Floristic Quality Index

Floristic Quality Assessments (FQA) utilize the vascular plant species at a site as an indicator of habitat quality. The fundamental assumption guiding the use of plants as indicators of habitat quality is that different species respond differently to the types and frequencies of disturbance. At one end of this disturbance spectrum are species which are able to persist, or may be introduced to an area, after certain types of anthropogenic disturbance - e.g. compaction of soil via heavy vehicles. On the other end of the spectrum are taxa which may only grow in areas which receive episodic natural disturbances characteristic of their ecosystem - e.g. a 100 year flood in a wetland. The subjectively estimated likelihood of a species either persisting, or being removed from a site due to disturbance is expressed as a Conservatism Value (C-Value). C-Values range from 1 to 10 for plants native to North America, and 0 for plants introduced to the continent since colonization by European Settlers, with plants which are not tolerant of disturbance being at the upper end of the spectrum.

The use of FQA are uncommon in the Bureau of Land Management, perhaps in part due to the FQA originating in the Midwest, and the assignment of C-values being a task which requires considerable amounts of resources (Spyreas (2019)). A further requirement which hampers the utilization of these metrics are that each individual plant species, often times including subspecies, is assigned a separate C-value, the number of land management professionals which are capable of distinguishing taxa at these resolutions, and have time assign C-values, are limited (Kramer & Havens (2015), Ahrends et al. (2011), Morrison (2016)). Other possible limiting factors are that the FQI indices have been traditionally associated with the portions of Natural Resources focused on designation of parcels for conservation and preservation, rather than management of these parcels.

While C-values exist for virtually all states East of the Continental Divide, Colorado is one of only two states with significant surface lands administered by the BLM which has existing C-values for every documented member of it's flora (Spyreas (2019)). In fact BLM Colorado staff, including the lead state botanist, assisted in developing the states C-values (Smith et al. (2020)). Here we utilize FQA, to supplement our formal AIM analysis, and to develop a map of the habitat quality for lands managed by the Uncompahgre Field Office.

The Floristic Quality (FQ) Assessments are comprised of two core indices Mean Coefficient of Conservatism (Mean CoC) and Floristic Quality Index (FQI) (see Appendix A for equations). While many novel permutations of these calculations exist, they seem to offer little additional insight to the main pair of indices and appear only useful for niche applications (Spyreas (2019)). As the general goal of an FQA is the assessment of parcels of land, the location of study areas across different habitat types is an accepted use, and as we used a weighted stratified sample design our points meet the assumptions implicit in the sample design (Spyreas (2019)). FQ Assessments have yet to converge on a standardized size for conducting the species inventory, and while Mean-CoC is affected by plot size FQI is relatively robust against small plot size. Regardless in similar systems plots of similar size as AIM plots have been shown to be adequate for noting enough species to calculate both main indices (Spyreas (2016)). In several applications C-values have been shown to be stable across sampling time, in part perhaps due to a propensity for many species at a site to share the same C-values (Spyreas (2016), Matthews et al. (2015), Bried et al. (2013)). C-values have also been shown capable of distinguishing habitat variability more effectively than two traditional diversity metrics (Simpsons and Shannons) (Taft et al. (2006)). Practitioners of varying degrees of skill are likely to have minimal effects on the estimates of the FQA indices due to the species encoding some degrees of redundant information (Bried et al. (2018), Spyreas (2019)). Further, the calculations of Mean C and FQI, are relatively independent of the species richness, or diversity at any one plot, and are capable of reflecting the C-values of the species present, rathern than other diversity metrics *per se*.

Utility and Limitations of FQA FQA scores are tied to the regional list of C-values, accordingly they cannot be compared across regions with independently developed C-values, in other words the FQA scores of sites in the mixed grass prairies of Kansas and Colorado are incomparable. Scores have the potential to be misleading if compared across major habitats, (e.g. comparing Sagebrush to Salt desert) but generally appear robust against this (Spyreas (2019)). The indices are relatively boundless, e.g. we could visit plots which we designate high quality and use them as benchmark for FQI values, but we cannot incorporate metrics e.g. land > 4 is ‘good’, until we perform these activities to determine locally relevant measurements.

> *‘. . . tolerance of anthropogenic disturbance and exclusivity to remnant habitats are the only validated criteria for defining FQA.’*
>
> *“. . . FQA conveys two things about high conservative species: (1) All else being equal, they have greater conservation value, and (2) they reflect a site’s history of minimal disturbance and degradation.”*
>
> — Spyreas 2019

Accordingly, in our scenario we are interested in using FQA to identify sites which have histories of evident degradation.

Methods

Cleaned AIM species richness data were imported from TerrAdat and joined to the CNHP C-values using the lookup keys developed for the Functional Diversity in Section 11. All species from the plot based species richness which were not unambiguously identified to terminal taxon were removed from analysis. The Mean Coefficient of Conservatism and Floristic Quality Index were calculated via the formulation in Appendix 1. To determine whether their were bias in the Mean-CoC values between the strata, a linear model with a single set categorical predictors and a single continuous response (an ‘ANOVA’, or Analysis of Variance) was used, with a Kruskall-Wallis Post-hoc test with 95% Confidence Intervals. The strata used for detecting differences were those developed in **SECTION XXX**.

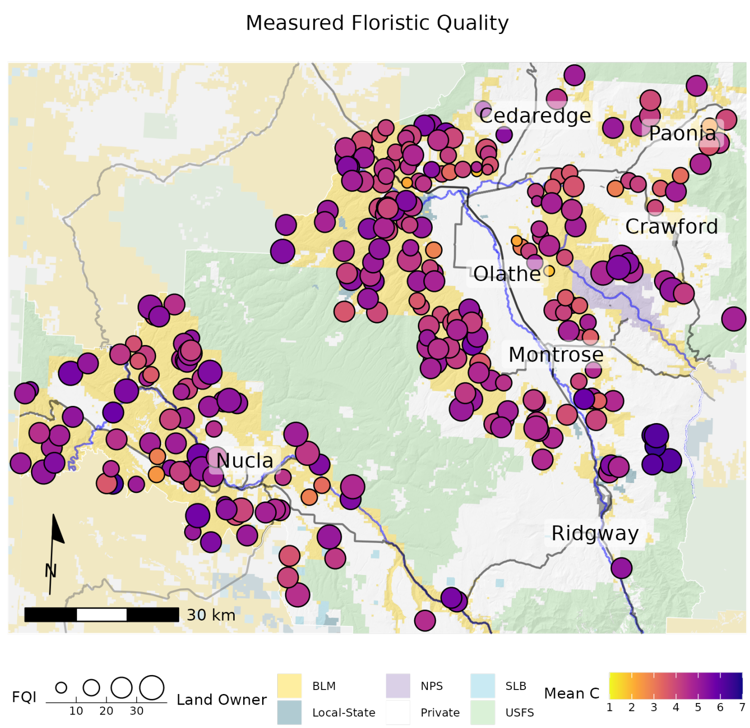

Figure 1: FQI Values observed at AIM Plots

The floristic quality assessment data are theoretically independent of ecological sites. Accordingly, we can create a statistical model using the plots which were sampled, and using this model predict the values of floristic quality across the field office onto a map. Based on Figure 1, and our field experience collecting these data, we suspected that the variables which most strongly predict Mean-CoC scores, and which we could readily acquire or create were: 1) distance to nearest road, 2) patch size of federal public lands, 3) human population density within ~ 10 km (~6 miles), 4) elevation. These analyses occurred across the extent of the mapped area in Figure 1), at a resolution of 90 meters. All of these variables were used as predictors of Mean-CoC scores in our statistical model.

Predictor variables were created using the following processes. To calculate the distance of each pixel of BLM land from the nearest road, the U.S. Census Bureau's 'roads' data set was acquired through 'tigris' and simplified using 'st_simplify'; the distance function of 'terra' was used to calculate the distance from each pixel to the nearest road (Walker (2022), Hijmans (2022), Pebesma (2018)). The U.S. Census Bureau roads data set contains nearly all of 'major' dirt and gravel roads across the field office, included many used for historic Uranium mining activities. To calculate the patch size of federal lands, i.e. all area across BLM, Forest Service, and National Park Service management areas etc., federal lands were queried from the PAD-US database and then areas dissected by roads were erased; the area of each parcel were then calculated using 'st_area' from sf, and converted from meters squared to hectares (Gergely & McKerrow (2022)). The patches were then converted into a raster using 'rasterize' from terra, and the size was saved as an attribute. To create estimates of population density within a distance of ~ 10km of each pixel of BLM Land, population density data at 30m resolution were downloaded from HDX, and summed to 90m resolution; these values were then summed using rolling windows in terra (Tiecke et al. (2017)). 90 meter resolution elevation data were acquired from EarthEnv (Tuanmu & Jetz (2014)). All data sets were cropped and re-sampled to a template to ensure optimal alignment of raster cells.

To determine whether any predictor variables would violate assumptions of independence, by being collinear, variance inflation factors were calculated using the 'function' vifstep in the package usdm with a theta cut off 10 (Naimi et al. (2014)). This function was used on a subset of 5000 random pixels regularly dispersed across the extent of the sample area, which had their value of all predictors extracted from them. No theta scores were high enough that they suggested any combination of variables should be removed from the analysis, but notable collinearity existed between 'elevation' and 'X coordinates'; this is likely due to the increase in elevation from the Colorado Plateau into the Rocky Mountains, which follows a longitudinal (West to East) trend.

In order to create a model which could capture the variation in our data and generalize them to predict Mean-CoC, a maximal model of the term *glm(mcoc_r ~ road_dist * pop_density * patch_size * elevation * xcoord * ycoord)* was created. This model was then fed into the 'dredge' function from the package MuMIn, which creates all smaller models, and evaluates them in an information theoretic framework (Bartoń (2022)). All tops models, those with ΔAIC < 2.0, were each selected for further analyses. Each model was checked for the effect of spatial autocorrelation in the residuals by identifying neighboring points using 'graph2nb' and converting them into three neighbor lists using 'nb2listw' (both functions from 'spdep'), the Moran's Index for each model was very low, despite the Moran's Index for the predictor variables being very high and showing significant clustering (Bivand (2022)). Subsequent to the interpretation of the maps produced by the stacked model outputs, considering that the model including the coordinates did not suffer from autocorrelation, and that the X coordinate was often selected as a predictor in top models, we believe that minor collinearity between the X coordinate and the strongest predictor Elevation indicated that the term was redundant.

A new full model, without the coordinates, was created in the form *glm(mcoc_r ~ road_dist * pop_density * patch_size * elevation)* and passed to the 'dredge' function in MuMIn. The results of the top models, those with ΔAIC $<= 2.13$ were checked for spatial autocorrelation, using the same methods as the models above. The general rule of thumb for only utilizing models with ΔAIC < 2, was bent in order to accommodate two more models which incorporated terms we have observed to be ecologically relevant, and are well supported in the landscape ecology primary literature (Symonds & Moussalli (2011)). After evaluation of the stacked prediction map from these top models, it was used for the final analysis, in part because removal of the x-coordinates made it more easy to interpret, and it was able to utilize the other predictors which are known to affect species compositions, both universally and strongly.

All six of the optimal models, which were in the ensemble model, were individually predicted into space and the weighted means, based on each models AIC weight, were used to generate an ensembled prediction layer (Symonds & Moussalli (2011), Bartoń (2022), Dormann et al. (2018)). Spatial predictions of the statistical model were performed using terra Hijmans (2022). The ensembled prediction was then clipped to the extent of UFO BLM administered lands.

Discussion

The Mean-CoC values varied by stratum (Analysis of Variance (ANOVA)), follow-up tests (Kruskall-Wallis) indicated that Salt Desert and Sagebrush sites did not differ from each other, nor Mixed Mountain Shrub and Pinon-Juniper from each other (Figure 1). Results indicate that all members of these two groups differed across them, e.g. Sagebrush differed from both Mixed Mountain Shrub and Pinon-Juniper and *vice versa*. The two lower elevation strata, Salt Desert and Sagebrush, are generally more accessible and have higher recreational land uses, and had the lowest Mean-CoC. While the values did vary, the extent of the variation was minor, with the range of variation from the median values 4.18 - 4.86 and a mean range of 4.21 - 4.76, given that these values occur over a range of 0-10, we considered this to be a negligible difference. Which was not indicative of major biases in the C-values ascribed to plants themselves, but rather indicative of actual land conditions, and that all main vegetation types in the UFO could be analysed together.

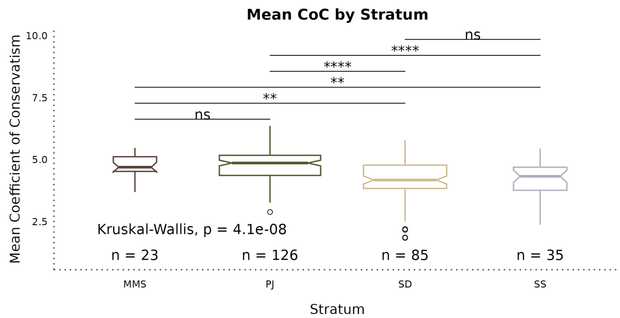

Figure 2: Comparision of median values by Stratum

The generalized linear models were used to predict the floristic quality of unsampled areas on BLM Land. Based on AIC model selection a total of 5 top models were retained, and these were then combined into a single model, of which the weights used to make a stacked ensembled prediction of the Mean-CoC. The results from these top models indicate that all predictors, elevation, human population density, Federal land patch size, and distance to nearest road, affect the Mean-CoC, but that by far the strongest predictor is elevation (Figure 4, Appendix). Elevation was such a useful predictor, that under certain statistical paradigms (e.g. forward or backward selection via null hypothesis) our models may be simplified to include it as the only predictor. However, given that C-values reflect the tolerance of a species to anthropogenic disturbances this approach would be questionable. Rather we suspect that elevation is strongly associated with disparities in land use intensity, but more importantly also the inability of lower elevation sites to recover from previous disturbances relative to higher elevation areas. The other predictors had

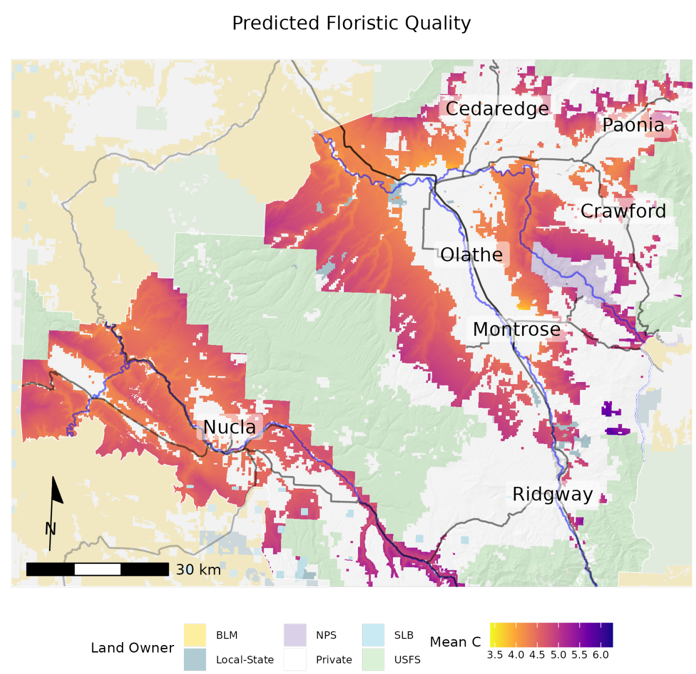

Figure 3: Consensus predictions of FQI Values across the UFO Field Office (excluding lands East of Paonia along the North Fork of the Gunnison

the expected effect on Mean-CoC as ex-
pected, Mean-CoC slowly increases as nearby human population density decreases, and FQI slowly increases as the size of the patch of natural lands and the distance from the nearest road increase.

From the predicted Mean-CoC values we can infer that habitat quality follows a consistent trend across the field office (Figure 2). In general, the lowest elevation sites display the lowest values of FQI and Mean-CoC. This is especially apparent between Montrose and Grand Junction on Highway 50. These and several other areas of the Salt Desert along Highway 50, and near Crawford are areas of concern. However, not all Salt Desert is inherently in this condition as can be seen from the values for a moderately sized parcel of BLM managed land immediately Southeast of Montrose; indicating that both anthropogenic disturbances and elevation in these areas are causal factors for these low values. Areas with low FQI values represent highlight potential restoration needs and the Mean-CoC prediction model can be a tool to help prioritize such actions. The highest elevation portions of BLM managed land are in the best habitat condition, and very good conditions range into Pinon-Juniper, including most of the disturbed areas around historic mining in the West End. By these metrics considerable portions of the sagebrush habitat occupied by Gunnison sage-grouse near Crawford are also in good condition.

Subjective interpretation suggests these results seem very similar to those generated by the much more time intensive comparison of Ecological Sites to Quantitative Benchmarks. This relationships considers serious consideration in the use of FQA as a proxy of site condition, and may warrant applications for rapid surveys prior to certain land use decisions, in time frames when AIM cannot be implemented. It also provides a second independent set of data which may help to contextualize sites which were unable to have their Ecological Sites correlated, are lacking Ecological Site Descriptions, or are difficult to interpret.

Appendix A - Indices

Mean Coefficient of Conservatism

$$\overline{C} = \frac{\sum C_i}{S}$$

Where:

$\overline{C}$ is the Mean Coefficient of Conservatism, or for short Mean C
S is the number of species included in the calculation
C_i in particular C is the Conservatism Value (C-Value), for each $_i$ of the S at the site
$\sum$ is an operator, meaning that we will calculate the sum of all C-Values, C

Floristic Quality Index

$$FQI = \overline{C} * \sqrt{S}$$

Where:

$\overline{C}$ is the Mean Coefficient of Conservatism, or for short Mean C
$\sqrt{S}$ is the square root of the number of species included in the calculation

Equations from Swink et al. (1994), and modified for simpler formulations.

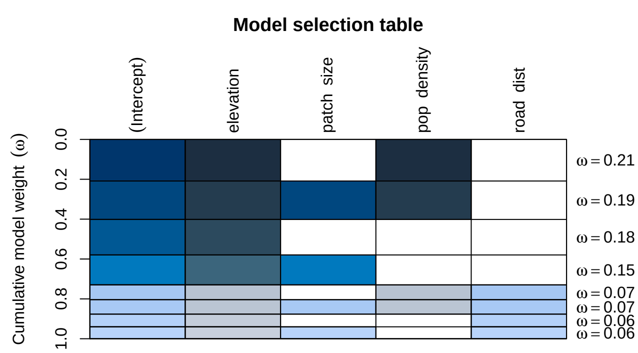

Figure 4: Contributions of each term to each model, and each model sized by contributions to ensemble prediction. Note the left-most column is not a predictor, and only the top 6 models were used.

References

Ahrends, A., Rahbek, C., Bulling, M. T., Burgess, N. D., Platts, P. J., Lovett, J. C., Kindemba, V. W., Owen, N., Sallu, A. N., Marshall, A. R., et al. (2011). Conservation and the botanist effect. *Biological Conservation*, *144*(1), 131–140.

Bartoń, K. (2022). *MuMIn: Multi-model inference*. https://CRAN.R-project.org/package=MuMIn

Bivand, R. (2022). R packages for analyzing spatial data: A comparative case study with areal data. *Geographical Analysis*, *54*(3), 488–518. https://doi.org/10.1111/gean.12319

Bried, J. T., Allen, B. E., Azeria, E. T., Crisfield, V. E., & Wilson, M. J. (2018). Experts and models can agree on species sensitivity values for conservation assessments. *Biological Conservation*, *225*, 222–228.

Bried, J. T., Jog, S. K., & Matthews, J. W. (2013). Floristic quality assessment signals human disturbance over natural variability in a wetland system. *Ecological Indicators*, *34*, 260–267.

Dormann, C. F., Calabrese, J. M., Guillera-Arroita, G., Matechou, E., Bahn, V., Bartoń, K., Beale, C. M., Ciuti, S., Elith, J., Gerstner, K., et al. (2018). Model averaging in ecology: A review of bayesian, information-theoretic, and tactical approaches for predictive inference. *Ecological Monographs*, *88*(4), 485–504.

Gergely, K. J., & McKerrow, A. (2022). PAD-US: National inventory of protected areas. *US Geological Survey*.

Hijmans, R. J. (2022). *Terra: Spatial data analysis*. https://CRAN.R-project.org/package=terra

Kramer, A. T., & Havens, K. (2015). Report in brief: Assessing botanical capacity to address grand challenges in the united states. *Natural Areas Journal*, *35*(1), 83–89.

Matthews, J. W., Spyreas, G., & Long, C. M. (2015). A null model test of floristic quality assessment: Are plant species' coefficients of conservatism valid? *Ecological Indicators*, *52*, 1–7.

Morrison, L. W. (2016). Observer error in vegetation surveys: A review. *Journal of Plant Ecology*, *9*(4), 367–379.

Naimi, B., Hamm, N. a.s., Groen, T. A., Skidmore, A. K., & Toxopeus, A. G. (2014). Where is positional uncertainty a problem for species distribution modelling. *Ecography*, *37*, 191–203. https://doi.org/10.1111/j.1600-0587.2013.00205.x

Pebesma, E. (2018). Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal*, *10*(1), 439–446. https://doi.org/10.32614/RJ-2018-009

Smith, P., Doyle, Georgia, & Lemly, J. (2020). *Revision of colorado's floristic quality assessment indices*. Colorado Natural Heritage Program. https://cnhp.colostate.edu/download/documents/2020/CO_FQA_2020_Final_Report.pdf

Spyreas, G. (2016). Scale and sampling effects on floristic quality. *PloS One*, *11*(8), e0160693.

Spyreas, G. (2019). Floristic quality assessment: A critique, a defense, and a primer. *Ecosphere*, *10*(8), e02825.

Swink, F., Wilhelm, G., et al. (1994). *Plants of the chicago region*. Indiana Academy of Science.

Symonds, M. R., & Moussalli, A. (2011). A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using akaike's information criterion. *Behavioral Ecology and Sociobiology*, *65*, 13–21.

Taft, J. B., Hauser, C., & Robertson, K. R. (2006). Estimating floristic integrity in tallgrass prairie. *Biological Conservation*, *131*(1), 42–51.

Tiecke, T. G., Liu, X., Zhang, A., Gros, A., Li, N., Yetman, G., Kilic, T., Murray, S., Blankespoor, B., Prydz, E. B., et al. (2017). Mapping the world population one building at a time. *arXiv Preprint arXiv:1712.05839*.

Tuanmu, M.-N., & Jetz, W. (2014). A global 1-km consensus land-cover product for biodiversity and ecosystem modelling. *Global Ecology and Biogeography*, *23*(9), 1031–1045.

Walker, K. (2022). *Tigris: Load census TIGER/line shapefiles*. https://CRAN.R-project.org/package=tigris