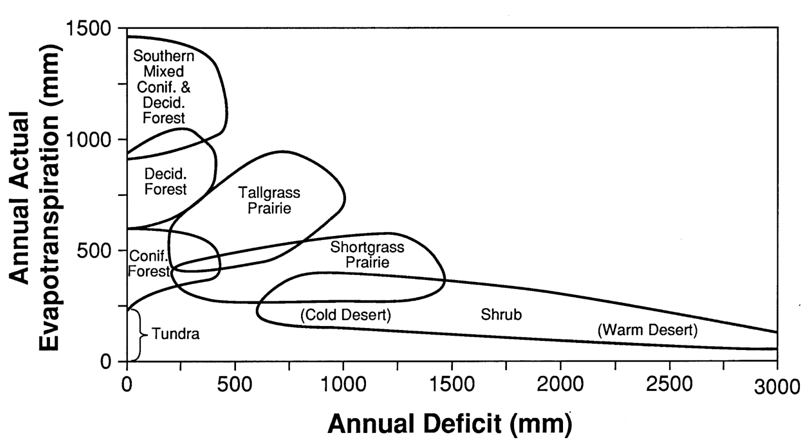

Figure 1: Deficit vs. Evapotranspiration graphic from Stephenson 1998

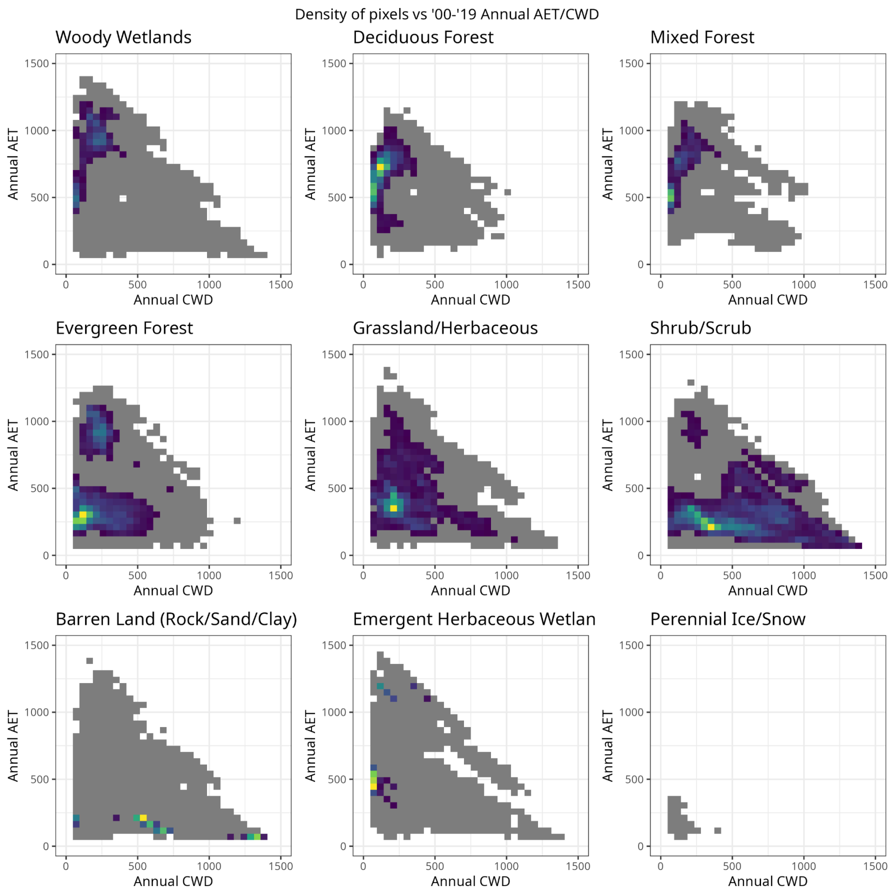

Figure 2: Annual CWD and AET pixel density for vegetation cover types. Areas of lower density are shown in grey to highlight areas of highest density. Compare areas of highest density to plant distributions shown in Figure 1.

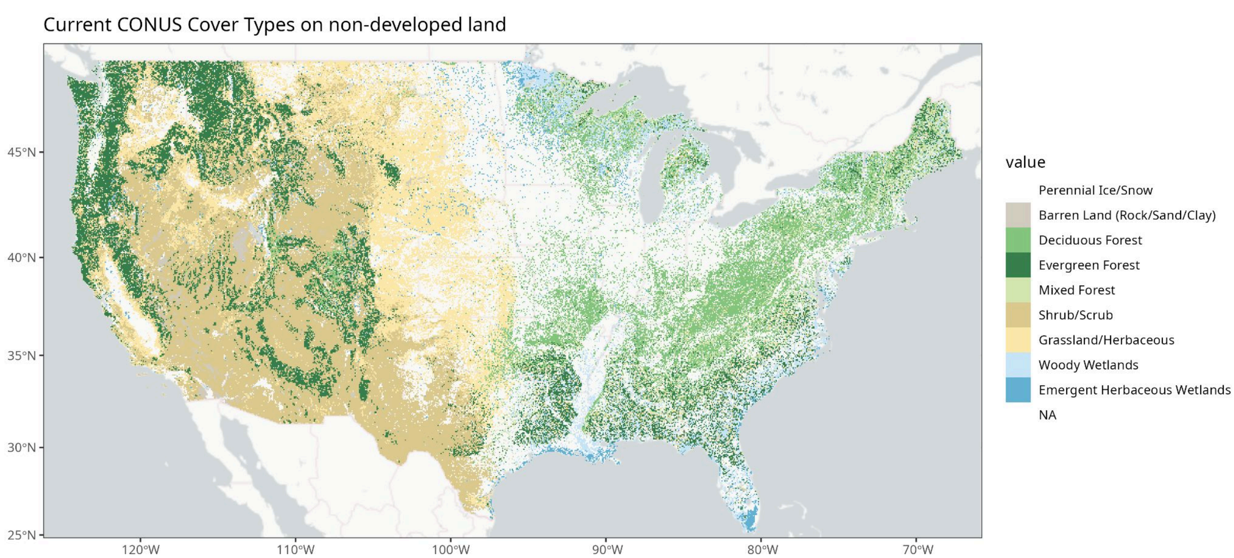

Figure 3: Raw response data - 2019 land cover types from NLCD. Developed land types were removed so that only natural land cover types (i.e., those more likely to be highly driven by climate) are shown. Note that the number of perennial ice/snow pixels is so small that they are virtually invisible at this scale and white pixels shown here are the underlying basemap.

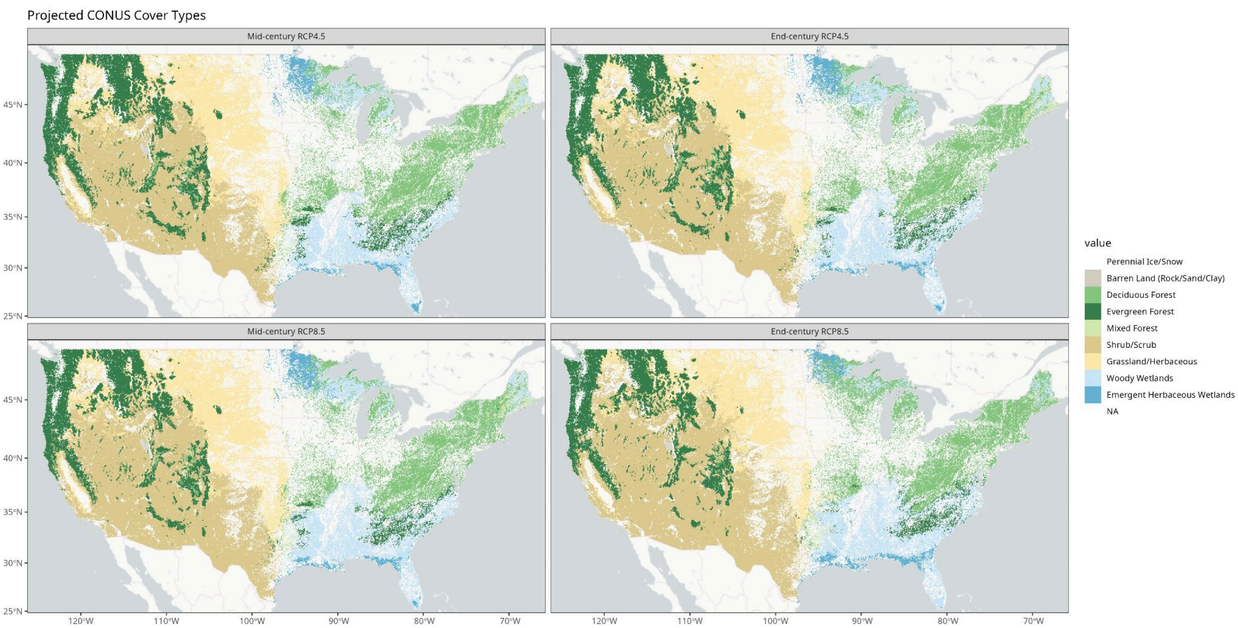

Figure 4: Projections of land cover change across CONUS. Mid-century and late-century time periods are on the left and right, respectively. RCP4.5 and RCP8.5 are shown on top and bottom, respectively.

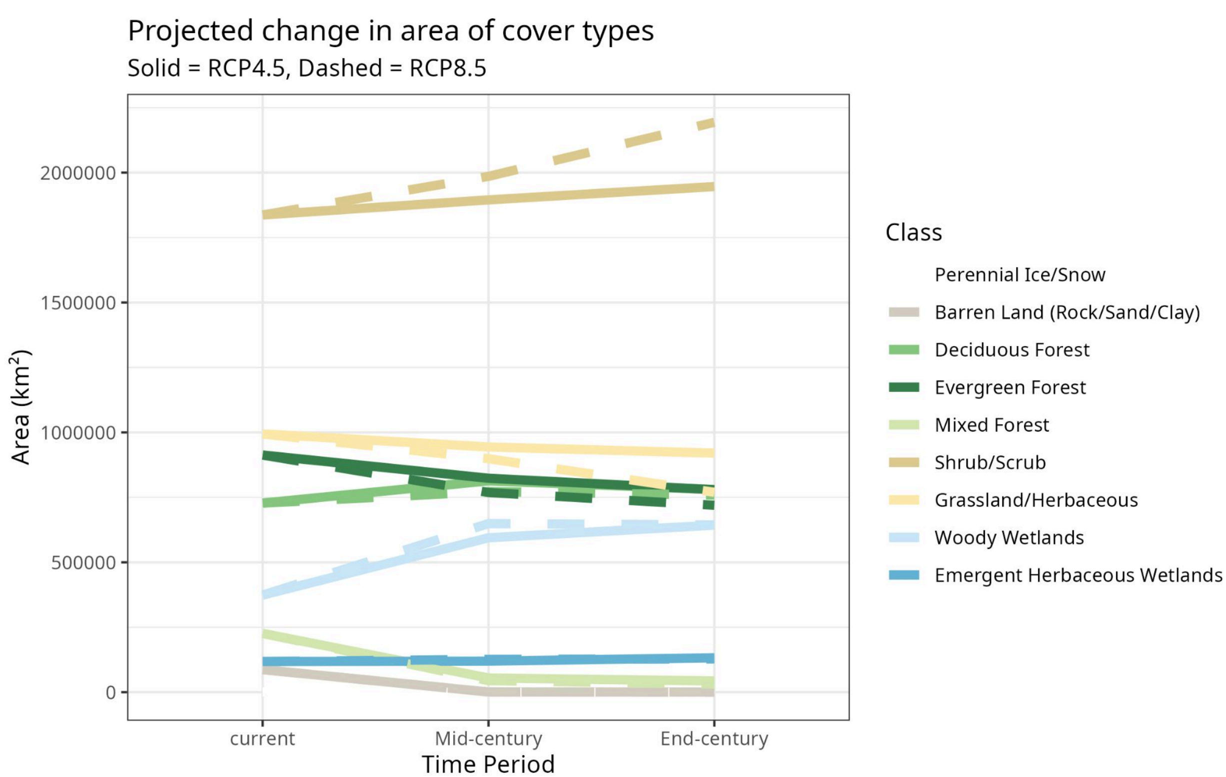

Figure 5: Areal sums (km^2) of land cover types across time periods and emissions pathways.

Variable	Time Period(s)
Potential Evapotranspiration (PET)	Spring, Summer, Fall, and Winter
Actual Evapotranspiration (AET)	Spring, Summer, Fall, and Winter
Climatic Water Deficit (CWD)	Spring, Summer, Fall, and Winter
Rain	Spring, Summer, Fall, and Winter
Soil Water	Spring, Summer, Fall, and Winter
Runoff	Spring, Summer, Fall, and Winter
Accumulated Snow Water Equivalents (accumSWE)	Annual
Soil Water Holding Capacity	N/A

Table 1: Water balance variables from the NPS Gridded Water Balance dataset (Tercek 2021) included as predictors. A total of 26 predictors were included.

True \ Predicted	Perennial Ice/Snow	Barren Land	Dec. Forest	Everg. Forest	Mixed Forest	Shrub/ Scrub	Grassland/Herbaceous	Woody Wetlands	Emergent Herbaceous Wetlands
Perennial Ice/Snow	145	254	0	103	0	61	23	0	0
Barren Land	194	37937	4614	4699	717	20473	4192	1514	781

Dec. Forest	0	893	529102	31509	41874	12533	21883	30880	4242
Everg. Forest	45	2239	40637	571069	31592	111521	32485	49149	1769
Mixed Forest	0	318	80829	38768	56313	6984	4068	20793	616
Shrub/ Scrub	45	11181	14667	127315	6899	1426354	91684	10425	3119
Grassland/Herbaceous	18	2300	21885	49183	4979	94499	729565	9205	5304
Woody Wetlands	0	706	42599	53515	18077	12472	8915	196872	10184
Emergent Herbaceous Wetlands	0	571	7295	4096	1008	8330	14756	15397	49896

Table 2: Confusion matrix of true vs predicted classifications from our random forest model.

Variable	Importance
aet_summer	275533.6221
rain_summer	270917.5229
cwd_summer	224760.0964
runoff_spring	209007.051
runoff_fall	182627.9193
runoff_winter	175450.0742
rain_spring	165814.9098
rain_fall	164605.5208
aet_fall	159639.7219
pet_summer	145584.6716
rain_winter	138920.4598
cwd_spring	137486.0428
cwd_fall	137112.0661
pet_spring	128640.0846
pet_fall	125718.196
aet_spring	121946.8688
runoff_summer	117401.8019
cwd_winter	109382.6815
aet_winter	108529.4888
pet_winter	108344.3547
soil_water_summer	103570.0563
soil_water_winter	98123.11789
soil_whc	98066.31854
soil_water_fall	93224.37371
accumswe	92594.93893
soil_water_spring	88778.36474

Table 3: Variable importance (Impurity metric) for random forest model.

Appendix:

The code used for this analysis is available here: https://github.com/shuysman/conus-veg-cover

References

1. IPCC, 2023: *Climate Change 2023: Synthesis Report.* Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]. IPCC, Geneva, Switzerland, pp. 35-115, doi: 10.59327/IPCC/AR6-9789291691647.
2. Chang, T., Hansen, A. J. & Piekielek, N. Patterns and Variability of Projected Bioclimatic Habitat for Pinus albicaulis in the Greater Yellowstone Area. *PLoS One* **9**, e111669 (2014).
3. Ellis, E. C. Anthropogenic transformation of the terrestrial biosphere. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **369**, 1010–1035 (2011).
4. Stephenson, N. L. Climatic Control of Vegetation Distribution: The Role of the Water Balance. *The American Naturalist* **135**, 649–670 (1990).
5. Stephenson, N. Actual evapotranspiration and deficit: biologically meaningful correlates of vegetation distribution across spatial scales. *Journal of Biogeography* **25**, 855–870 (1998).
6. Tercek, M. T. *et al.* Historical changes in plant water use and need in the continental United States. *PLOS ONE* **16**, e0256586 (2021).
7. Tercek, M. T., Gross, J. E. & Thoma, D. P. Robust projections and consequences of an expanding bimodal growing season in the western United States. *Ecosphere* **14**, e4530 (2023).
8. Thoma, D. P. *et al.* Water balance as an indicator of natural resource condition: Case studies from Great Sand Dunes National Park and Preserve. *Global Ecology and Conservation* **24**, e01300 (2020).
9. Halofsky, J. E., Peterson, D. L., Dante-Wood, S. K., Linh Hoang, L. H., Ho, J. J., & Joyce, L. A. (2018). Climate change vulnerability and adaptation in the Northern Rocky Mountains, Part 1.
10. R Core Team (2024). _R: A Language and Environment for Statistical Computing_. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
11. Marvin N. Wright, Andreas Ziegler (2017). ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. Journal of Statistical Software, 77(1), 1-17. doi:10.18637/jss.v077.i01
12. Hernangómez D (2023). "Using the tidyverse with terra objects: the tidyterra package." _Journal of Open Source Software_, *8*(91), 5751.
13. Hill, A. P. & Field, C. B. Forest fires and climate-induced tree range shifts in the western US. *Nat Commun* **12**, 6583 (2021).
14. Hijmans R (2024). _terra: Spatial Data Analysis_. R package version 1.7-83, <https://CRAN.R-project.org/package=terra>.
15. Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H (2019). "Welcome to the tidyverse." _Journal of Open Source Software_, *4*(43), 1686. doi:10.21105/joss.01686 <https://doi.org/10.21105/joss.01686>.