sciendo

Journal of Environmental Geography 13 (3–4), 51–60.

DOI: 10.2478/jengeo-2020-0012
ISSN 2060-467X

A SOLUTION TO TREAT MIXED-TYPE HUMAN DATASETS FROM SOCIO-ECOLOGICAL SYSTEMS

Lisa B. Clark[1], Eduardo González[1,2*], Annie L. Henry[1], Anna A. Sher[1*]

[1]Department of Biological Sciences, University of Denver, Denver, CO, 80208-9010
[2]Department of Biology, Colorado State University, Fort Collins, CO, 80523
*Corresponding author, e-mail: edusargas@hotmail.com

Research article, received 26 August 2020, accepted 8 December 2020

Abstract

Coupled human and natural systems (CHANS) are frequently represented by large datasets with varied data including continuous, ordinal, and categorical variables. Conventional multivariate analyses cannot handle these mixed data types. In this paper, our goal was to show how a clustering method that has not before been applied to understanding the human dimension of CHANS: a Gower dissimilarity matrix with partitioning around medoids (PAM) can be used to treat mixed-type human datasets. A case study of land managers responsible for invasive plant control projects across rivers of the southwestern U.S. was used to characterize managers' backgrounds and decisions, and project properties through clustering. Results showed that managers could be classified as "federal multitaskers" or as "educated specialists". Decisions were characterized by being either "quick and active" or "thorough and careful". Project goals were either comprehensive with ecological goals or more limited in scope. This study shows that clustering with Gower and PAM can simplify the complex human dimension of this system, demonstrating the utility of this approach for systems frequently composed of mixed-type data such as CHANS. This clustering approach can be used to direct scientific recommendations towards homogeneous groups of managers and project types.

Keywords: Gower's similarity coefficient, partition around medoids clustering, human dimension, coupled human and natural systems, land management

INTRODUCTION

The human dimension of biological conservation, ecological restoration, and environmental management in a broad sense is a recent, growing focus in the scientific literature as an important component of coupled human and natural systems (CHANS, a.k.a. socio-ecological systems, Liu et al., 2007). When considering human systems, both within and separate from CHANS, large datasets are often involved due to the complex and varied nature of survey data. The CHANS framework can also yield data that are challenging to work with due to the interconnections between systems and data that encompasses multiple scales. Multivariate analyses are therefore frequently used, however survey and ecological data often include mixed types of variables (i.e., continuous, ordinal, and categorical), which cannot be treated by most conventional multivariate tests. For example, while cluster analyses are commonly used in social sciences, most use well known distance metrics such as Euclidean (e.g., García-Llorente et al., 2011) or Bray Curtis (e.g., Higuera et al., 2013), which cannot handle mixed data.

Gower dissimilarity matrices with clustering using partitioning around medoids (PAM) have been used recently as a new solution to the problem of mixed data in other disciplines, such as biomedical sciences, ecology, and socioeconomics (Table 1), but never in CHANS before now. The Gower similarity coefficient is specifically designed to deal with mixed data, which becomes even more likely when combining human and natural variables as found in CHANS. Gower also has additional advantages such as allowing for missing values and for different weights to be assigned to each variable (Gower, 1971; Legendre and Legendre, 2012). PAM is an alternative for the popular, non-hierarchical k-means method. Unlike those methods, PAM accepts other distance metrics besides Euclidean and is useful for relatively small sample sizes with outliers (Borcard et al., 2011; Kaufman and Rousseeuw, 1990).

This study illustrates the use of Gower distances with PAM to investigate the human dimension of coupled systems in a case study of managers involved in the control of an invasive tree (*Tamarix spp.*) in the riparian southwestern U.S. Of specific interest was whether there were profiles of managers or projects (based on their education, management role, experience, etc.) that were associated with particular management decisions. More generally, this study examines how characteristics and decisions of this population of managers and their projects could be more easily described through clustering. While many restoration ecology studies have inventoried management actions in river restoration projects (Bernhardt et al., 2007; Morandi et al., 2014), to the authors' knowledge, an in-depth, quantitative exploration of the characteristics of managers and their projects has not been done (except see Sher et al., 2020). Previous literature on the human dimension of

Table 1 Examples of papers using Gower similarity coefficients with partitioning around medoids (PAM) clustering to treat mixed data types.

Field	Examples
Biomedical science	Han et al. 2014, Canul-Reich et al. 2015, Hummel et al. 2017
Genetics	Krichen et al. 2008, Stefani et al. 2014
Marketing/Analytics	Silva et al. 2016, Lismont et al. 2017, Arunachalam and Kumar 2018
Sports research	Akhanli and Hennig 2017
Ecology	Williams et al. 2011, Pimenta et al. 2017
Socioeconomics	Kühne et al. 2010, Gellynck et al. 2011, Hennig and Liao 2013, Iparraguirre et al. 2013, Maione et al. 2018
Sociology	Bohensky et al. 2016, King et al. 2016

This selection was obtained from a search in Google Scholar using the chain "Gower and partitioning around medoids" done on Mar 21, 2018 that yielded 410 results. The list is not exhaustive.

restoration ecology has only focused on one aspect of decision-making such as partnerships or political input (e.g., Kallis et al., 2009; Oppenheimer et al., 2015), despite the myriad of aspects that may be important, including education level, governing organization, and collaboration. It was hypothesized that the proposed statistical method would give interpretable, meaningful clusters of managers, types of projects, and types of management decisions. We then tested the hypothesis that management decisions could be predicted by characteristics of managers and/or projects. This is important because if managers with particular characteristics are consistently making specific management decisions such as choosing to monitor their projects, then scientific recommendations regarding those decisions can be more accurately targeted toward the relevant managers.

We believe that this novel application of Gower distances with PAM will be useful to within any field that may study the natural-human interface with mixed-type data sets, including not only restoration ecology but also human geography, environmental sociology, and environmental psychology.

METHODS

Case study

Tamarix spp. (tamarisk, saltcedar) is a shrubby tree native to Eurasia that can grow in monocultures along riverways and impacts wildlife habitat (Bateman et al., 2013; Sogge et al., 2013; Strudley and Dalin, 2013), soil salinity (Ohrtman and Lair, 2013), and native plant communities (Friedman et al., 2005; Merritt and Poff, 2010). *Tamarix* is one of the most pervasive invasive riparian plants across the southwest U.S. and has also invaded other arid and semi-arid world regions such as Mexico, Argentina, Australia and South Africa (Sher, 2013). Removal of *Tamarix* is a common practice in river management (González et al., 2015), and there are many methods managers use to remove this species, including a broadly-dispersed biological control (Bean and Dudley, 2018). These projects are conducted on lands owned by a variety of agencies including federal (e.g., Bureau of Reclamation), state (e.g., state natural resource departments), local (e.g., conservancy districts), non-profit (e.g., The Nature Conservancy), and private (e.g., individual landowners).

In order to investigate the human dimension of the restoration of *Tamarix*-dominated lands, land managers of *Tamarix* removal projects were identified from a large dataset originally collected to assess the effects of removal method on vegetation (Fig. 1; see González et al., 2017). This was a collaborative effort of 16 research institutions; sites included all locations across the southwestern U.S. where data were available, distributed across the Upper Colorado, Lower Colorado, and Middle Rio Grande river basins. These managers were invited to participate in an online survey and in-person interviews in order to assess whether management decisions in these projects were associated with individual characteristics of those managers or projects. Information about the managers' backgrounds was needed, as well as the approach to restoration specific to each of their projects. The online survey was administered through Qualtrics to land managers. The 20-minute survey was tested through multiple iterations using mock interviews and through Qualtrics by trusted land managers and collaborators to ensure clarity.

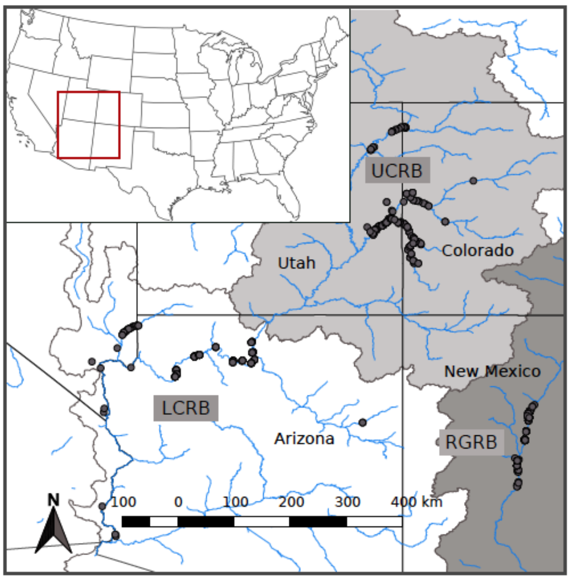

Fig. 1 Map of study area.
UCRB – Upper Colorado River Basin; LCRB – Lower Colorado River Basin; RGRB – Rio Grande River Basin. Points are *Tamarix* removal project sites

The survey was approved by the University of Denver Institutional Review Board (#816375-5), and it was fielded from August 2016 to March 2017 (Clark et al., 2019). We contacted 46 managers via email or phone; only one manager who was contacted did not complete the survey, thus our final sample size was 45 managers. The subsequent survey results encompassed 78 projects including 227 sites where *Tamarix* was treated (93% of treated sites originally sampled for vegetation data reported by González et al., 2017a; 2017b). See Table 2 for diversity of respondents. Seventeen managers had more than one project and 54 projects had multiple managers. Some of the variables were related to each manager, and others to specific projects. Thus, the data were considered in terms of managers (n=45) and projects (n=78). As this study represents nearly all *Tamarix* removal projects in the southwestern U.S. over the last 20 years, our sample size can be considered highly representative of this population.

The survey results produced continuous, ordinal, and categorical variables, organized into two general categories: characteristics and decisions (Table 3). Within each of these, some variables were specific to projects, while others were specific to managers regardless of the project, such as education level. The characteristics variables included: governing agency or organization ("agency"; Table 3), education, experience level, and management role. Agency was considered both in relation to the manager and to the project, as it often differed. Experience was also considered for the manager as an overall measure of management experience and for the project as a measure of location-specific experience. The decision variables covered the manager's goals for each project, degree of collaboration across agencies, information sources, *Tamarix* removal method, and monitoring methods. For information sources, managers were asked to rate the influence of information provided by particular agencies or organizations (e.g., formal: scientific articles; informal: peer conversations) on their decision-making, resulting in a count of the number of influential sources rated "somewhat influential" or higher. In the survey, managers also selected monitoring frequency for each type of monitoring method (e.g., physical, chemical, biological) but because most managers used more than one type and we were interested in how frequently any type of monitoring was done rather than each type, we created an ordinal variable for overall monitoring frequency where the highest frequency for any method was recorded.

Cluster analysis

Four cluster analyses were run for each of the variable categories – manager characteristics, project-specific characteristics, general management decisions, and project-specific management decisions – using partitioning around medoids (PAM method; Borcard et al., 2011; Kaufman and Rousseeuw, 1990) on a Gower dissimilarity matrix (Gower, 1971; Legendre and Legendre, 2012). The weighting of each set of variables was adjusted for each cluster analysis to give equal

Table 2 Summary of respondent characteristics

Characteristic	Proportion of each category
Gender	Men 47% | Women 53%
Education level	High school 4% | Bachelors 33% | Masters 47% | Doctorate 18%
Experience level	< 11 years 24% | 11-20 years 22% | > 20 years 42% | Did not identify as a land manager 9%
Owning agencies	*Federal* 49% | *State* 18% | *Local* 12% | *Private/Non-profit* 14% | *More than one agency* 8%
	Federal includes: National Park Service, Bureau of Land Management, US Fish and Wildlife Service, Bureau of Reclamation State includes: 1 state park service, 2 state natural resource departments, 3 state fish and wildlife services Local includes: 3 municipalities, 1 tribe, 1 conservation district Private/Non-profit includes: 1 non-profit, 1 private company, 1 university, individuals
Managing agencies	*Federal* 29% | *State* 12% | Local 9% | *Private/Non-profit* 14% | *More than one agency* 36%
	Federal includes: National Park Service, Bureau of Land Management, US Fish and Wildlife Service, Bureau of Reclamation State includes: 1 state natural resource department, 1 state fish and wildlife service, 1 conservancy district Local includes: 2 municipalities, 1 tribe Private/Non-profit includes: 1 non-profit, 1 private company, individuals
Employing agencies	*Federal* 47% | *State* 11% | *Local* 16% | *Non-profit* 20% | *Private* 4%
	Federal includes: National Park Service, Bureau of Land Management, US Fish and Wildlife Service, Bureau of Reclamation, US Army Corps of Engineers State includes: 4 state natural resource departments Local includes: 2 municipalities, 1 county, 2 conservation districts Non-profit includes: 3 non-profits, 1 private company, 1 university

Percentages are calculated based on the total number of managers (for gender, education, experience, and employing agency) or projects (for owning and managing agency).

Table 3 Twenty-four survey variables used for analysis, by cluster category: characteristics vs. decisions for managers and for specific projects.

Category	Variable type	Variable	Description
Manager characteristics	continuous	Management role	Number of roles out of: directly make decisions, implement decisions made by others, oversee projects with input from a partnership, collect data, other
	ordinal	Overall experience	<11 years, 11-20 years, >20 years
	ordinal	Education	High school, Bachelors, Masters, PhD
	categorical	Employing agency	Private, non-profit, local, state, federal
Project-specific charac.	ordinal	Experience in project area	<11 years, 11-20 years, >20 years
	categorical	Managing agency of proj.	Private/non-profit, local, state, federal, collaborative
	categorical	Owning agency of project	Private/non-profit, local, state, federal, collaborative
Management decisions	categorical	Type of inform. sources	Formal (e.g., peer-reviewed literature, conference talks), informal (e.g., conversations, past experience), mixed
	continuous	Number of information sources	-
	continuous	Number of monitoring methods	Includes visual, biological, physical, and chemical
	ordinal	Frequency of monitoring	Variable or <every 4 yrs, every 1-2 yrs, >annual
	continuous	Number of monitoring groups	Includes self, other personnel within agency, collaborators, university scientists, private consultants, other
	continuous	Number of collaborating groups	Includes federal personnel, state personnel, private consultants, scientists, neighbors/peers, other
	continuous	Number of collab. scientist groups	Includes federal, state, county, private consultants, non-profit agency, university, other
	continuous	Number of researching groups	Includes self, university scientists, other scientists
	continuous	Removal method (four variables)	Proportion of sites with each method (biocontrol only, cut-stump, heavy machinery, and burning) by manager
Project-specific decisions	categorical	Goals (14 variables)	Yes/no for each of 14 goals within the following categories: Plant, Wildlife, Water, People, Other
	continuous	Removal method (four variables)	Proportion of sites with each method (biocontrol only, cut-stump, heavy machinery, and burning) within a project

weight to each variable as the number of sub-variables was not consistent. However, for the project-specific management decisions, the goals of "none", "other", and "livestock forage" were only rarely selected in the surveys and drove the clustering in preliminary analyses, so were given a lower weight than the other goals.

All clustering methodologies assign observations to the same cluster based on algorithms that consider the distance (or similarity) between observations. The clustering algorithm used by the PAM method is an extension of the popular K-means algorithm, which uses Euclidean distances only and therefore cannot deal with categorical data. Unlike K-means, the PAM algorithm can be fed with a dissimilarity matrix, a matrix that contains all pairwise distances between the observations, instead of the raw data. This broadens the choice of distance measures to others that allow continuous as well as ordinal and categorical variables. The PAM algorithm computes k representative observations, called medoids, through an iterative process that ends when the average dissimilarity of the medoids to all the observations in the cluster is the minimal possible. As in K-means, the number of clusters (k) has to be defined *a priori*. We used the optimum average silhouette width (ASW) method to estimate the best number of clusters (Kaufman and Rousseeuw, 1990). In this study, the number of cluster groups was based on the highest average silhouette width that had a feasible logical interpretation, determined by significant differences on the survey variables between cluster groups using chi-square or Mann-Whitney U comparisons for each variable.

The dissimilarity matrix that was used to feed the PAM procedure was computed using the Gower similarity coefficient (Gower, 1971; Legendre and Legendre, 2012). The Gower metric has other advantages besides allowing mixed data (i.e., data that includes continuous, ordinal, and categorical variables). First, all variables including ordinal and categorical are scaled to [0,1] so the requirement of the PAM method of all variables being dimensionally homogeneous (Borcard et al., 2011) is met. Normality for continuous variables is not required. Second, missing values are discarded from the calculation without the need of removing the observation or the variable; so the dataset can include missing values and no power is lost. Third, it is possible to set different weights for each variable. In our case study, we assigned weights to the variables so that each set of variables (i.e., collaboration, role, etc.) was equal in weight. All continuous variables were scaled before calculating the Gower coefficients. The Gower's similarity coefficient between two observations (s_{ij}) is calculated following the equation:

$$S_{ij} = \frac{\sum_{k=1}^{p} w_{ijk}\, s_{ijk}}{\sum_{k=1}^{p} w_{ijk}} \quad (1),$$

where s_{ij} denotes the similarity of observations i and j for the k^{th} variable, and wij is the weight given to the k^{th} variable (a weight of 0 is given in case of missing values for i or j). The similarity s_{ijk} is defined for continuous and ordinal variables as

$$1 - \left| x_{ik} - x_{jk} \right| / r \quad (2),$$

where r is the range of the variable. For categorical variables, s_{ijk} is defined as 0 if x_{ik} and x_{jk} differ and 1 if x_{ik} and x_{jk} are the same. The Gower's dissimilarity matrix is computed by transforming the similarities of all pairs of observations as

$$\sqrt{(1 - s_{ij})} \quad (3).$$

The Gower coefficients were calculated using the function daisy of the package cluster (Maechler et al. 2018) and the PAM clustering were run using the function pamk of the package fpc (Hennig, 2013) in R 3.4.1 (R Core Team, 2017).

Cluster assessment

In order to define the profiles of the resulting cluster groups, the mean response to each of the variables used to run the cluster analysis for each of the four cluster group pairs were compared Pearson's chi-square tests, for categorical data, or Mann-Whitney U, for continuous data, in JMP 13.0.0 (SAS Institute, 2014). Mann-Whitney is a non-parametric test, selected because our continuous variables were rarely normally distributed. To determine if the "characteristic" cluster groups (both for managers and for projects) helped explain "decisions" variables, the same approach was used with characteristic clusters as the independent variable and individual decisions variables as the dependent variables. To account for the increased risk of a Type I error due to the large number of tests, a Bonferroni adjustment was applied to the alpha based on the number of analyses for each sub-question.

RESULTS AND DISCUSSION

In this case study, Gower similarity coefficients and PAM clustering was used to summarize survey data comprised of mixed variable types in a coupled human and natural system. This approach created four clear sets of clusters relating to manager characteristics, project characteristics, management decisions, and project-specific decisions based on survey responses by managers of invasive *Tamarix* removal projects. Surprisingly, the characteristic clusters did not, for the most part, explain management decisions, suggesting that individual managers did not make choices based on their background, but instead that these decisions may be more the product of the agency or collaborative group and determined by their resources and/or priorities (Sher et al., 2020). These results demonstrate the utility of this analysis approach and provide insight into the structure of this specific system, which can assist understanding of and thus communication with managers.

Previous research that has included surveys of managers has rarely investigated the linkage between managers backgrounds and management actions taken as determined by combinations of factors (but see Sher et al., 2020). More often, surveys of approaches (Bernhardt et al., 2007; Morandi et al., 2014) or attitudes of managers has been assessed (e.g., Curtis and de Lacy, 1998, Padgett and Imani, 1999), typically with very little if any quantitative hypothesis testing (but see Clark et al., 2019). At least one such study has implemented a multivariate clustering method for identifying opinions and attitudes of land managers toward implementing conservation initiatives (Knight et al., 2010), but no subsequent analysis appears to have been done with these clusters. In another, background was linked to management approach, but these traits were only considered singly, rather than contributing to a multi-dimensional profile (Raymond and Brown, 2011). It is our hope that this method of using Gower similarity coefficients and PAM clustering can help facilitate more studies of the hypothesized causal relationships between elements, as was done here.

Cluster results

For each of the four variable groups, distinct pairs were created by the cluster analysis (Table 4-7). Coefficients are either Mann-Whitney U (continuous variables) or Pearson's chi-square (ordinal and categorical variables) and indicate significant differences between the cluster groups if bolded ($p < 0.05$). The manager characteristics cluster groups were explained primarily by employing agency, education, and management role, with an equal number of managers in each group (Table 4). Managers in group 1 ("federal multitaskers") had lower education, worked for mostly federal agencies, and had more management roles including overseeing projects with input from a partnership, relative to group 2 ("educated specialists").

Projects were distinguished by all of the variables used in the analysis: local experience, managing agency, and owning agency (Table 5). Most projects in the first group were characterized by having more locally-experienced managers and tending to be owned and managed by larger or collaborative entities ("public") whereas group 2 projects ("private") were owned and

Table 4 Description of cluster groups created from manager characteristics (ASW=0.22), Group 1: Federal multitaskers (n=22), Group 2: Educated specialists (n=23)

Variable	Weight	Group 1	Group 2	coefficient	P
Role					
Direct management role	0.05	64%	65%	0.11	0.92
Implement decisions made by others	0.05	36%	9%	3.38	0.07
Oversee projects with input from a partnership	0.05	**86%**	**39%**	**11.62**	**<0.001**
Collect data	0.05	50%	26%	2.83	0.09
Median breadth of management roles (0-4)	0.05	**3**	**1**	**8.52**	**0.004**
Experience					
Most common experience level	0.25	>20 years	11-20 years	5.92	0.12
Education					
Most common education level	0.25	**Bachelors**	**Masters**	**18.20**	**<0.001**
Agency					
Most common employing agency	0.25	**Federal**	**Non-profit/ University**	**19.21**	**<0.001**

Table 5 Description of cluster groups created from project-specific characteristics (ASW=0.51), Group 1: Public (n=49), Group 2: Private (n=25)

Variable	Weight	Group 1	Group 2	coefficient	P
Experience					
Most common local experience level	0.33	**11-20 years**	**<11 years**	**8.84**	**0.01**
Managing agency					
Most common managing agency	0.33	**Collaborative**	**Private/Non-profit**	**63.28**	**<0.001**
Owning agency					
Most common owning agency	0.33	**Federal**	**Private/Non-profit /University**	**53.87**	**<0.001**

managed mostly by smaller organizations such as private companies or non-profits.

There were two groups from the general management decision cluster analysis; these were significantly distinguished by information sources, monitoring, and the use of heavy machinery to remove *Tamarix* (Table 6). The first group ("quick and active") used fewer sources of information but those sources were a mix of formal and informal; they used less comprehensive but more frequent monitoring and more heavy machinery than the other group ("thorough and careful"). The project-specific management decisions (Table 7) were characterized by the selection of goals related to ecosystem health such as native plant diversity or habitat improvement and more removal by burning (group 1: "ecocentric") while group 2 ("limited scope") had few goals selected but did select "none" or "other" (e.g., community involvement, water conservation, research) goals more often and had more removal by heavy machinery. These groupings provided an overview of the managers involved in *Tamarix* removal projects and the decisions they make, helping us understand which traits or aspects of projects are likely to be aligned.

This clustering tool also facilitated the analysis of relationships between variables. Numbers in Table 8 are the coefficients from either Mann-Whitney U or Pearson's chi-square tests depending on the type of variable. No significant relationships were found with Bonferroni adjusted α=0.004 and α=0.003 for general and project-specific decisions, respectively. Counter to predictions, no strong relationships between manager characteristics and decisions made about projects (as shown by non-significant pairwise comparisons with individual variables; Table 8a) were found (e.g., Hagger et al., 2017; Martin-Lopez et al., 2007; Roche et al., 2015). This result suggests that either managers exhibit no bias in decision making in these restoration projects based on their own backgrounds, and/or that there are enough other controls in place through mechanisms to overwhelm any such bias (Clark et al., 2019, Sher et al., 2020). These controls are likely to include the constraints and goals of specific agencies, the influence of collaborators, and the availability of resources for a given project. It is also possible that any influence of manager characteristics on decisions were too small to be detected by a sample of this size.

Similarly, whether projects were "public" or "private" did not strongly predict management decisions made about those projects, although private projects were more likely to have the listed goals, especially aesthetics and native plant diversity, than public projects (Table 8b). Public projects were more likely to have used biological control, but these results were not statistically

Table 6 Description of cluster groups created from general management decisions (ASW=0.18), Group 1: Quick and active (n=24), Group 2: Thorough and careful (n=20)

Variable	Weight	Group 1	Group 2	coefficient	P
Information source					
Most common type of information sources	0.125	**mix**	**formal**	**15.40**	**<0.001**
Median breadth of information sources (0-22)	0.125	**13**	**21**	**7.82**	**0.005**
Monitoring					
Median breadth of monitoring methods (0-4)	0.125	**2**	**3**	**10.11**	**0.002**
Most common monitoring frequency	0.125	**> once a year**	**> once a year or < every 4 years**	**8.09**	**0.04**
Collaboration					
Median breadth of monitoring groups (0-6)	0.0625	2	3	2.54	0.11
Median breadth of collaborating groups (0-7)	0.0625	3	2.5	0.10	0.75
Median breadth of science collaborators (1-7)	0.0625	4	4.5	0.78	0.38
Median breadth of researching groups (0-4)	0.0625	2	1	1.40	0.24
Removal method					
Mean proportion of biocontrol only	0.0625	0.18	0.05	2.98	0.08
Mean proportion of cut-stump	0.0625	0.18	0.28	3.42	0.06
Mean proportion of heavy machinery	0.0625	**0.39**	**0.09**	**4.07**	**0.04**
Mean proportion of burning	0.0625	0.16	0.19	0.42	0.52

Table 7 Description of cluster groups created from project-specific decisions (ASW=0.34), Group 1: Ecocentric (n=43), Group 2: Limited scope (n=29)

Variable	Weight	Group 1	Group 2	coefficient	P
Plant-related goals					
Native plant diversity	0.0606	**93%**	**28%**	**33.37**	**<0.001**
Ecosystem resilience	0.0606	**79%**	**7%**	**36.09**	**<0.001**
Exotic plant removal	0.0606	**95%**	**76%**	**6.01**	**0.01**
Wildlife-related goals					
Habitat improvement	0.0909	**100%**	**31%**	**41.06**	**<0.001**
Endangered species	0.0909	**65%**	**0%**	**30.90**	**<0.001**
Water-related goals					
Channel maintenance	0.0606	21%	14%	0.60	0.44
Restore over-bank flooding	0.0606	**60%**	**17%**	**13.20**	**<0.001**
Water quality	0.0606	**28%**	**3%**	**7.00**	**0.008**
People-related goals					
Aesthetics	0.0606	40%	38%	0.02	0.89
Recreation	0.0606	28%	34%	0.35	0.55
Wildfire mitigation	0.0606	**63%**	**21%**	**12.37**	**<0.001**
Other goals					
Livestock forage	0.0303	9%	0%	2.86	0.09
Other	0.0303	**2%**	**28%**	**10.10**	**0.002**
None	0.0303	**0%**	**14%**	**6.28**	**0.01**
Removal method					
Mean proportion of biocontrol only	0.0455	0.09	0.14	0.04	0.85
Mean proportion of cut-stump	0.0455	0.22	0.33	0.53	0.47
Mean proportion of heavy machinery	0.0455	**0.23**	**0.33**	**6.48**	**0.01**
Mean proportion of burning	0.0455	**0.25**	**0.11**	**4.57**	**0.03**

significant with an adjusted alpha of p<0.003. However, it should be noted that as this study sample represented a near-census of *Tamarix* projects in the southwestern U.S., such descriptive statistics may still be meaningful. Projects that were owned and managed privately or by non-profits may have had the flexibility to have more specific and customized goals than those projects that required buy-in from larger or more diverse stakeholders.

Taken together, this information can inform future collaborations with managers and scientists in this coupled system by giving context to their interactions. For example, managers who are federal multitaskers may not have the capacity to try new methods but educated specialists may be more willing and able to do so. Thus, educated specialists may be the best candidates to try innovative new practices and could be more directly targeted in communications and dissemination. Additionally, increased understanding of managers by scientists is essential for building trust in relationships with managers, which is crucial to the success of any collaboration (Vangen and Huxham, 2003).

Method assessment

When treating mixed-type data from CHANS systems, there are many advantages to using cluster analyses and PAM with Gower in particular. When cluster analysis is used, the whole dataset can be utilized rather than having to choose a priori which variables will be the most important, which has been the usual practice to treat mixed-type data to date. With PAM clustering and Gower, categorical variables – which are very common when assessing characteristics of people (e.g., education level or gender) – do not need to be omitted or converted

Table 8 Pairwise comparisons between manager (a) and project (b) characteristic cluster groups (columns) and management decisions (rows)

	Characteristic cluster groups			
a) General approach	**Federal multitaskers**	**Educated specialists**	**coefficient**	**P**
Information type	mixed/formal	mixed	0.79	0.67
Number of information sources	19	17	2.07	0.15
Number of monitoring groups	3	2	0.02	0.89
Number of monitoring methods	2	2	0.01	0.92
Monitoring frequency	> once a year	every 1-2 years	2.07	0.56
Number of collaborating groups	3	3	0.0006	0.98
Number of science collaborators	4	4	0.59	0.44
Number of researching groups	2	1	3.06	0.08
Biocontrol	0.15	0.09	2.18	0.14
Cut-stump	0.25	0.20	0.91	0.34
Heavy machinery	0.25	0.25	0.03	0.87
Burning	0.14	0.20	0.006	0.94
b) Project-specific approach	**Public**	**Private**	**coefficient**	**P**
Native plant diversity	**57%**	**84%**	**5.18**	**0.03**
Ecosystem resilience	47%	56%	0.55	0.46
Exotic plant removal	83%	96%	2.53	0.11
Habitat improvement	68%	80%	1.16	0.28
Endangered species	36%	80%	0.42	0.52
Channel maintenance	19%	16%	0.11	0.74
Restore over-bank flooding	40%	48%	0.38	0.54
Water quality	17%	20%	0.10	0.75
Aesthetics	**28%**	**60%**	**7.18**	**0.01**
Recreation	28%	36%	0.54	0.46
Wildfire mitigation	38%	60%	3.10	0.08
Livestock forage	2%	12%	3.03	0.08
Other	17%	4%	2.53	0.11
None	9%	0%	2.25	0.13
Biocontrol	**0.13**	**0.06**	**4.44**	**0.04**
Cut-stump	0.29	0.19	2.21	0.14
Heavy machinery	0.15	0.27	0.07	0.79
Burning	0.18	0.20	0.0006	0.98

in some way to a nominal numerical value that is then improperly represented. In addition to dealing with mixed type variables, PAM is also more robust to outliers than traditional methods (Arunachalam and Kumar, 2018; Maione et al., 2018), can deal with non-symmetrical data (Gellynck et al., 2011), and can be used for relatively small sample sizes like ours as is also common with human datasets (Iparraguirre et al., 2013; King et al., 2016). The Gower coefficient also allows for weighting of variables and missing values in the dataset. Unlike PAM, Gower is sensitive to outliers (Sander and Lubbe, 2018). Despite this drawback, this analysis method is one of the best solutions to dealing with mixed data types in a multivariate setting.

CONCLUSIONS

This study has demonstrated the application of a clustering method used in other fields of study to a CHANS context. PAM with Gower is useful in this study due to the need to comprehensively reflect complex data. In this way, managers and their decisions can be understood in a holistic manner and the cluster groupings can inform future recommendations and the allocation of resources. This method also has the potential to be useful in other CHANS studies such as endangered species management, grazing management, or water management where there are even more factors involved with the addition of politics and federal or state regulations.

Acknowledgements

This research was supported by a National Science Foundation Dynamics of Coupled Natural and Human Systems award (project number 1617463), a Pustmueller Grant, and the University of Denver. The authors thank all managers who participated in the survey and/or interview. The assistance of Carly McGuire, Madison Sligh, and Kayleigh Kearnan with data collection and entry was much appreciated. Thanks also go to N. Sayre and R. Lave for help with data collection protocols.

References

Akhanli, S.E., Hennig, C. 2017. Some Issues in Distance Construction for Football Players Performance Data. Archives of Data Science 2(1). DOI: 10.5445/KSP/1000058749/09

Arunachalam, D., Kumar, N. 2018. Benefit-based consumer segmentation and performance evaluation of clustering approaches: An evidence of data-driven decision-making. *Expert Systems with Applications* 111, 11–34. DOI: 10.1016/j.eswa.2018.03.007

Bateman, H.L., Paxton, E.H., Longland, W.S. 2013. Tamarix as Wildlife Habitat. In Sher, A.A., Quigley, M.T. (Eds.) Tamarix: A Case Study of Ecological Change in the American West. Oxford University Press, New York, 168–188. DOI: 10.1093/acprof:osobl/9780199898206.003.0010

Bean, D., Dudley, T. 2018. A synoptic review of *Tamarix* biocontrol in North America: tracking success in the midst of controversy. *BioControl* 63(3), 361–376. DOI: 10.1007/s10526-018-9880-x

Bernhardt, E.S., Sudduth, E.B., Palmer, M.A., Allan, J.D., Meyer, J.L., Alexander, G., Follastad-Shah, J., Hassett, B., Jenkinson, R., Lave, R., Rumps, J., Pagano, L. 2007. Restoring rivers one reach at a time: Results from a survey of U.S. river restoration practitioners. *Restoration Ecology* 15, 482–493. DOI: 10.1111/j.1526-100X.2007.00244.x

Bohensky, E.L., Kirono, D.G.C., Butler, J.R.A., Rochester, W., Habibi, P., Handayani, T., Yanuartati, Y. 2016. Climate knowledge cultures: Stakeholder perspectives on change and adaptation in Nusa Tenggara Barat, Indonesia. *Climate Risk Management* 12, 17–31. DOI: 10.1016/j.crm.2015.11.004

Borcard, D., Gillet, F., Legendre, P. 2011. Numerical Ecology with R. Springer, New York.

Canul-Reich, J., Hernández-Torruco, J., Frausto-Solis, J., Méndez Castillo, J.J. 2015. Finding relevant features for identifying subtypes of Guillain-Barré Syndrome using Quenching Simulated Annealing and Partitions Around Medoids. *International Journal of Combinatorial Optimization Problems and Informatics* 6(2), 11–27.

Clark, L.B., Henry, A.L., Lave, R., Sayre, N.F., González, E., Sher, A.A. 2019. Successful information exchange between restoration science and practice. *Restoration Ecology* 27(6), 1241–1250. DOI: 10.1111/rec.12979

Curtis, A., de Lacy, T. 1998. Landcare, stewardship and sustainable agriculture in Australia. *Environmental Values* 7, 59–78. DOI: https://www.jstor.org/stable/30302269

Friedman, J.M., Auble, G.T., Shafroth, P.B., Scott, M.L., Merigliano, M.F., Freehling, M.D., Griffin, E.R. 2005. Dominance of non-native riparian trees in western USA. *Biological Invasions* 7(4), 747–751. DOI: 10.1007/s10530-004-5849-z

García-Llorente, M., Martín-López, B., Nunes, P.A.L.D., González, J.A., Alcorlo, P., Montes, C. 2011. Analyzing the Social Factors That Influence Willingness to Pay for Invasive Alien Species Management Under Two Different Strategies: Eradication and Prevention. *Environmental Management* 48(3), 418–435. DOI: 10.1007/s00267-011-9646-z

Gellynck, X., Kühne, B., Weaver, R.D. 2011. Relationship quality and innovation capacity of chains: the case of the traditional food sector in the EU. *Proceedings in Food System Dynamics* 2(1), 1–22. DOI: 10.22004/ag.econ.100498

González, E., Sher, A.A., Anderson, R.M., Bay, R.F., Bean, D.W., Bissonnete, G.J., Bourgeois, B., Cooper, D.J., Dohrenwend, K., Eichhorst, K.D., El Waer, H., Kennard, D.K., Harms-Weissinger, R., Henry, A.L., Makarick, L.J., Ostoja, S.M., Reynolds, L.V., Robinson, W.W., Shafroth, P.B. 2017a. Vegetation response to invasive Tamarix control in southwestern U.S. rivers: A collaborative study including 416 sites. *Ecological Applications* 27(6), 1789–1804. DOI: 10.1002/eap.1566

González, E., Sher, A.A., Anderson, R.M., Bay, R.F., Bean, D.W., Bissonnete, G.J., Cooper, D.J., Dohrenwend, K., Eichhorst, K.D., El Waer, H., Kennard, D.K., Harms-Weissinger, R., Henry, A.L., Makarick, L.J., Ostoja, S.M., Reynolds, L.V., Robinson, W.W., Shafroth, P.B., Tabacchi, E. 2017b. Secondary invasions of noxious weeds associated with control of invasive Tamarix are frequent, idiosyncratic and persistent. *Biological Conservation* 213, 106–114. DOI: 10.1016/j.biocon.2017.06.043

González, E., Sher, A. A., Tabacchi, E., Masip, A., Poulin, M. 2015. Restoration of riparian vegetation: a global review of implementation and evaluation approaches in the international, peer-reviewed literature. *Journal of Environmental Management* 158, 85-94. DOI: 10.1016/j.jenvman.2015.04.033

Gower, J.C. 1971. A general coefficient of similarity and some of its properties. *Biometrics* 27(4), 857–871. DOI: 10.2307/2528823

Hagger, V., Dwyer, J., Wilson, K. 2017. What motivates ecological restoration? *Restoration Ecology* 25, 832–843. DOI: 10.1111/rec.12503

Han, S., Sung, K.R., Lee, K.S., Hong, J.W. 2014. Outcomes of laser peripheral iridotomy in angle closure subgroups according to anterior segment optical coherence tomography parameters. *Investigative Ophthalmology & Visual Science* 55, 6795–6801. DOI: 10.1167/iovs.14-14714

Hennig, C. 2013. fpc: Flexible procedures for clustering. R package version 21-5.

Hennig, C., Liao, T.F. 2013. How to find an appropriate clustering for mixed-type variables with application to socio-economic stratification. *Journal of the Royal Statistical Society. Series C: Applied Statistics* 62, 309–369. DOI: 10.1111/j.1467-9876.2012.01066.x

Higuera, D., Martín-López, B., Sánchez-Jabba, A. 2013. Social preferences towards ecosystem services provided by cloud forests in the neotropics: Implications for conservation strategies.

Regional Environmental Change 13, 861–872. DOI: 10.1007/s10113-012-0379-1

Hummel, M., Edelmann, D., Kopp-Schneider, A. 2017. Clustering of samples and variables with mixed-type data. *PloS ONE* 12(11), 1–24. DOI: 10.1371/journal.pone.0188274

Iparraguirre, J., Gentry, T., Pena, D. 2013. Vulnerability of Primary Care Organizations to the National Health Service Reform in England. *Applied Economic Perspectives and Policy* 35(4), 634–660. DOI: 10.1093/aepp/ppt021

Kallis, G., Kiparsky, M., Norgaard, R. 2009. Collaborative governance and adaptive management: Lessons from California's CALFED Water Program. *Environmental Science & Policy* 12, 631–643. DOI: 10.1016/j.envsci.2009.07.002

Kaufman, L., Rousseeuw, P. 1990. Finding Groups in Data: And Introduction to Cluster Analysis. Wiley seri. John Wiley and Sons Inc.

King, M.L., Hering, A.S., Aguilar, O.M. 2016. Building predictive models of counterinsurgent deaths using robust clustering and regression. *Journal of Defense Modeling and Simulation* 13(4), 449–465. DOI: 10.1177/1548512916644074

Knight, A. T., Cowling, R. M., Difford, M., & Campbell, B. M. 2010. Mapping human and social dimensions of conservation opportunity for the scheduling of conservation action on private land. *Conservation Biology* 24(5), 1348–1358. DOI: https://www.jstor.org/stable/40864035

Krichen, L., Martins, J.M.S., Lambert, P., Daaloul, A., Trifi-Farah, N., Marrakchi, M., Audergon, J.M. 2008. Using AFLP Markers for the Analysis of the Genetic Diversity of Apricot Cultivars in Tunisia. *Journal of the American Society for Horticultural Science* 133(2), 204–212. DOI: 10.21273/JASHS.133.2.204

Kühne, B., Vanhonacker, F., Gellynck, X., Verbeke, W. 2010. Innovation in traditional food products in Europe: Do sector innovation activities match consumers' acceptance? *Food Quality and Preference* 21(6), 629–638. DOI: 10.1016/j.foodqual.2010.03.013

Legendre, P., Legendre, L. 2012. Numerical Ecology. 3rd English. Elsevier Science, Amsterdam.

Lismont, J., Vanthienen, J., Baesens, B., Lemahieu, W. 2017. Defining analytics maturity indicators: A survey approach. *International Journal of Information Management* 37, 114–124. DOI: 10.1016/j.ijinfomgt.2016.12.003

Liu, J., Dietz, T., Carpenter, S.R., Alberti, M., Folke, C., Moran, E., Pell, A.N., Deadman, P. 2007. Complexity of Coupled Human and Natural Systems. *Science* 317, 1513–1516. DOI: 10.1126/science.1144004

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., Hornik, K. 2018. Cluster Analysis Basics and Extensions. R package version 2.0.7-1.

Maione, C., Nelson, D.R., Barbosa, R.M. 2018. Research on social data by means of cluster analysis. *Applied Computing and Informatics* 15(2), 153–162. DOI: 10.1016/j.aci.2018.02.003

Martin-Lopez, B., Montes, C., Benayas, J. 2007. The non-economic motives behind the willingness to pay for biodiversity conservation. *Biological Conservation* 139(1-2), 67–82. DOI: 10.1016/j.biocon.2007.06.005

Merritt, D.M., Poff, N.L. 2010. Shifting dominance of riparian *Populus* and *Tamarix* along gradients of flow alteration in western North American rivers. *Ecological Applications* 20(1), 135–152. DOI: 10.1890/08-2251.1

Morandi, B., Piégay, H., Lamouroux, N., Vaudor, L. 2014. How is success or failure in river restoration projects evaluated? Feedback from French restoration projects. *Journal of Environmental Management* 137, 178–188. DOI: 10.1016/j.jenvman.2014.02.010

Ohrtman, M.K., Lair, K.D. 2013. Tamarix and Salinity: An Overview. In: Sher, A.A., Quigley, M.T. (Eds.) Tamarix: A Case Study of Ecological Change in the American West. Oxford University Press, New York, 123–145. DOI: 10.1093/acprof:osobl/9780199898206.003.0008

Oppenheimer, J.D., Beaugh, S.K., Knudson, J.A., Mueller, P., Grant-Hoffman, N., Clements, A., Wight, M. 2015. A collaborative model for large-scale riparian restoration in the western United States. *Restoration Ecology* 23(2), 143–148. DOI: 10.1111/rec.12166

Padgett, D.D., Imani, N.O. 1999. Qualitative and quantitative assessment of land-use managers' attitudes towards environmental justice. *Environmental Management* 24(4), 509–515. DOI: 10.1007/s002679900250

Pimenta, V., Barroso, I., Boitani, L., Beja, P. 2017. Wolf predation on cattle in Portugal: Assessing the effects of husbandry systems. *Biological Conservation* 207, 17–26. DOI: 10.1016/j.biocon.2017.01.008

Raymond, C.M., Brown, G. 2011. Assessing conservation opportunity on private land: socio-economic, behavioral, and spatial dimensions. *Journal of Environmental Management* 92(10), 2513–2523. DOI: 10.1016/j.jenvman.2011.05.015

R Core Team. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Online available at: http://www.R-project.org/

Roche, L.M., Schohr, T.K., Derner, J.D., Lubell, M.N., Cutts, B.B., Kachergis, E., Eviner, V.T., Tate, K.W. 2015. Sustaining Working Rangelands: Insights from Rancher Decision Making. *Rangeland Ecology and Management* 68(5), 383–389. DOI: 10.1016/j.rama.2015.07.006

Sander, U., Lubbe, N. 2018. The potential of clustering methods to define intersection test scenarios: Assessing real-life performance of AEB. *Accident Analysis & Prevention* 113, 1–11. DOI: 10.1016/j.aap.2018.01.010

SAS Institute, I. 2014. JMP version 13.0.

Sher, A.A. 2013. Introduction to the Paradox Plant. In: Sher, A.A. and Quigley, M.T. (Eds.) Tamarix: A Case Study of Ecological Change in the American West. Oxford University Press, New York:1–18. DOI: 10.1093/acprof:osobl/9780199898206.003.0001

Sher, A.A., Clark, L., Henry, A.L., Goetz, A.R.B., González, E., Tyagi, A., Simpson, I., Bourgeois, B. 2020. The Human Element of Restoration Success: Manager Characteristics Affect Vegetation Recovery Following Invasive *Tamarix* Control. *Wetlands*. DOI: 10.1007/s13157-020-01370-w

Silva, F., Teixeira, B., Pinto, T., Santos, G., Vale, Z., Praça, I. 2016. Generation of realistic scenarios for multi-agent simulation of electricity markets. *Energy* 116, 128–139. DOI: 10.1016/j.energy.2016.09.096

Sogge, M. K., E. H. Paxton, Van Riper, C. 2013. Tamarisk in Riparian Woodlands: A Bird's Eye View'. In: Sher, A.A., Quigley, M.T. (Eds.) Tamarix: A Case Study of Ecological Change in the American West. Oxford University Press, New York, 189–206. DOI: 10.1093/acprof:osobl/9780199898206.003.0011

Stefani, F.O.P., Jones, R.H., May, T.W. 2014. Concordance of seven gene genealogies compared to phenotypic data reveals multiple cryptic species in Australian dermocyboid Cortinarius (Agaricales). *Molecular Phylogenetics and Evolution* 71, 249–260. DOI: 10.1016/j.ympev.2013.10.019

Strudley, S., Dalin, P. 2013. Tamarix as Invertebrate Habitat'. In: Sher, A.A. Quigley, M.T. (eds.) Tamarix: A Case Study of Ecological Change in the American West. Oxford University Press, New York, 207–224. DOI: 10.1093/acprof:osobl/9780199898206.003.0012

Vangen, S., Huxham, C. 2003. Nurturing Collaborative Relations: Building Trust in Interorganizational Collaboration. *The Journal of Applied Behavioral Science* 39(1), 5–31. DOI: 10.1177/0021886303039001001

Williams, J.N., Hollander, A.D., Geen, A.T.O., Thrupp, L.A., Hanifin, R., Steenwerth, K., Mcgourty, G., Jackson, L.E. 2011. Assessment of carbon in woody plants and soil across a vineyard-woodland landscape. *Carbon Balance and Management* 6(1), 1–14. DOI: 10.1186/1750-0680-6-11