Supplementary materials for:

Williams, R. N., D. C. Dauwalter, R. F. Thurow, D. P. Philipp, J. E. Williams, and C. A. Walser. 2019. Identification and utility of native fish conservation areas in the Upper Snake River basin. Pages 135-159 *in* D. C. Dauwalter, T. W. Birdsong, and G. P. Garrett, editors. Multispecies and watershed approaches to freshwater fish conservation. American Fisheries Society, Bethesda, Maryland.

Prepared by:

Richard N. Williams, Federation of Fly Fishers, 524 W Two Rivers Driver, Eagle, ID 83616; Email: troutdna@cableone.net

Daniel C. Dauwalter, Trout Unlimited, 910 Main Street, Suite 342, Boise, ID 83702; Email: ddauwalter@tu.org

Contents:

Supplement A. Results from a survey of fisheries professionals with expertise in the Upper Snake River basin.

We surveyed fishery professionals familiar with the resources in the Upper Snake River basin to assess the relative importance of various factors when identifying watersheds with potential as native fish conservation areas. Questions were designed to assess: the importance of different native trout species (and forms); importance of different non-game fish species; relative importance of current distribution versus historical distribution; relative importance of landscape indicators of habitat integrity (riparian condition, watershed connectivity, watershed condition, water quality, flow regime) and future habitat security (land conversion, energy development, resource extraction, climate change, introduced species); relative importance of land ownership; relative importance of land protection status; acceptable levels of genetic introgression; and minimum watershed sizes for NFCA function. The survey was sent by email (with one emailed follow up reminder) to 42 fishery professionals from Idaho Department of Fish and Game, Wyoming Game and Fish Department, Oregon Department of Fish and Wildlife, Idaho Department of Environmental Quality, Idaho Office of Species Conservation, U.S. Bureau of Land Management, U.S. Forest Service, U.S. Geological Survey, U.S. Fish and Wildlife Service, Henrys Fork Foundation, Friends of the Teton River, Trout Unlimited, College of Idaho, Idaho State University, Boise State University, industry consultants, and Idaho Power. The survey was implemented in Google Documents and all responses were anonymous. Twenty people responded and mean responses are shown in Figure A1.

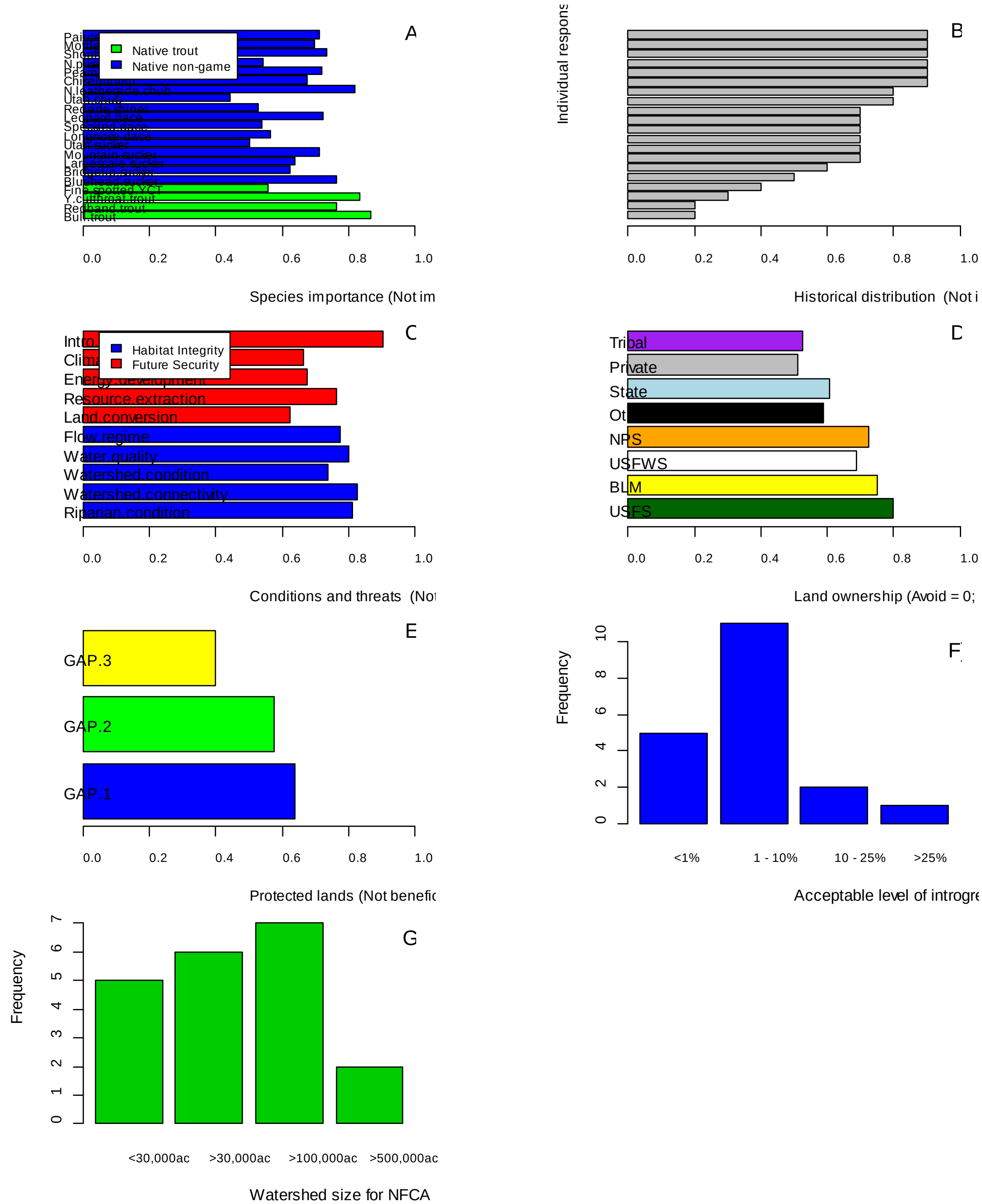

Figure A1. Mean survey response or frequency of survey responses to questions regarding the relative importance of different factors when identifying native fish conservation areas in the Upper Snake River basin: different native trout and non-game species (A); historical species distributions relative to current distributions (B); indicators of habitat integrity and future threats (C); land ownership (D); land protection status (E); levels of genetic introgression (F); and watershed size (G). Number of survey responses was 20.

Supplement B. Probability of occurrence models for native non-game species in the Upper Snake River basin.

Comprehensive assessment of the value of places for species conservation is not only dependent on known species localities, but also where they have the potential to occur based on habitat suitability when field data are sparse. Modeling the distribution of species from a suite of habitat characteristics available for the entire landscape of interest can be used to predict how likely a species is to be found there during sampling given existing habitat conditions. Species distribution modeling, where the presence-absence of species can be predicted from a suite of habitat variables, represents an efficient way to predict where species are likely to occur across entire landscapes when the entire landscape cannot be sampled with field sampling techniques. Species distribution modeling has been successfully applied to stream fishes in riverine networks (Dauwalter and Rahel 2008), and have been used to identify watersheds with a high potential for multiple species conservation in the Colorado River Basin (Dauwalter et al. 2011; Strecker et al. 2011). Species distribution models were developed for each species in the Upper Snake River basin as an important data component in identification of potential native fish conservation areas.

Methods

Presence-absence data for native non-game fishes were used to develop models that predict the probability of occurrence of each species for each stream segment in the Upper Snake River basin. Specifically, random forest models (Breiman 2001) were developed that predict probability of occurrence as a function of: mean annual streamflow (cfs), stream slope (%), mean annual precipitation (mm), and canal density in watershed (km / km^2) from the National Hydrography Dataset (NHD) Plus (USEPA and USGS 2005); mean August air temperature (C) (PRISM Group 2008); % converted land (urban, agriculture, hay/pasture) in watershed and % converted land in a 100-m stream buffer within watershed (USEPA 2001); total reservoir storage per watershed size (m^3 / km^2)(USACE 2008); and road density in watershed (km / km^2)(U.S. Census 2001) (Figure B1); random forests are a flexible modeling approach that can model high order interactions and non-linear relationships between environmental variables and species occurrence without a priori model specification (i.e., predefined variable interactions or non-linear responses between predictor and response variables) that is needed in traditional modeling approaches such as generalized linear models (Olden et al. 2008; Evans et al. 2011). Models were fit using only data from sub-basins representing probable native species distributions as reported in Meyer et al. (2013), as well as Gamett (2003) for the sinks drainages. The approach of Murphy et al. (2010) was used to screen variables based on their contribution to explaining species occurrences; final models were refit with the remaining variables that best explained species occurrences. Final species-specific models were then used to predict probability of occurrence for all species for perennial stream segments with drainage areas greater than 159,100 km^2 in the NHDPlus; again, predictions were only made in sub-basins within the probable native range reported by Meyer et al. (2013) and Gamett (2003). Model predictive ability was assessed by using the Area Under the Curve (AUC) of a plot of a receiver operating characteristic plot (i.e., a plot of 1-specificity vs. sensitivity across a range of probability values), which is a measure of model performance ranging for 0.5 (no discrimination ability) to 1.0 (complete discrimination) that is unaffected by species prevalence (proportion of sites where species was observed) (Hosmer and Lemeshow 2000; Manel et al. 2001). Both in-sample and 10-fold cross-validated AUC values were computed.

Results

Random forest models developed for all non-game fish species, except leopard dace, showed good predictive ability with 10-fold cross-validated AUC values >0.75 (Table B1); the model for leopard dace had the poorest predictive ability (AUC =0.695). Mean annual flow (a measure of stream size), stream slope, August air temperature, and mean annual precipitation influenced species presence most often (Table B2; Figure B2), but contribution of each variable differed by species (Table B2) and relationships between species probability of occurrence and each environmental variable was often non-linear (Figure B3). Probability of occurrence predictions were typically higher in areas where the species was present in the field survey data, again indicating models with good predictive ability (Figure B4).

References

Breiman, L. 2001. Random forests. Machine Learning 45:5-32.

Dauwalter, D. C., and F. J. Rahel. 2008. Distribution modeling to guide stream fish conservation: an example using the mountain sucker in the Black Hills National Forest, USA. Aquatic Conservation: Marine and Freshwater Ecosystems 18:1263-1276.

Dauwalter, D. C., J. S. Sanderson, J. E. Williams, and J. R. Sedell. 2011. Identification and implementation of Native Fish Conservation Areas in the Upper Colorado River Basin. Fisheries 36:278-288.

Evans, J. S., M. A. Murphy, Z. A. Holden, and S. A. Cushman. 2011. Modeling species distribution and change using random forest. Pages 139-159 *in* C. A. Drew, Y. F. Wiersma, and F. Huettmann, editors. Predictive species and habitat modeling in landscape ecology. Springer, New York.

Gamett, B. L., and D. W. Zaroban. 2003. The distribution and potential origin of sculpin species in the sinks drainages of southeastern Idaho. Pages 10-12 *in* R. W. Van Kirk, J. M. Capurso, and B. L. Gamett, editors. The Sinks symposium: exploring the origin and management of fishes in the Sinks Drainages of southeastern Idaho. Idaho Chapter American Fisheries Society, Boise.

Gamett, B. L. 2003. A summary: the origin of fishes in the Sinks Drainages of southeastern Idaho. Page 16 *in* R. W. Van Kirk, J. M. Capurso, and B. L. Gamett, editors. The Sinks symposium: exploring the origin and management of fishes in the Sinks Drainages of southeastern Idaho. Idaho Chapter American Fisheries Society, Boise.

Hosmer, D. W., and S. Lemeshow. 2000. Applied logistic regression, second edition. John Wiley & Sons, Inc., New York.

Manel, S., H. C. Williams, and S. J. Ormerod. 2001. Evaluating presence-absence models in ecology: the need to account for prevalence. Journal of Applied Ecology 38:921-931.

Meyer, K. A., J. A. Jr. Lamansky, D. J. Schill, and D. W. Zaroban. 2013. Nongame fish species distribution and habitat associations in the Snake River basin of southern Idaho. Western North American Naturalist 73:20-34.

Murphy, M. A., J. S. Evans, and A. Storfer. 2010. Quantifying *Bufo boreas* connectivity in Yellowstone National Park with landscape genetics. Ecology 91:252-261.

PRISM Group (Oregon State University). 2008. PRISM 800m Normals (1971 - 2000). Oregon State University, Corvallis, Oregon. http://www.prism.oregonstate.edu/.

Strecker, A. H., J. D. Olden, J. B. Whittier, and C. P. Paukert. 2011. Defining conservation priorities for freshwater fishes according to taxonomic, functional, and phylogenetic diversity. Ecological Applications 21:3002-3013.

U.S.Census Bureau. 2001. Census 2000. U.S. Census Bureau, Population Division, U.S. Deparment of Commerce, Washinton, D. C.

USACE (U.S. Army Corps of Engineers). 2008. National Inventory of Dams. U.S. Army Corps of Engineers. http://crunch.tec.army.mil/nidpublic/webpages/nid.cfm.

USEPA. 2001. National Landcover Dataset 2001. http://www.epa.gov/mrlc/nlcd-2001.html.

USEPA and USGS (U.S. Environmental Protection Agency and U.S. Geological Survey). 2005. National Hydrography Dataset Plus - NHDPlus (1:100,000 scale). U.S. Environmental Protection Agency and U.S. Geological Survey, Sioux Falls, South Dakota. http://www.horizon-systems.com/nhdplus/ .

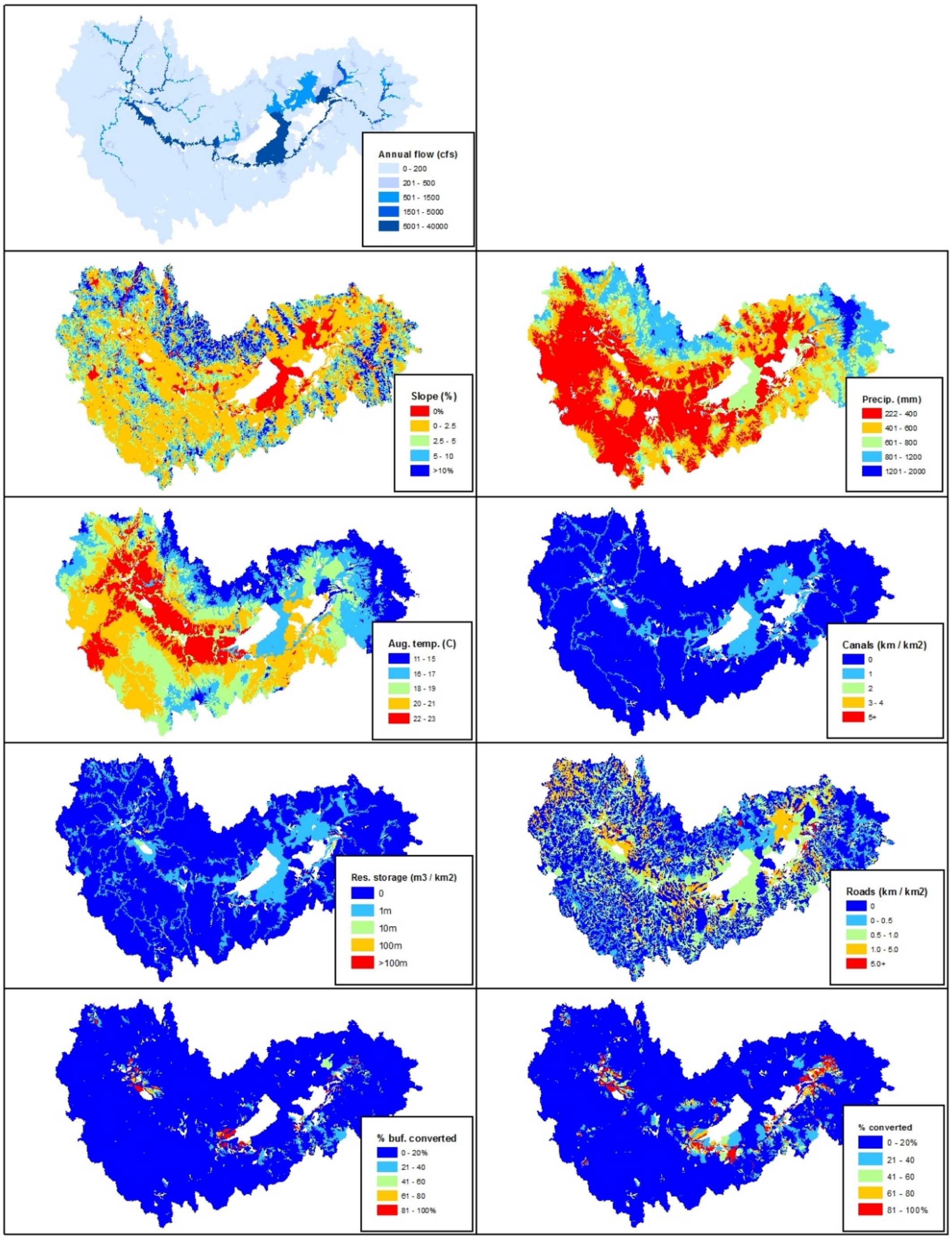

Figure B1. Environmental variables used to develop models used to predict probability of occurrence for native non-game species in the Upper Snake River basin.

Table B1. Number of sites observed, total sites, prevalence (observed / total) for model training data, and in-sample Area-Under-the-Curve (AUC) of a Receiver-Operating-Characteristic (ROC) plot and 10-fold cross-validated AUC of species niche models.

Species	Observed	Total sites	prevalence	In-sample AUC	10-fold AUC
N. leatherside chub	39	1210	0.032	0.999	0.925
Bluehead sucker	46	1296	0.035	0.998	0.915
Bridgelip sucker	290	1609	0.180	0.992	0.885
Largescale sucker	101	1609	0.063	0.998	0.899
Mountain sucker	118	2736	0.043	0.992	0.834
Utah sucker	55	1310	0.042	0.999	0.875
Chiselmouth	88	1452	0.061	0.999	0.931
Longnose dace	298	3002	0.099	0.996	0.886
Speckled dace	640	3047	0.210	0.994	0.882
Leopard dace	10	660	0.015	0.999	0.695
Redside shiner	462	3045	0.152	0.986	0.883
N. pikeminnow	123	1448	0.085	0.997	0.907
Utah chub	23	1280	0.018	0.998	0.829
Mottled sculpin	263	2628	0.100	0.995	0.846
Paiute sculpin	237	2456	0.096	0.998	0.827
Shorthead sculpin	239	1609	0.149	0.995	0.923
Wood River sculpin	44	154	0.286	0.996	0.768
Mountain whitefish	100	1354	0.074	0.997	0.851

Table B2. Mean decrease in accuracy of nine environmental variables for random forest models by species indicating relative importance of each variable in predicting species occurrence. Blank values indicate low predictive ability of that variable for that species and was excluded in final models.

Species	Annual flow (cfs)	Annual Precip. (mm)	August Temp. (C)	Roads (km / km^2)	% watershed converted	Dam storage (m^3/km^2)	Slope (%)	Canals (km / km^2)	% 100-m buffer converted
N. leatherside chub	14.15	15.98	14.81	11.81	12.10				
Bluehead sucker	27.21	19.30	18.63			15.81			
Bridgelip sucker	78.46	67.46	52.19	47.79	43.54	16.84	78.63	11.86	33.30
Largescale sucker	74.68						39.33	34.30	27.02
Mountain sucker	47.01	43.96	40.48		23.45		39.70		
Utah sucker	13.23	14.78	12.70	9.13	8.40	10.54	15.56	1.59	9.54
Chiselmouth	42.01	29.50			23.27			34.16	26.09
Longnose dace	157.79	101.28	125.73				116.17		
Speckled dace	231.91	212.02	219.51				278.76		
Leopard dace		3.05	3.48		3.93			2.23	3.96
Redside shiner	151.36	129.38	141.16	119.48			181.19		
N. pikeminnow	63.57		35.60			30.81	47.87		29.57
Utah chub	10.43		6.72	7.82			13.48		
Mottled sculpin	71.36	72.54	68.14	58.45	40.45	13.16	75.69	11.19	35.16
Paiute sculpin		100.37	93.52	74.93	42.18		93.64		
Shorthead sculpin	85.68	77.62	97.05	51.84			13.75		
Wood River sculpin	12.79	13.45	16.02				12.94		3.61
Mountain whitefish	91.32	80.02							

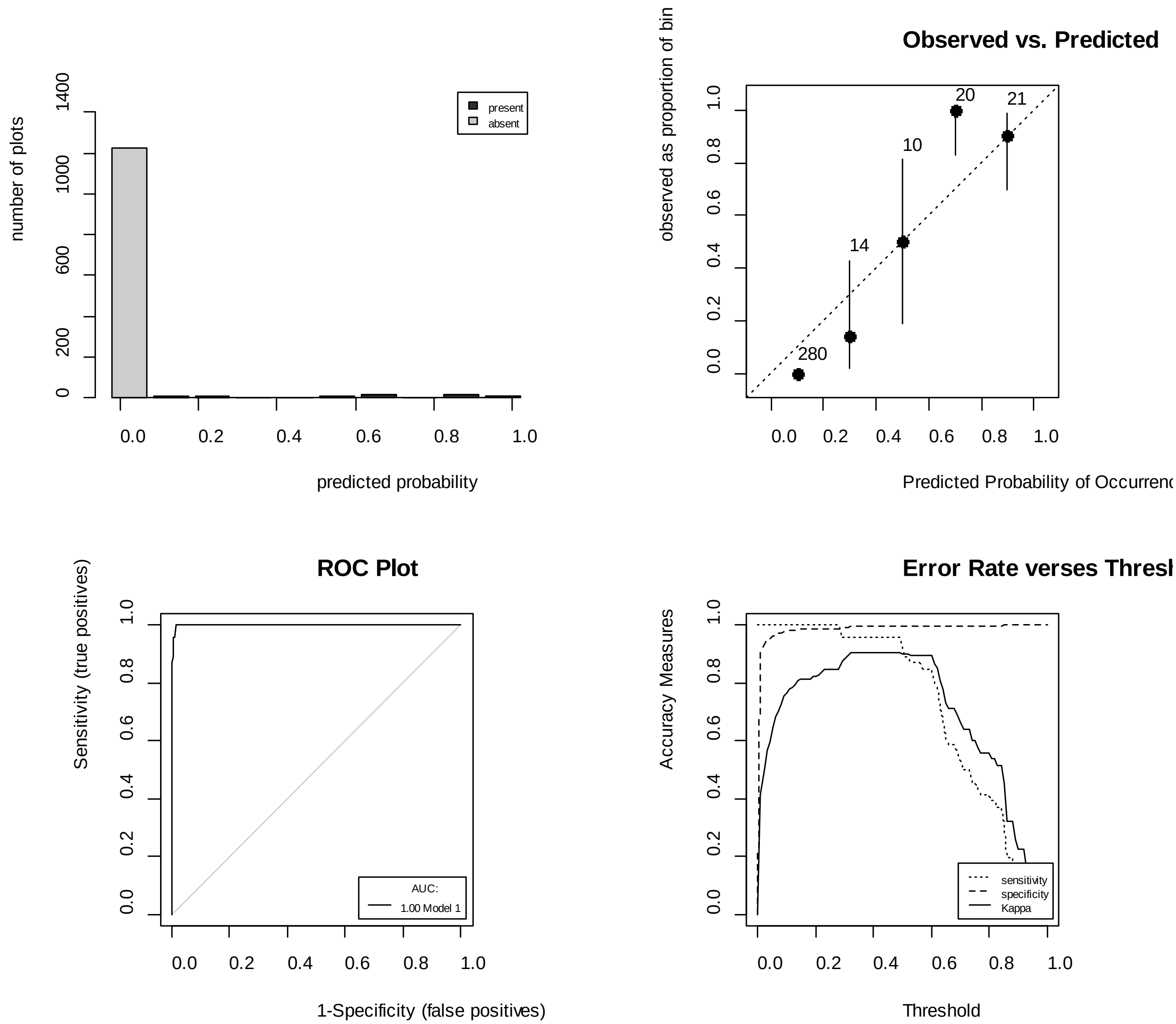

Figure B2. Bluehead sucker random forest model diagnostics. Frequency of bluehead sucker presences and absences versus predicted probability of occurrence (upper left); proportion of sites with presences versus predicted probability of occurrence bins (upper right); Receiver operating characteristic (ROC) plot of model sensitivity versus 1 – specificity (lower left); sensitivity, specificity, and Kappa statistic versus probability threshold used to assign species presence or absence based on model prediction (lower right). * change y axis on upper left so the low bars are visible**

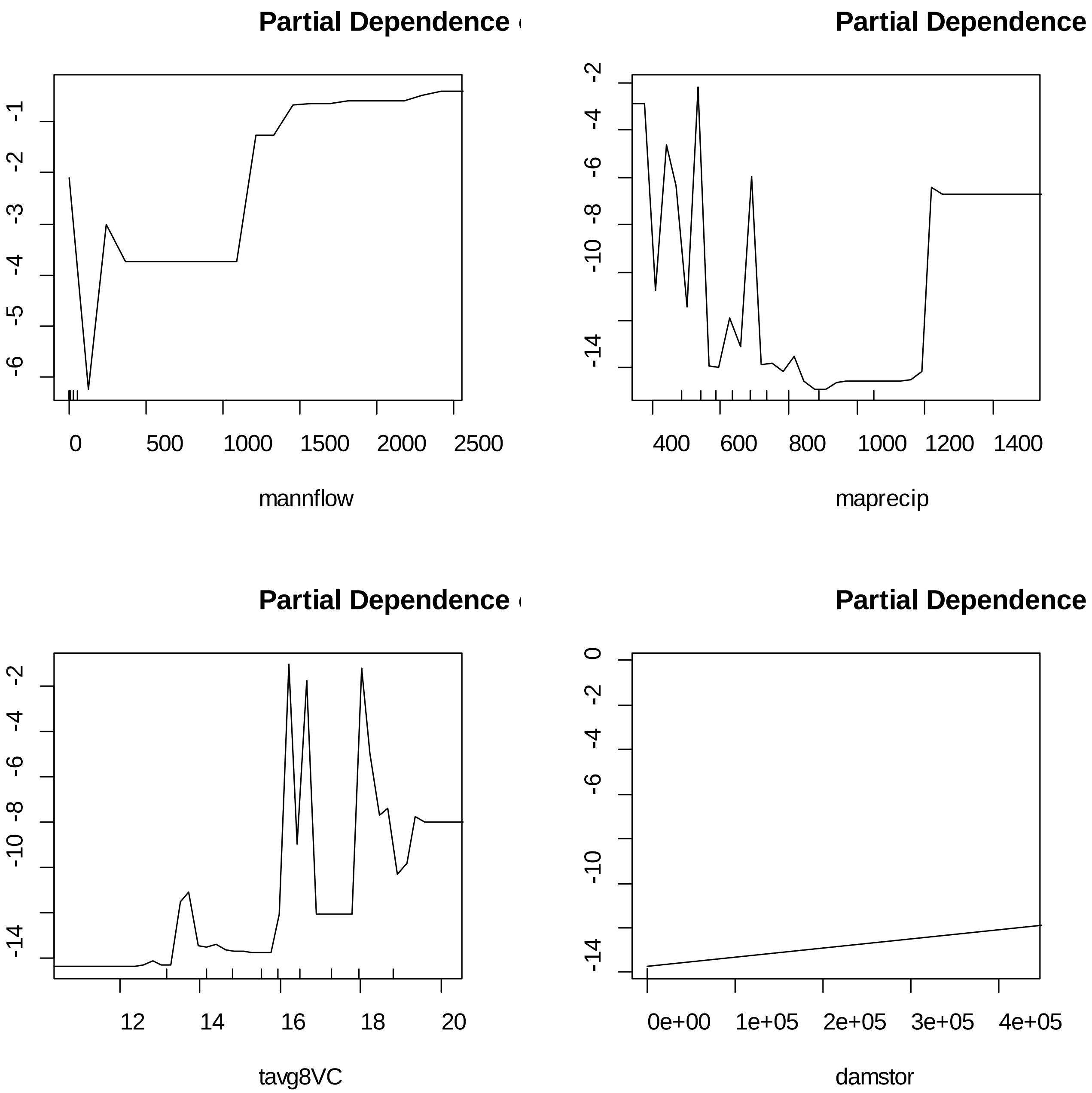

Figure B3. Partial dependence plots from a random forest model showing how bluehead sucker probability of occurrence varies with individual environmental variables at the mean values of all other variables. Plots ordered left to right in order of variable importance. Mannflow = mean annual flow (cfs); tavg8VC = mean August temperature (C); cpconv = % converted land in watershed; maprecip = mean annual precipitation (mm); slope = stream segment slope (%); damstor = total dam storage above segment (m^3 / km^2); roadden = road density in watershed (km / km^2); buf_pconv = % converted land in a 100-m stream buffer; canalssqkm = canal density in watershed (km / km^2).

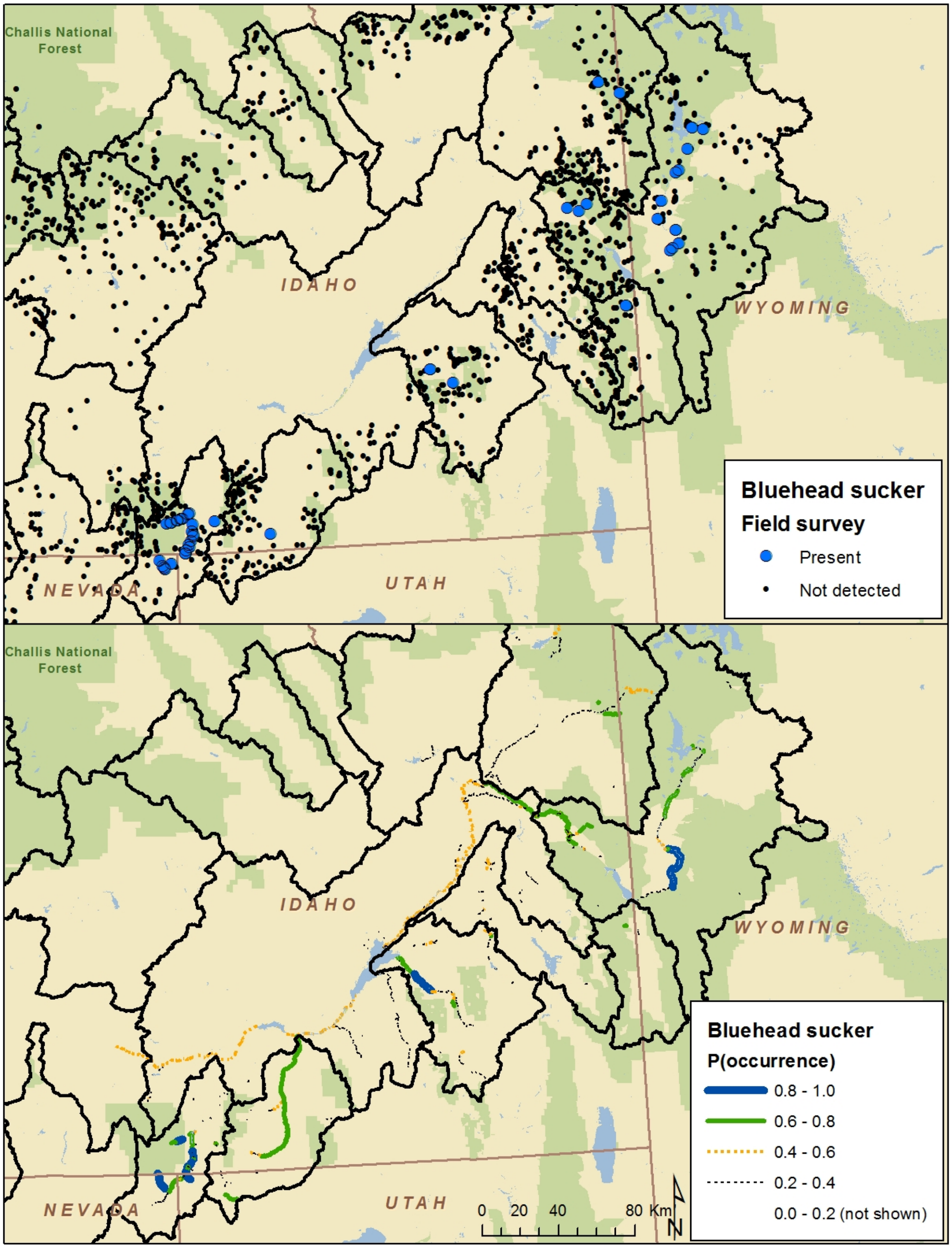

Figure B4. Probability of occurrence predictions from a random forest model for bluehead sucker in the Upper Snake River basin above Shoshone Falls.

Supplement C. Application of Habitat Integrity and Future Security indicators from Trout Unlimited's Conservation Success Index (www.tu.org/csi) in the Upper Snake River basin.

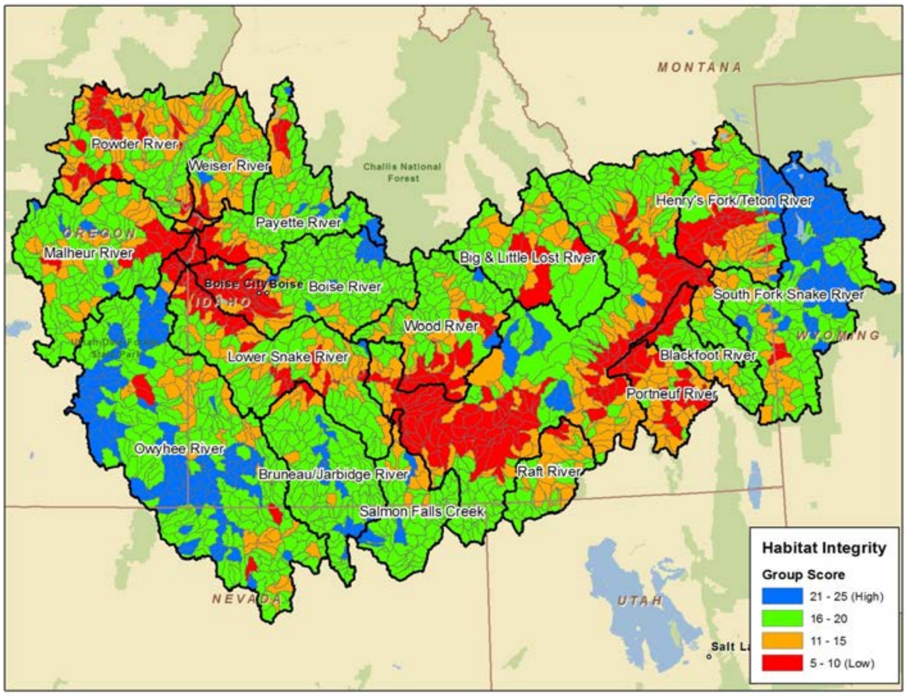

Figure C1. Spatial distribution of Habitat Integrity scores for subwatersheds (HUC 12: n = 2079) in the Upper Snake River basin above Hells Canyon. Habitat Integrity indicators were based on Trout Unlimited's Conservation Success Index.

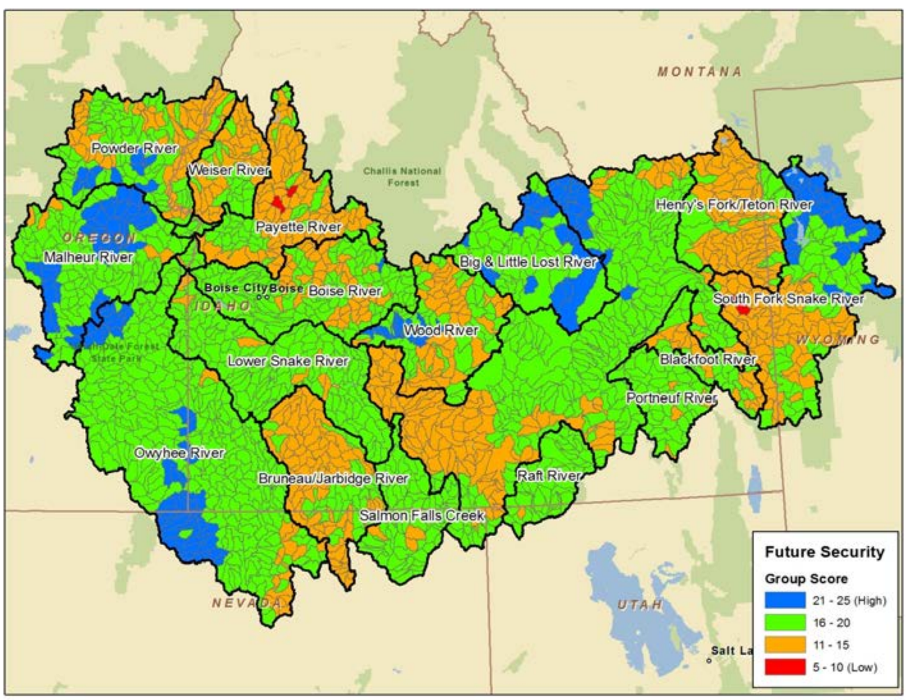

Figure C2. Spatial distribution of Future Security scores for subwatersheds (HUC 12: n = 2079) in the Upper Snake River basin above Hells Canyon. Future Security indicators were based on Trout Unlimited's Conservation Success Index.

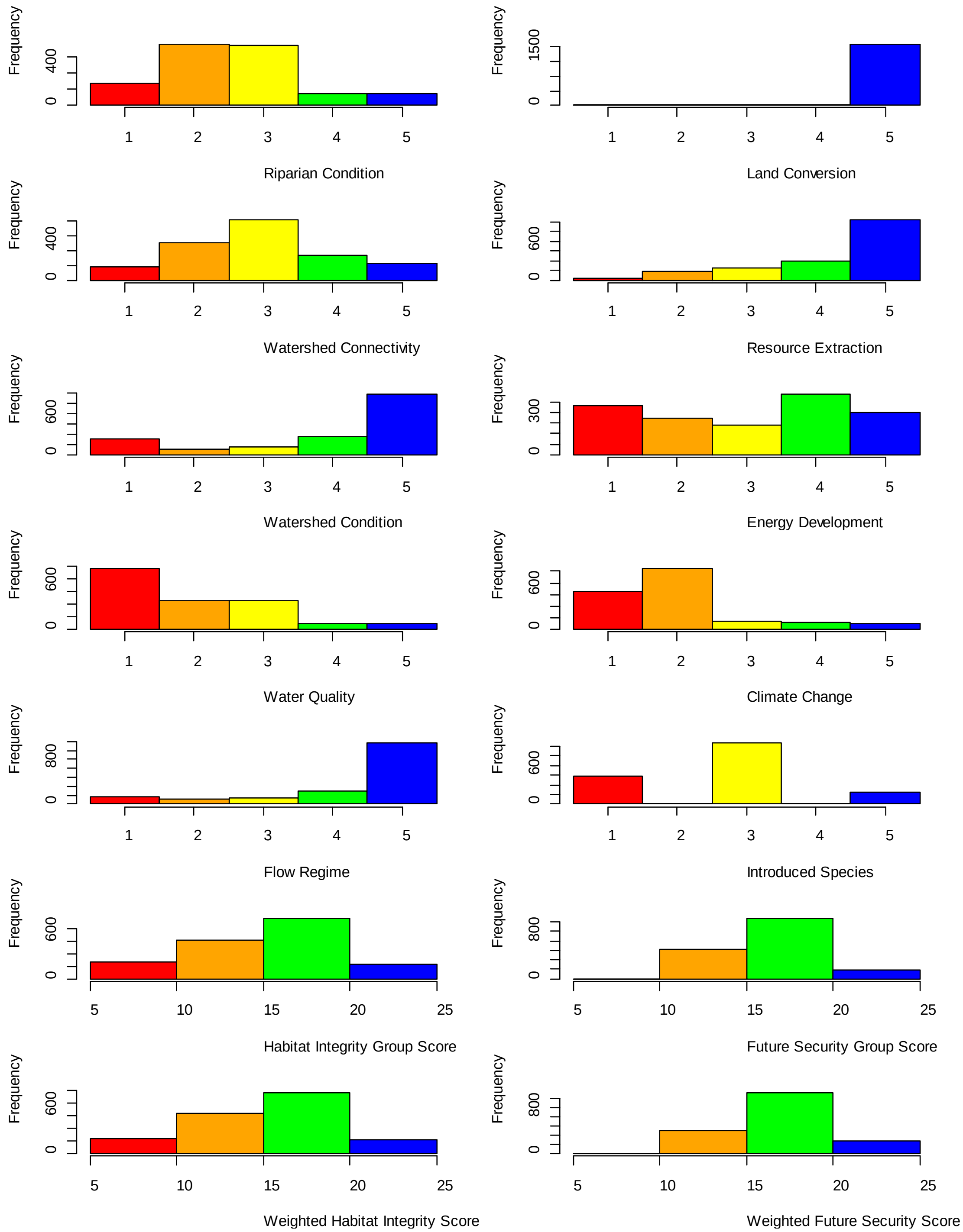

Figure C3. Frequency distribution of Habitat Integrity (Left) and Future Security (Right) indicator scores, group scores, and weighted group scores for subwatersheds (HUC 12; n = 2079) in the Upper Snake River basin above Hells Canyon. Habitat Integrity and Future Security indicators were based on Trout Unlimited's Conservation Success Index.

Supplement D. Land ownership in the Upper Snake River basin.