A Statistical Approach for Watershed Pairing to Inform Semi-Arid Watershed Restoration and Develop a Regional Ecosystem Services Program

Roy E. Petrakis[1], Kurt Vaughn[2], Laura M. Norman[3], Caleb Weaver[2], H. Ron Pulliam[2], and Richard Pritzlaff[4]

Borderlands Restoration Network, Tucson, AZ[1]; Borderlands Restoration Network, Patagonia, AZ[2]; U.S. Geological Survey, Tucson, AZ[3]; Biophilia Foundation, Chester, MD[4]

Abstract--*Watersheds within the Madrean Archipelago have been severely degraded by a combination of both anthropogenic and natural processes, including unsustainable grazing and timber harvests, long-term droughts, and acute flooding events. We are investigating the effectiveness of various holistic watershed restoration techniques, used by land managers to combat the negative effects of these processes. Our study site, Smith Canyon, located in the Sonoita Creek Watershed and part of the Nogales Ranger District of the Coronado National Forest in Arizona, U.S.A., consists of numerous (~90) structurally-similar sub-basins, each covering approximately 2 to 5 ha. This unique and structurally repetitive landscape presents an opportunity for rigorous large-scale experimentation when considering each sub-basin as a replicate unit. We are using a statistical approach to group subsets of the sub-basins based on structural and biophysical traits. Restoration treatments within the selected sub-basins will focus on the use of check dams and other erosion control methods designed to reduce erosion impacts and reduce nonpoint source pollution downstream. We will monitor the empirical results of water and sediment flows following the treatments to assess the effectiveness and impacts on ecosystem services. We propose these research results be incorporated into a new protocol for large-scale restoration, with the overarching goal to document ecosystem services and a potential payment mechanism for the region.*

INTRODUCTION

A combination of anthropogenic and natural processes, including unsustainable grazing, long-term droughts, and acute flooding events, have degraded the quality of semi-arid and arid watersheds within the Madrean Archipelago and across the Southwest over the past several decades (BLM and USFS, 1994; Denevan, 1967). Watershed restoration structures have been used to combat the negative effects of these processes. Benefits resulting from restoration structures include, but are not limited to: sediment control (Polyakov et al. 2014), aquifer regeneration (Norman et al., 2018), water flow dynamics (Baker et al. 1995; Norman et al. 2010a,b, 2015; Norman and Niraula, 2016), and improving habitat and vegetation conditions (Debano and Schmidt, 1990; Norman et al. 2014; Wilson and Norman, 2018). Much of the knowledge about management effects on the hydrologic cycle and vegetation comes from paired catchment studies, though often it is difficult to match watersheds with good controls (Neary, 2016).

Our overarching goal is to investigate the effectiveness of holistic watershed restoration techniques and the impacts they may have on a suite of ecosystem services, with the aim of possibly developing viable payments, though yet to be determined, for an ecosystem services program in the region. We plan to address these goals through the following research stages. First, we developed a spatial database consisting of structural and biological data to allow for a robust and statistically valid experimental design. Using statistical algorithms with this dataset, we paired groups of sub-basins that are most similar, leaving groups of treatment and control basins. Within the theme of watershed pairing, sub-basins are considered as separate watersheds. Because of funding and time constraints, a portion of these grouped sub-basins will be selected for monitoring based on practitioner input. The amount of sediment and water flow will be monitored prior to the treatments for our selected study sub-basins during the 2018 monsoon season (July – September). In the summer of 2019, erosion-control and water retention restoration structures will be constructed and additional monitoring of sediment and water flows will take place. Finally, based on the quantified impact of these structures on general watershed conditions, we will develop a protocol to improve future restoration efforts and a develop an ecosystem services program. This paper describes this unique methodological approach for statistically paired watersheds and future experimental design.

STUDY AREA

Our study area is Smith Canyon Watershed (SCW), which is located within the Nogales Ranger District of the Coronado National Forest, north of Patagonia, Arizona, U.S.A. (Figure 1). This watershed provides a unique opportunity for rigorous large-scale experimentation due to its feather-like landscape consisting of roughly 90 structurally similar sub-basins (SUB(s)), each approximately 2 to 5 ha. Smith Canyon is a tributary of Sonoita Creek, within the Sonoita Creek 5th Hydrologic Unit Code (HUC). Water flows within SCW are dependent on rainfall events, particularly during the summer monsoon season (Adams and Comrie, 1997).

DATA AND METHODS

Spatial Database Development

We collected a combination of biophysical and structural variables for Smith Canyon Watershed and the surrounding Sonoita Creek Watershed. A sub-3 m Digital Elevation Model (DEM) was used to derived from airborne LiDAR to develop elevation statistics, including: mean, minimum, maximum, and range (Tyson Swetnam, Tucson, AZ, personal communication, January 25, 2018). We also included the following structural input variables: 1) aspect (percent north, south, east, and west facing), and 2) slope (percent gradual, moderate, and steep), derived from the Soil and Water Assessment Tool (SWAT – SWAT Landuse, 2018). Additionally, we collected general structural traits such as area, length, width, and perimeter for each SUB.

To comprise the set of biophysical variables, we included a combination of various land use/land cover classifications for the region, each intended to map certain aspects of the landscape, including: land cover from 2009 (Villarreal et al. 2011), the translated landuse derivatives for Villarreal et al. (2011) created in SWAT (SWAT Landuse, 2018), and existing vegetation designations (Wallace et al. 2011). We also included the Soil Survey Geographic (SSURGO) soil database for southeast Arizona developed by the USDA (Soil Survey Staff). Finally, we produced a five-year mean (2013 – 2017) Normalized Difference Vegetation Index (NDVI – Tucker, 1979) image for the pre-monsoon period of stressed vegetation (mid-June) in Google Earth Engine (GEE – Gorelick et al. 2017). The images GEE used to develop this include Landsat 8 Operational Land Imager (OLI) images from overlapping scenes (Path 35/36; Row 38). Using the resulting NDVI dataset, we quantified various statistics including: minimum, maximum, mean, range, and standard deviation, for each SUB.

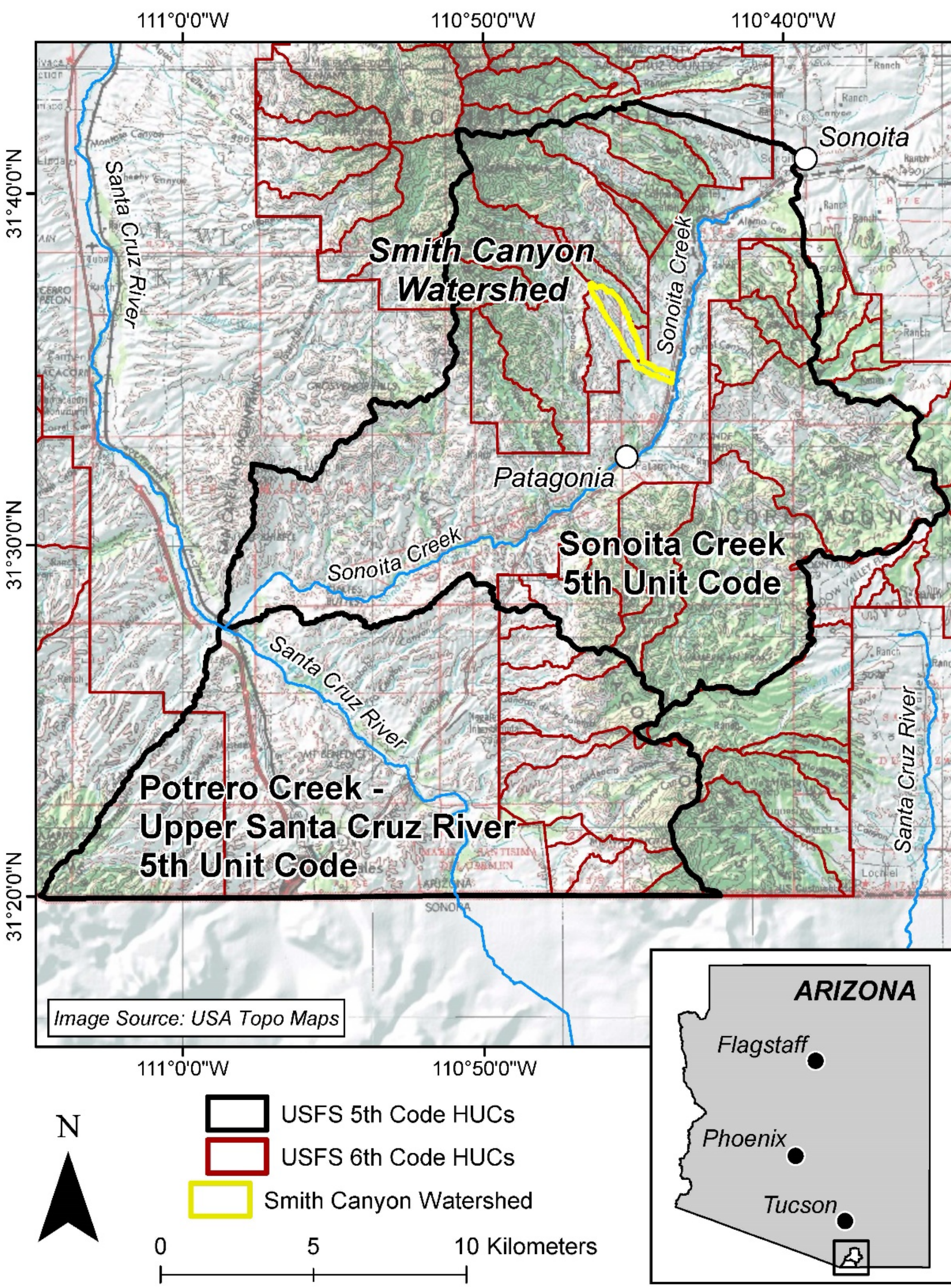

Figure 1--Location of the Smith Canyon Watershed (SCW – in yellow) in relation to the Sonoita Creek 5th Unit Code, approximately 5 km north of Patagonia, AZ.

Clustering Analysis

To pair various SUBs that were most similar based on the input biophysical and structural variables, we used hierarchical clustering techniques in R (Murtagh and Contreras, 2011; R Core Team, 2015). Several processing steps were completed for this analysis. First, we separated both the structural and biophysical variable types quantified for each SUB. For each variable type, we then completed a factor extraction and reduction process using a combination of variable correlations, Maximum Likelihood, and Principal Component Analysis (PCA). We completed this step for each variable type separately to see how the variables interacted. We also shaped our variable selection by considering the real-world impact each variable may have in predicting dynamics of water and sediment flows. For example, gradual and steep slopes were chosen because they will cause increased variability of water flows, whereas moderate slopes will have comparatively less impact on flow dynamics. We then combined the most effective variables for each type and completed the variable reduction analysis with the full suite of variables. Once variable selection was complete, we ran several tests within R to quantify the clustering tendency of the SUBs based on variables included, as well as the optimal number of clusters present for the SUBs. Finally, we built a dendrogram to visualize the SUB clusters. Final selection of study SUBs will use a combination of the dendrogram results and a visual review by local land managers, where tendencies of water flow within the SUBs as well as general characteristics of erosion and flows will be considered.

Restoration Structures

We will develop a protocol for the implementation of restoration structures, ranging from cross-channel check dams to larger-scale head-cut armaments, depending on erosion severity (Heede, 1978). Many of the SUBs have developed large head-cuts. Restoration structures will be designed for erosion-control and water retention – varying by location, size, and construction material. We plan to construct structures using a combination of local dead vegetation and available rocks. Within the defined clusters, 12 SUBs were selected to create three replicates of the four restoration treatment approaches: 1) control (no structures), 2) low (head-cut armament), 3) moderate (head-cut armament; check dams), and 4) high density (head-cut armament; check dams; upland structures).

RESULTS AND DISCUSSION

Factor Extraction, Reduction, and Final Selection

We determined through our statistical analysis approach that a reduced combination of the initial input biophysical and structural variables was optimal for our clustering analysis (Table 1). Each of the final variables can have a unique influence on sediment and water flows within the SUBs. Additionally, the selected variables group SUBs that are most similar within Smith Canyon.

Dendrogram

Our dendrogram consisted of three main clades, or clusters, each with multiple sub-hierarchies (Figure 2). Each of the clades was characterized by unique qualities present within their respective SUBs.

CONCLUSION

We used a statistic-based clustering analysis to group similar sub-basins centered on their biophysical and structural traits, for use as paired control and restoration watersheds. This robust experimental design can improve the validity of statistically valid restoration science. It will form the basis for an analysis of the impacts and effectiveness of restoration structures. This research is intended to stand as an initial step in developing a viable payment program for ecosystem services in the region, including services such as downstream flood insurance rates and carbon sequestration payments.

Improving watershed conditions provides greater flood control and better water quality with indirect benefits to surrounding habitats, and monetization of these ecosystem services could even provide funding for future projects.

Table 1--The final combination of biophysical and structural variables used for the clustering analysis.

Variable	Variable Data Type	Variable Type
Percent steep slope	Slope	Structural
Mean slope	Slope	Structural
Percent North facing	Aspect	Structural
Percent East Facing	Aspect	Structural
Length	Trait	Structural
Area	Trait	Structural
Percent Range Shrubland (RNBG)	SWAT Landuse	Biophysical
Percent moderate slope gravelly sandy loam (1421630)	Soil	Biophysical
Percent steep slope gravelly sandy loam (1421631)	Soil	Biophysical
NDVI Maximum	Vegetation Index	Biophysical
NDVI Minimum	Vegetation Index	Biophysical

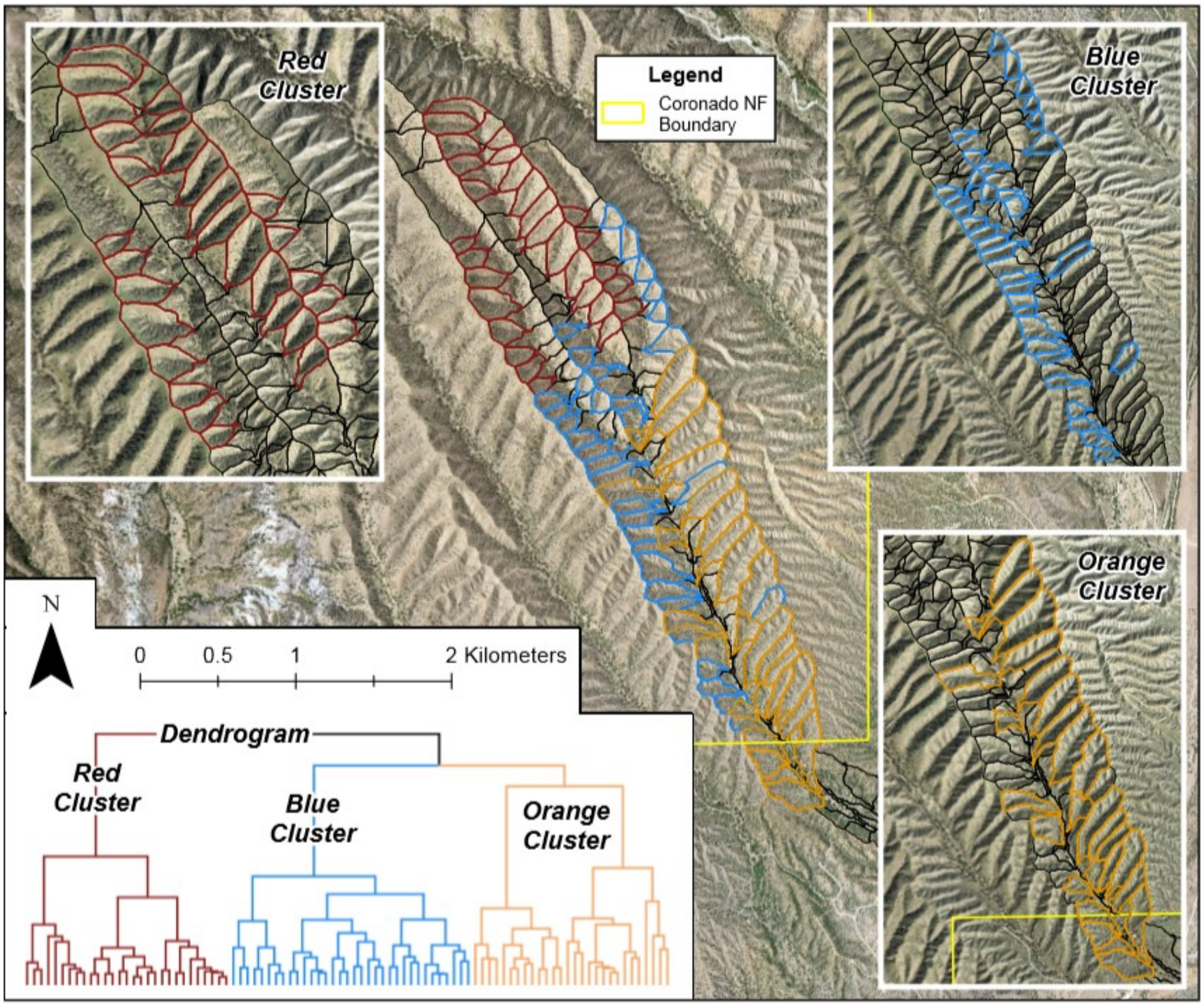

Figure 2--Dendrogram results and the location of clusters within Smith Canyon Watershed.

REFERENCES

Adams, D.K.; Comrie, A.C. 1997. The North American Monsoon. Bulletin of the American Meteorological Society. 78: 2197–2213.

Baker Jr., M.B.; DeBano, L.F.; Ffolliott, P.F. 1995. Hydrology and watershed management in the Madrean Archipelago. In: DeBano, L.H.; Ffolliott, P.H.; Ortega-Rubio, A.; Gottfried, G.J.; Hamre, R.H.; Edminster, C.B.; tech. coords., Biodiversity and Management of the Madrean Archipelago: The Sky Islands of Southwestern United States and Northwestern Mexico, Tucson, AZ, September 19–23, 1994. USDA Forest Service, Rocky Mountain Forest and Range Experiment Station, Fort Collins, CO, pp. 329–337.

BLM (U.S. Bureau of Land Management); USFS (U.S. Forest Service). 1994. Rangeland reform '94 : draft environmental impact statement : executive summary. Washington, DC. U.S. Department of the Interior. Bureau of Land Management.

Debano, L.F.; Schmidt, L.J. 1990. Potential for enhancing riparian habitats in the southwestern United States with watershed practices. Forest Ecology and Management. 33-34: 385–403.

Denevan, W.M. 1967. Livestock Numbers in Nineteenth-Century New Mexico, and the Problem of Gullying in the Southwest. Annals of the Association of American Geographers. 57: 691–703.
Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sensing of Environment. 202: 18-27.
Heede, B.H. 1978. Designing gully control systems for eroding watersheds. Environmental Management. 2: 509–522.
Murtagh, F.; Contreras, P. 2011. Algorithms for hierarchical clustering: an overview. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. 2: 86–97.
Neary, D.G. 2016. Long-Term Forest Paired Catchment Studies: What Do They Tell Us That Landscape-Level Monitoring Does Not? Forests. 7: 1-15.
Norman, L.M.; Niraula, R. 2016. Model analysis of check dam impacts on long-term sediment and water budgets in Southeast Arizona, USA. Ecohydrology & Hydrobiology. 16: 125–137.
Norman, L.M.; Huth, H.; Levick, L.; Burns, I.S.; Guertin, D.P.; Lara-Valencia, F.; Semmens, D. 2010a. Flood hazard awareness and hydrologic modelling at Ambos Nogales, United States–Mexico border. Journal of Flood Risk Management. 3: 151–165.
Norman, L.M.; Levick, L.; Guertin, D.P.; Callegary, J.; Guadarrama, J.Q.; Anaya, C.Z.G.; Prichard, A.; Gray, F.; Castellanos, E.; Tepezano, E.; Huth, H.; Vandervoet, P.; Rodriguez, S.; Nunez, J.; Atwood, D.; Granillo, G.P.O.; Ceballos, F.O.G. 2010b. Nogales flood detention study (Open-File Report No. 2010-1262). U.S. Geological Survey.
Norman, L.M.; Villarreal, M.L.; Pulliam, H.R.; Minckley, R.; Gass, L.; Tolle, C.; Coe, M. 2014. Remote sensing analysis of riparian vegetation response to desert marsh restoration in the Mexican Highlands. Ecological Engineering. 70: 241–254.
Norman, L.M.; Brinkerhoff, F.; Gwilliam, E.; Guertin, D.P.; Callegary, J.; Goodrich, D.C.; Nagler, P.L.; Gray, F. 2015. Hydrologic response of streams restored with check dams in the Chiricahua Mountains, Arizona. River Research and Applications. 32: 519–527.
Norman, L.M.; Callegary, J.B.; Lacher, L.J.; Wilson, N.R.; Fandel, C.A.; Forbes, B.; Swetnam, T. 2018, In Press. Impacts of restoration on groundwater recharge at the Babocomari Ranch, SE Arizona. Journal of Hydrology.
Polyakov, V.O.; Nichols, M.H.; McClaran, M.P.; Nearing, M.A. 2014. Effect of check dams on runoff, sediment yield, and retention on small semiarid watersheds. Journal of Soil and Water Conservation. 69: 414–421.
R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. [online]. https://www.R-project. Org/
Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. Web Soil Survey. [online]. http://websoilsurvey.nrcs.usda.gov/
SWAT Landuse. 2018. [online]. https://oldgeni.isnew.info/landuse.html#_SWAT_landuse_classification [June 12, 2018].
Tucker, C.J. 1979. Red and Photographic Infrared Linear Combinations for Monitoring Vegetation. Remote Sensing of Environment. 8: 127–150.
Villarreal, M.L.; Norman, L.M.; Wallace, C.S.A.; Van Riper III, C. 2011. A Multitemporal (1979-2009) Land-Use/Land-Cover Dataset of the Binational Santa Cruz Watershed (Open-File Report No. 2011-1131). U.S. Geological Survey.
Wallace, C.S.A.; Villarreal, M.L.; Norman, L.M. 2011. Development of a High-Resolution Binational Vegetation Map of the Santa Cruz River Riparian Corridor and Surrounding Watershed, Southern Arizona and Northern Sonora, Mexico (Open-File Report No. 2011-1143). U.S. Geological Survey.
Wilson, N.R.; Norman, L.M. 2018. Analysis of vegetation recovery surrounding a restored wetland using the normalized difference infrared index (NDII) and normalized difference vegetation in