CrossMark

# A practitioner's guide for exploring water quality patterns using principal components analysis and Procrustes

C. J. Sergeant · E. N. Starkey · K. K. Bartz ·
M. H. Wilson · F. J. Mueter

**Abstract** To design sustainable water quality monitoring programs, practitioners must choose meaningful variables, justify the temporal and spatial extent of measurements, and demonstrate that program objectives are successfully achieved after implementation. Consequently, data must be analyzed across several variables and often from multiple sites and seasons. Multivariate techniques such as ordination are common throughout the water quality literature, but methods vary widely and could benefit from greater standardization. We have found little clear guidance and open source code for efficiently conducting ordination to explore water quality patterns. Practitioners unfamiliar with techniques such as principal components analysis (PCA) are faced with a steep learning curve to summarize expansive data sets in periodic reports and manuscripts. Here, we present a seven-step framework for conducting PCA and associated tests. The last step is dedicated to conducting Procrustes analysis, a valuable but rarely used test within the water quality field that describes the degree of concordance between separate multivariate data matrices and provides residual values for similar points across each matrix. We illustrate the utility of these tools using three increasingly complex water quality case studies in US parklands. The case studies demonstrate how PCA and Procrustes analysis answer common applied monitoring questions such as (1) do data from separate monitoring locations describe similar water quality regimes, and (2) what time periods exhibit the greatest water quality regime variability? We provide data sets and annotated R code for recreating case study results and as a base for crafting new code for similar monitoring applications.

C. J. Sergeant (✉)
National Park Service, Inventory and Monitoring Program, Southeast Alaska Network, 3100 National Park Road, Juneau, AK, USA
e-mail: christopher_sergeant@nps.gov

E. N. Starkey
National Park Service, Inventory and Monitoring Program, Upper Columbia Basin Network, 105 E Second St, Suite 7, Moscow, ID, USA

K. K. Bartz
National Park Service, Inventory and Monitoring Program, Southwest Alaska Network, 240 W 5th Ave, #114, Anchorage, AK, USA

M. H. Wilson
National Park Service, Inventory and Monitoring Program, Chihuahuan Desert Network, New Mexico State University, MSC 3ARP 3655, Research Drive, Genesis Building D, Las Cruces, NM, USA

F. J. Mueter
School of Fisheries and Ocean Sciences, University of Alaska Fairbanks, 17101 Point Lena Loop Rd, Juneau, AK, USA

�（ Springer

## Introduction

The physical, chemical, and biological characteristics of water bodies have been described using a diverse array of analytical techniques. To design relevant and sustainable water quality monitoring programs, practitioners must have the ability to choose meaningful variables, justify the temporal and spatial extent of monitoring, and demonstrate that program objectives are successfully achieved after implementation. Consequently, data must be examined across several variables and often from multiple sites and seasons. These aspects of sampling design and water quality pattern analysis have commonly been explored using multivariate (or chemometric) techniques (e.g., Ouyang 2005; Olsen et al. 2012). Multivariate analyses have been accepted for decades as efficient tools for characterizing the range and variability of multiple water quality variables across broad temporal and spatial scales, and the number of monitoring papers using these techniques appears to be accelerating (Muangthong and Shrestha 2015).

Principal components analysis (PCA) is one of the simplest and most widely applied multivariate ordination techniques. Ordination seeks to display a set of multivariate observations within a lower two- or three-dimensional frame (Fig. 1). The hope is that points within this frame are arranged along interpretable axes corresponding to known environmental gradients (e.g., warming water temperature and decreasing dissolved oxygen). Within each axis, it is possible to determine which individual variables are responsible for the greatest observed variation. This useful characteristic

of PCA may assist monitoring programs with prioritizing limited resources by measuring variables explaining the majority of water quality regime variation.

While many articles in water quality literature have described results from PCA, we have found little guidance dedicated to clear explanations of the analytical decision-making process and annotated open source code necessary for successfully implementing PCA and associated tests. Since water quality practitioners have heavy workloads, including extensive field work and frequent reporting, clear and logical guidance for conducting PCA is a necessary contribution to water quality literature that will help specialists efficiently implement analyses in which they may not have time to learn independently. Here, we provide seven steps for conducting and interpreting PCA and Procrustes analysis using water quality data.

While many of the individual analytical steps for conducting PCA present multiple choices, we promote a generalized methodology that should be appropriate in almost all cases for the exploratory analysis of water quality data sets and point to other key sources for readers seeking more detail. We further illustrate the seven steps using three case studies of increasing complexity and provide the data files and R code necessary for reproducing the results (Online Resources 1–9). The complementary R code is based on existing multivariate statistical packages that have been widely used in environmental science literature. The latest version of R can be downloaded free of charge at the following web link: https://cran.r-project.org/bin/windows/base/.

In addition to PCA, we include guidance for conducting Procrustes analysis, a post hoc technique
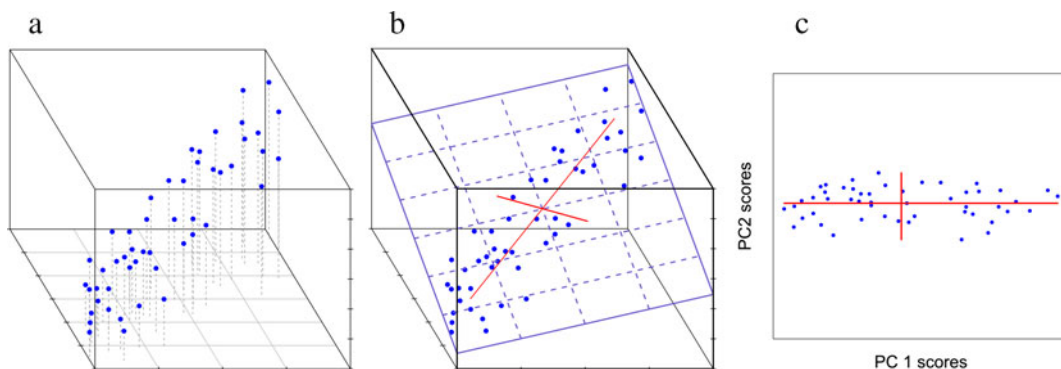


**Fig. 1** Generalized graphical representation of PCA, showing reduction from multivariate space (here, three-dimensional) to two dimensions (principal components). **a** Data in three-dimensional space. **b** Data are projected onto a two-dimensional surface; the first component, PC1, is drawn through the cloud's longest direction in order to maximize variation along that axis; the second component, PC2, is perpendicular to PC1 and maximizes the remaining variation. **c** The two-dimension surface is rotated to new axes defined by PC1 and PC2

that describes the degree of concordance between separate multivariate data matrices and individual points within each matrix. We assume that readers have a basic understanding of the utility of multivariate analysis and have established specific study or program objectives to be explored using PCA. For readers seeking more detailed background, McGarigal et al. (2000) offer a clearly written introduction to multivariate ordination and the underlying assumptions. Our goal was to create an analytical toolbox arming the first-time user of PCA with the knowledge to confidently adapt our case study data and complete a multivariate water quality analysis using open source statistical software.

## Seven steps for conducting PCA and associated tests

### Step 1: prepare data

The first step to conducting PCA is to select sampling units and the multivariate attributes (variables) measured for each unit. Data must be numeric and are typically continuous (infinite number of values); however, PCA has been used with categorical data by converting categories to scaled numeric values (Linting et al. 2007). Although we do not know of any formal analyses assessing the performance of mixed (continuous/categorical) data sets in PCA, it is generally not an accepted practice (McGarigal et al. 2000). In a typical PCA data matrix (also known as R-mode PCA), rows ($r$) represent the sampling unit (such as location), while columns ($c$) represent the measured variables (such as water temperature). For example, to determine whether 12 sites within a watershed demonstrate multivariate differences using median values for four water quality variables, the $r/c$ ratio would be 12:4. No missing values are allowed in the data matrix. Methodological descriptions of studies using PCA should always report how blank matrix cells were assigned a value (imputed) or include the rationale for removing missing data from the final data matrix (Olsen et al. 2012).

Data matrices for R-mode PCA should have more rows than columns. There are no universally accepted rules of thumb, but it is preferable to have at least two times more rows than columns. PCA results from data matrices with lower $r/c$ ratios should be interpreted more cautiously. Unlike many multivariate techniques, PCA performs well when several variables in the data set are correlated. Still, for analyses using large numbers of

variables, eliminating highly correlated variables in the data matrix (for example, chlorophyll-$a$ measured with two different methods) may improve the interpretation of results (see Olsen et al. 2012 for an expanded discussion of this topic).

Many water quality studies calculate variable means, medians, and/or standard deviations across daily to annual time scales for inclusion within a data matrix. The appropriate time scale for data aggregation and analysis should be chosen to adequately address study objectives (Sergeant and Nagorski 2014).

### Step 2: transform and standardize data

While data are typically assumed to exhibit multivariate normality, this is not strictly required and PCA results can still be interpreted if data are non-normal. Multivariate normality is difficult to formally assess, and no single test is widely accepted, but a useful surrogate measure is to check the univariate normality of each variable within the data matrix. Typical diagnostics for checking univariate normality include box plots, normal Q-Q plots, and the Shapiro-Wilk test of normality (see Zar 2010 for extended discussion of checking for univariate normality). We have included R code for a "check.norm" function that will examine the distribution of a single variable and check it against a normal distribution (Online Resource 1). Quantitative ecologists generally agree that using non-normal data is acceptable for *exploratory* analysis using PCA (Legendre and Legendre 1998; McGarigal et al. 2000), but the assumption of normality becomes more important when extending ordination results by using hypothesis testing with principal component (PC) scores (e.g., creating a linear regression of PC1 scores versus environmental variables). Scores are discussed further in step 6.

For severely non-normal data, a data transformation such as natural log, square-root, fourth-root, or Box-Cox may improve normality (see Zar 2010 for extended discussion of data transformations). A subset or all of the analyzed variables can be transformed. Methodological descriptions of studies using PCA should always state why certain variables were transformed before conducting PCA. A previous review of water quality studies using PCA found that sampling design and analytic methods were often poorly documented (Olsen et al. 2012). Since there are no universal rules for transforming variables, it may be advantageous for researchers to conduct separate PCA runs using transformed and non-transformed data to assess whether

ordination patterns lead to differing interpretation of results.

After constructing the final data matrix, values must be standardized to 0 mean and 1 standard deviation (SD). This can be accomplished using the "scale" function in R or basing the PCA on the correlation matrix (as opposed to the variance-covariance matrix; see lines for "princomp" in Online Resources 4, 6, and 9). Using the correlation matrix in PCA forces the variable values to be standardized. Since water quality variables are quantified on disparate measurement scales (e.g., dissolved oxygen in mg/L and specific conductance in mS/cm), data standardization ensures that no single variable with higher absolute values dominates the PCA, but that all variables are considered equally. We do not know of a case in water quality analysis in which it would be appropriate to use the variance-covariance matrix or non-standardized data.

Step 3: calculate eigenvalues and eigenvectors

The previous two steps lead to the development of the final PCA data matrix. Applying a PCA to this matrix (e.g., using "princomp" function in R; R Core Team 2014) will generate eigenvalues (variance explained by each PC) and eigenvectors (loadings or the importance of each variable to explaining variation in each PC).

The number of PCs in the output ("Comp. 1-4", Fig. 2) matches the number of variables in the data matrix. If variables in the data matrix are highly correlated, the proportion of variance explained by the initial PCs ("Proportion of variance", Fig. 2a) will be high. It is tempting to take advantage of this pattern by adding more variables to a PCA data matrix, but this often leads to redundancies in the data set and complicates pattern

**a**

```
Importance of components:
                        Comp. 1 Comp. 2 Comp. 3 Comp. 4
Standard deviation      1.39474 1.08079 0.90046 0.27517
Proportion of variance  0.48632 0.29203 0.20270 0.01893
Cumulative proportion   0.48632 0.77835 0.98106 1.00000
```

**b**

```
Loadings:
        Comp. 1 Comp. 2 Comp. 3 Comp. 4
SC      -0.083   0.716  -0.691   0.050
DO       0.686  -0.141  -0.179   0.691
pH      -0.181  -0.682  -0.695  -0.140
Temp    -0.700  -0.047   0.086   0.707
```

Fig. 2 Example R output from "princomp" function using data from case study B (Lapwai Creek, 2008). **a** Summary of eigenvalues produced by calling "summary" in R. **b** Summary of eigenvectors produced by calling "print(loadings())" in R

interpretation. In the same way, environmental models should strive to add additional variables cautiously (Burnham and Anderson 2002), PCA data matrices should only include variables that directly address study objectives and stated hypotheses.

Step 4: determine the number of principal components to interpret

This step seeks to identify principal components containing interpretable information. Some or most components may simply reflect noise (e.g., measurement error) in the underlying variables. Indeed, the goal of PCA is to separate interpretable information (signal) from noise or minor sources of variability. Several methods have become popular for determining the number of PCs to interpret, but the most robust and easily calculated approach is the broken-stick model, which we define here by paraphrasing Legendre and Legendre (1998). The model considers the variance across all PCs to be contained within a stick of unit length. If PCA divided the variance randomly across all PCs, the variation explained by each PC would be similar to the lengths of pieces obtained by breaking the unit stick into as many random pieces as there are PCs. Therefore, PCs explaining a proportion of the variance smaller than predicted by the broken stick null model should not be interpreted. To allow readers to consider this visually, we have included an annotated spreadsheet describing the calculation of the broken-stick model and a graphical representation of the results (Online Resource 2). This is a key analysis for justifying the number of interpretable PCs for a given PCA.

Another less conservative but widely used method for determining the number of PCs to interpret is the scree plot, in which eigenvalues are plotted in descending order against their corresponding PCs. In an ideal scree plot, the highest and most influential eigenvalues create a steep curve, followed by a relatively straight horizontal line formed by the lowest eigenvalues. Only PCs corresponding to eigenvalues creating the curved portion of the plot should be included for interpretation. However, scree plots with obvious break points are rare, especially in analyses with few PCs (Jackson 1993). A modification of the scree plot involves graphing the eigenvalues as described above and overlaying a second scree plot of mean eigenvalues generated with $N \approx 1000$ iterated matrices of random, uncorrelated data. The PCs to the left of where these lines cross are considered

interpretable (Horn 1965). Case study B illustrates the use of the broken-stick model, and Case study C considers both the broken-stick model and the modified scree plot. Jackson (1993) provides detailed descriptions of these two tests and additional tests not covered in this article.

Step 5: examine eigenvectors

After determining the appropriate number of PCs to interpret, it is critical to determine which variables significantly contribute to directional trends along each PC (see biplot description in "Step 6: plot and interpret ordination results" section) based on their loadings (Fig. 2b). Rules of thumb vary in complexity (McGarigal et al. 2000), but a widely accepted and simple qualitative rule proposes that loadings greater than 0.30 or less than −0.30 are significant, loadings greater than 0.40 or less than −0.40 are more important, and loadings greater than 0.50 or less than −0.50 are very significant (Hair et al. 1987; McGarigal et al. 2000). Similar to choosing appropriate variables for the data matrix (steps 1 and 3), it is the responsibility of individual researchers to properly justify the interpretation for each loading and not over-interpret findings. Qualitative rules of thumb such as the one presented here are generally acceptable for exploratory analyses, but researchers using PCA for statistical hypothesis testing should consider more rigorous approaches such as bootstrapped confidence intervals for loadings (McGarigal et al. 2000; Peres-Neto et al. 2003).

Step 6: plot and interpret ordination results

Biplots are the main tool for visualizing the results of a PCA in two dimensions (Figs. 3, 4, and 5). Individual points on a biplot represent one sample (row) of multivariate data from the data matrix. The horizontal and vertical axes represent the principal components and the coordinates of the points along these axes are referred to as PC1 and PC2, respectively. Vectors based on the loadings (step 5) are usually superimposed on a biplot to aid interpretation of the spread of samples in the ordination diagram (Figs. 3, 4, and 5). Vectors radiate out from the origin and represent the influence of individual variables on the variability among samples. Longer vectors represent larger absolute values for particular loadings. Points that plot farther from the origin in the direction of a particular vector are more strongly
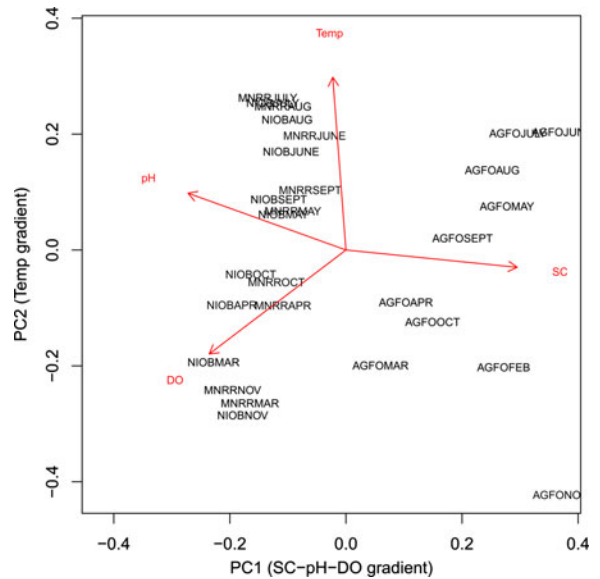


**Fig. 3** PCA biplot based on the monthly means of four water quality parameters measured on the Niobrara River at Agate Fossil Beds NM (AGFO), Niobrara National Scenic River (NIOB), and Missouri National Recreational River (MNRR) in 2014. Each PCA score is labeled as 'PARKMMM' where 'PARK' stands for the park alpha code and 'MMM' stands for the first three letters of each month

associated with that variable. The location of each point relative to each PC, other points, and each vector are the main considerations for visually interpreting a biplot. Default biplots in R can be difficult to interpret due to overlapping labels. We provide code for default and slightly polished versions of biplots (Figs. 3 and 4; Online Resources 4, 6, and 9), but producing a highly refined biplot figure often involves customized R code (Fig. 5) or manipulating default R visualizations with vector graphics software such as Adobe Illustrator or Inkscape.

Step 7 (optional): compare multiple data matrices using Procrustes analysis

Often, specific study objectives require comparing two or more multivariate data sets that differ across time or space. For example, researchers might be interested in comparing data matrices collected at different time scales within a single site (e.g., determine whether PCA results are similar between hourly and weekly data collected at the same monitoring site; Sergeant and Nagorski 2014) or data matrices collected simultaneously from different sites (see "Case study C" section below). Both the Mantel test and the Procrustes

**Fig. 4** PCA biplot based on the weekly means of four water quality parameters measured in Lapwai Creek (2008, 2011, and 2014). Each PCA score is labeled as 'MMMW', where MMM are the first three letters of the month and W denotes the letters a–e representing the week of the month. For figure clarity, font size for individual scores has been increased and 2-digit year code removed. Figures produced using case study B R code will have slight differences in appearance
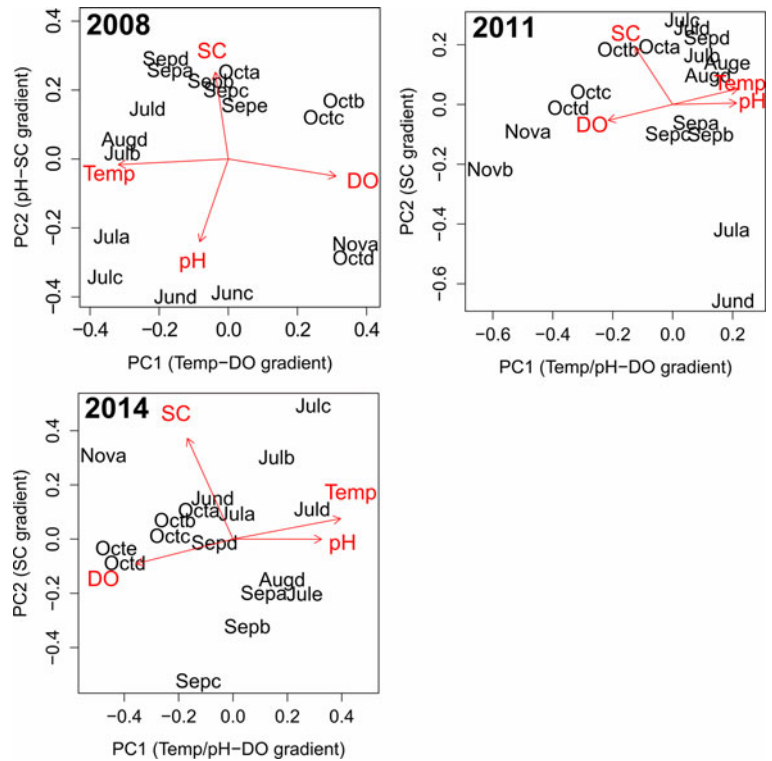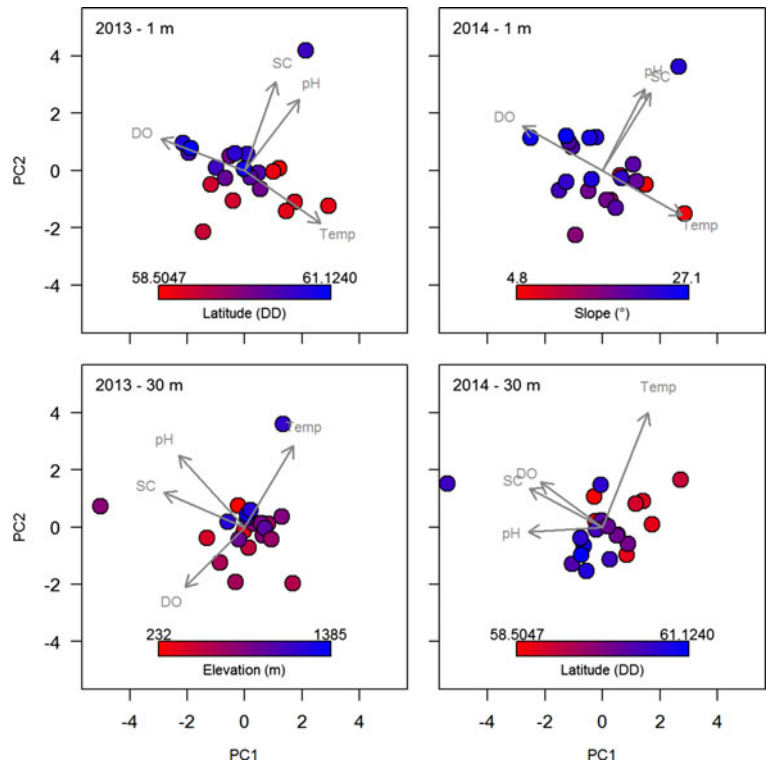


**Fig. 5** PCA biplots from annual water quality data collected across 2 years (2013 and 2014), two depths (1 and 30 m), and 21 lakes in the Southwest Alaska Network. Color-coding integrates gradients in select environmental variables correlated with PC1 and PC2

superimposition provide statistics describing the similarity between two data (distance) matrices (Peres-Neto and Jackson 2001). The Mantel test, which provides a single statistic describing the degree of similarity between two matrices, is relatively simple and widely used. However, Procrustes analysis not only produces a correlation-like statistic ($m^2$) for two matrices, but also provides residual values that allow for the comparison of individual scores (Olden et al. 2001; Peres-Neto and Jackson 2001; Sergeant and Nagorski 2014). Residual values can often clearly illustrate when a single multivariate observation is much more variable in comparison to the complete set of observations in an ordination. A Procrustes randomization test, also known as PROTEST, can be used to assess whether the fit between two matrices ($m^2$) is statistically significant (see a more detailed synopsis of Procrustes analysis and PROTEST here: http://jackson.eeb.utoronto.ca/procrustes-analysis/). We believe Procrustes analysis has been under-used in water quality studies and holds great promise for many applications, including describing short-term anomalies in water quality patterns before sufficient data are collected for more rigorous trend assessment. We provide interpretations for the results of Procrustes analysis in case studies B and C.

## Case studies

The following three case studies, which steadily increase in analytical complexity, are organized according to the seven steps for conducting PCA described earlier. Each step below is denoted by (1) through (7) to allow

readers to easily cross-reference case studies with the generalized steps above.

Each case study analyzes data collected by long-term monitoring networks within the National Park Service Inventory and Monitoring Program (NPS I&M; Table 1). Each NPS I&M network (32 nationally) collects long-term monitoring data for ecologically similar clusters of parks (Fancy et al. 2009). Within NPS I&M, the four variables of water temperature, pH, dissolved oxygen (DO), and specific conductance (SC) are referred to as the "core parameters," and are measured in nearly every long-term water quality program (NPS 2002). Due to the increasing affordability, reliability, and broad usage of water quality sensors, we use these core parameters as the basis for comparing results of multivariate analyses from freshwater systems. Note that in the seven steps leading up to these case studies, we have referred to water quality parameters as measured "variables" to remain consistent with other statistical literature. Within the following case studies, the terms "variables" and "parameters" are used interchangeably.

### Case study A

#### Objectives

In 2014, the NPS Northern Great Plains Network, in collaboration with the US Geological Survey (USGS), monitored the NPS core water quality parameters at three sites along the Niobrara River in northern Nebraska moving downstream from the headwaters, in order: Agate Fossil Beds National

**Table 1**  Summary of case study characteristics and associated online resources

| Case study | Spatial-temporal extent | Matrix dimensions (rows × columns) | Exploratory questions | Online resources |
|---|---|---|---|---|
| A | One river; three sites; 1 year | $28 \times 4$ | Do water quality regimes differ across a single watershed? | 1, 2, 3, 4 |
| B | One river; one site; 3 years | $17 \times 4$ | Do seasonal water quality regimes at the same site differ across years? Does shifting the collection site downstream change water quality regime characterization? | 1, 2, 5, 6 |
| C | 21 lakes; two depths; 2 years | $21 \times 4$ | Does multivariate ordination create intuitive lake groupings based on water quality regimes? Do the groupings differ among years? Do the groupings differ among depths? Do any environmental variables correlate with the groupings? | 1, 2, 7, 8, 9 |

Monument (AGFO; USGS gage 06454100), Nio-
brara National Scenic River (NIOB; USGS gage
06461500 at river km 229), and Missouri River
National Recreational River (MNRR; USGS gage
06465500 at river km 24). At each site, multi-
parameter sondes logged water quality data at 15-
min intervals from February (AGFO) or March
(NIOB, MNRR) through mid-November. Our main
objective for conducting PCA was to explore basic
differences in water quality regimes across a single
river continuum and relate these differences to
known geologic variability throughout the water-
shed. Detailed methods can be found in Wilson
and Wilson (2014).

*Approach*

(1) Continuous measurements for each of the four core
parameters were aggregated into monthly aver-
ages. Data matrix rows represented the sampling
location and month, while columns represented the
monthly mean values of the four core parameters
(row/column ratio 28:4; Online Resource 3).

(2) Data for each location/month combination were
evaluated for normality using the "check.norm"
function (Online Resource 1). Results of the
Shapiro-Wilk test of normality indicated that
SC and pH were non-normally distributed (both
$p = 0.0001$). However, natural log data transfor-
mations did not significantly improve normality
for SC or pH. Since the objective of this anal-
ysis was exploratory rather than strict hypothe-
sis testing, we did not transform the data for the
final analysis. Data were standardized to 0
mean and 1 SD to ensure all variables were
considered equally.

(3) The PCA function "princomp" (R Core Team
2014; Online Resource 4) generated the eigen-
values, which are the proportion of the variance
explained by each PC. For the Niobrara River sites,
92 % of the variance was explained by the first two
principal components (PC1 = 57 % PC1,
PC2 = 35 %; Table 2).

(4) Using the broken-stick model (Online Resource 2),
the observed variances accounted for by PC1
(2.28) and PC2 (1.38) were greater than expected
if variance was randomly divided across all com-
ponents (PC1 = 2.08, PC2 = 1.08). In contrast, the
observed variance for PC3 (0.22) and PC4 (0.11)

**Table 2**　Summary of the percent variance attributed to PC1–PC4 for each PCA

| Case study | Year (depth) | PC1 | PC2 | PC3 | PC4 |
|---|---|---|---|---|---|
| A | 2014 | *57* | *35* | 6 | 3 |
| B | 2008 | 49 | *29* | *20* | 2 |
| | 2011 | *74* | 20 | 5 | 1 |
| | 2014 | *62* | 23 | 14 | 1 |
| C | 2013 (1 m) | 46 | *38* | 11 | 6 |
| | 2013 (30 m) | 45 | *30* | 16 | 10 |
| | 2014 (1 m) | 44 | *40* | 10 | 7 |
| | 2014 (30 m) | *57* | 22 | *15* | 7 |

Italicized values denote that the associated PC was interpretable
based on broken-stick model results

were less than the expected broken-stick values
(PC3 = 0.58 and PC4 = 0.25). These calculations
provide clear support for focusing interpretation
on PC1 and PC2 trends.

(5) We used simple threshold value rules of thumb to
determine which variables contributed most
strongly to the trends along PC1 and PC2. Load-
ings greater than 0.50 or less than −0.50 are 'very
significant' (McGarigal et al. 2000). SC, pH, and
DO were most strongly associated with PC1 and
similar in loading magnitude (Table 3). Along
PC2, water temperature had a much greater loading
than the three other variables in that component
(Table 3).

(6) In order to visualize the results of this PCA, ordi-
nation scores were graphed using "biplot" (R Core
Team 2014; Fig. 3; Online Resource 4). Individual
variable loadings indicated that SC along PC1 and
water temperature along PC2 were most strongly
associated with variability in water quality across
samples. Pairs of monthly scores for NIOB and
MNRR grouped closely on the left side of the
graph, illustrating a high degree of water quality
concordance between these two sites despite
200 km of separation along the river channel.
AGFO scores plotted separately from NIOB and
MNRR on the right side of the biplot indicating the
influence of increasing SC values on
distinguishing the water quality regime of this
headwater site. In general, higher SC is correlated
with increased amounts of dissolved material in the
water column. At AGFO, the underlying geologic
formation is comprised of the Arikaree Group, a
fine-grained sandstone with localized beds of

**Table 3** Summary of the loadings attributed to PC1 and PC2 for each PCA. Also, see Fig. 3 from Olsen et al. (2012) for an alternative graphical presentation of loading values

| Case study | Year (depth) | PC1 loadings | | | | PC2 loadings | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | DO | pH | SC | Temp | DO | pH | SC | Temp |
| A | 2014 | −0.506 | −0.584 | 0.633 | 0.049 | −0.495 | 0.270 | −0.082 | 0.822 |
| B | 2008 | 0.686 | −0.181 | −0.083 | −0.700 | −0.141 | −0.682 | 0.716 | −0.047 |
| | 2011 | −0.547 | 0.540 | −0.318 | 0.554 | −0.259 | 0.023 | 0.931 | 0.257 |
| | 2014 | −0.552 | 0.504 | −0.259 | 0.612 | −0.234 | −0.001 | 0.953 | 0.194 |
| C | 2013 (1 m) | −0.647 | 0.422 | 0.240 | 0.588 | 0.245 | 0.551 | 0.687 | −0.407 |
| | 2013 (30 m) | −0.461 | −0.509 | −0.624 | 0.373 | −0.467 | 0.557 | 0.268 | 0.632 |
| | 2014 (1 m) | −0.617 | 0.325 | 0.369 | 0.615 | 0.347 | 0.630 | 0.602 | −0.346 |
| | 2014 (30 m) | −0.477 | −0.574 | −0.565 | 0.351 | 0.346 | −0.040 | 0.300 | 0.888 |

volcanic ash and silty sands (Gutentag et al. 1984) that are likely to produce a higher SC. At NIOB and MNRR, the principal geologic unit is the younger Ogallala Formation which is comprised of loosely cemented coarse-grained sand and gravels thus, producing a lower SC in this part of the Niobrara River (Gutentag et al. 1984).

Monthly scores from the warmer months (May through September) for all three parks along the Niobrara River were plotted along the top of the biplot, which corresponds with higher water temperature and lower DO (Fig. 3). Water temperatures for these warmer monthly averages ranged from 14.4 °C at AGFO to 24.6 °C at MNRR, while DO ranged from 3.3 mg/L at AGFO to 9.6 mg/L at MNRR. The scores for the cooler months (February, March, April, October, and November) were plotted at the bottom of the biplot for all three parks. In November, the AGFO site had extremely cool water and high SC relative to all other scores. Water temperatures for these cooler monthly averages ranged from 0.5 °C at AGFO to 12.2 °C at MNRR, while DO ranged from 5.7 mg/L at AGFO to 12.6 mg/L at NIOB. The stretch of the river that flows through AGFO is narrow with heavily vegetated banks. Potential factors influencing the water quality regime of the AGFO reach include vegetative decay, low dissolved oxygen, and water withdrawal.

(7)  Since this case study only examined one data matrix, a Mantel test or Procrustes analysis was not relevant. Case studies B and C illustrate the utility of Procrustes analysis.

*Summary*

Our simplest case study demonstrates that differences and similarities in multivariate patterns of water quality within discrete reaches of the same river can be explored quickly and efficiently using PCA. We concluded that SC and water temperature were important drivers of the Niobrara River water quality regime variation. The observable differences in SC at AGFO versus NIOB and MNRR are likely attributable to known differences in underlying geologic composition. Additionally, vegetative decay, low dissolved oxygen, and water withdrawal may be affecting the health of the Niobrara River reach within and adjacent to AGFO (Bowles et al. 2013; Spurgeon et al. 2014).

Case study B

*Objectives*

The Upper Columbia Basin Network seasonally monitored the NPS core water quality parameters in Lapwai Creek, Nez Perce National Historical Park, Idaho, during 2008, 2011, and 2014. Data were recorded hourly between June and November (see detailed methods in Starkey et al. 2008). The long-term objective of the monitoring effort is to determine the status and trends of the core water quality parameters, but assessing meaningful trends requires a longer time series than the existing 3 years. Using PCA and Procrustes analysis, we (1) tested whether the seasonal water quality regimes from three monitoring years were significantly different and (2) described unique water quality patterns at a

single site across the years before formal trend analyses were feasible. These analyses provided insight into whether the monitoring program was functioning as intended and effectively characterizing water quality regimes.

*Approach*

(1)   Each monitoring year, weekly means were calculated for core parameters from the end of June through the first week in November. Data were included in each of the three final data matrices (17 rows representing each week of monitoring by 4 columns representing weekly mean values of each core parameter; Online Resource 5) if they met predefined data quality standards determined by the degree of fouling and sensor drift (Wagner et al. 2006; Starkey et al. 2008). All rows with missing values were removed from the data matrix. Once these rows were removed for each year, the yearly data matrices were trimmed to be the same dimensions (Online Resource 5).

(2)   Data for each parameter/year combination were evaluated for normality using the "check.norm" function (Online Resource 1). Results of the Shapiro-Wilk test of normality indicated that the following parameters were non-normally distributed: 2008 SC ($p = 0.002$), 2011 SC ($p = 0.002$), 2008 DO ($p = 0.039$), 2011 DO ($p = 0.009$), 2014 DO ($p = 0.040$), 2008 pH ($p = 0.017$), and 2014 pH ($p = 0.038$). Common data transformations such as natural log did not improve normality. Therefore, to simplify interpretation of the proceeding ordinations, no transformations were used in the final analysis. Data were standardized to 0 mean and 1 SD.

(3)   PCA was implemented separately for each of the three yearly data matrices and summarized using the "princomp" function (R Core Team 2014; Online Resource 6). For each year, greater than 98 % of the observed variation was explained by the first three principal components (Table 2).

(4)   Using the annotated spreadsheet template (Online Resource 2), we examined a broken-stick model for each year's PCA results. During 2011 and 2014, model results suggested that interpretation should be focused mainly on PC1 trends. In 2008, the observed variance explained in PC1 (1.95) was slightly less than if the expected variance was

divided randomly across all components (2.08). But, the observed variance for PC2 (1.17) and PC3 (0.80) were greater than the expected broken-stick values (PC2 = 1.08 and PC3 = 0.58). These results, in combination with examining a standard scree plot for the 2008 PCA, which drops quickly after PC3 (Online Resource 6), suggested that PC1 through PC3 trends were interpretable. The PCA interpretations using the same data matrices supplemented with weekly standard deviations were similar to the results presented here. For other data sets, adding a measure of variability to the data matrix may increase clarity for interpreting ordination results (e.g., Sergeant and Nagorski 2014).

(5)   We focused our interpretation on loadings greater than 0.40 or less than −0.40 ("significant" or "very significant;" McGarigal et al. 2000). Water temperature and DO were the highest PC1 loadings across all years. PC1 loadings for pH were the third highest in 2011 and 2014. SC loadings for PC1 were never greater than 0.40 or less than −0.40. The highest loading for PC2 each year was SC, and in 2008, pH was also less than −0.4 (Table 3).

(6)   PCA ordinations for each year's data matrix were visualized using biplot (R Core Team 2014; Fig. 4). PC1 scores followed similar patterns each year. DO and water temperature loading vectors radiated in opposite horizontal directions, indicative of a strong negative correlation between these variables. Weekly scores from the warmest months (July/August) were plotted at the extreme end of the water temperature vector (highest mean weekly water temperature) and opposite the direction of the DO vector (lowest mean weekly DO). Weekly scores from the coolest months (October/November) were plotted at the extreme end of the DO vector (highest mean weekly DO) and opposite the direction of the water temperature vector (lowest mean weekly water temperature). Weekly scores from August and September 2014 were plotted in the opposite direction of the SC vector and were likely influenced by an uncharacteristic stream discharge pulse in late August (E. N. Starkey, unpublished data). Similarly, weekly scores from late June and early July 2011 scores were influenced by rain events that lead to a sudden drop in SC (Starkey 2012). The 2014 biplot and loadings are very similar to both 2008 and 2011, despite the

monitoring site moving approximately 150 m downstream in 2014 due to changes in the stream channel preventing deployment of the data logger (E. N. Starkey, unpublished data). It is important to note that the orientation of PCs is arbitrarily assigned during PCA, so the water temperature and DO vectors for 2008 (Fig. 4) can be flipped to appear more similar to 2011 and 2014.

(7) After interpreting biplots, we implemented Procrustes analyses to address two questions: (1) did water quality regimes defined by the core parameters vary significantly among pairs of years, and (2) was the 2014 water quality regime at the monitoring station moved downstream significantly different from the upstream station location in 2008 and 2011? Procrustes analyses were performed on sets of PCA scores for each combination of paired years (Online Resource 6). There was significant similarity between the multivariate ordinations produced by data matrices from each paired year combination (2008 versus 2011, $m^2 = 0.453$, $p = 0.001$; 2008 versus 2014, $m^2 = 0.369$, $p = 0.001$; 2011 versus 2014, $m^2 = 0.522$, $p = 0.001$). Across all comparisons, six of the seven highest residuals between individual sample points were from June or July, indicating these months may be inherently variable as the stream stabilizes from spring and early summer high flows (Table 4). If monitoring program capacity became limited, this analysis shows that it may be important to at least continue collecting data during these months to ensure that the range of physical-chemical stream dynamics is accurately characterized each year. Except for one high residual value in the 2011 versus 2014 comparison (the fourth week of June 2011; Table 4), there was no indication that the water quality regime at the 2014 monitoring station differed significantly from the upstream site used the previous two monitoring years (Fig. 4).

Summary

This case study demonstrated the utility of combining PCA with Procrustes analyses to evaluate the similarity of water quality regimes at a single monitoring site across years. Our main conclusions were as follows: (1) water quality regimes did not differ across 3 years, (2) DO and

**Table 4** The five highest Procrustes residuals for each paired annual comparison (2008 versus 2011, 2008 versus 2014, and 2011 versus 2014) in case study B (Lapwai Creek)

| Week[a] | Comparisons | | |
| | 2008 versus 2011 | 2008 versus 2014 | 2011 versus 2014 |
| --- | --- | --- | --- |
| Jund11 | | | 0.416 |
| Julc08 | 0.334 | | |
| Jula11 | | | 0.311 |
| Novb11 | | | 0.251 |
| Jula08 | 0.227 | | |
| Jund08 | | 0.220 | |
| Julb08 | | 0.216 | |
| Octb08 | | 0.209 | |
| Sepd08 | | 0.202 | |
| Sepb08 | 0.199 | | |
| Octa08 | | 0.192 | |
| Sepa08 | 0.190 | | |
| Nova11 | | | 0.186 |
| Sepd11 | | | 0.185 |
| Octd08 | 0.177 | | |
| Mean[b] | 0.148 | 0.131 | 0.146 |

[a] Weeks are labeled MMMWYY, where MMM is the first three letters of the month, W is the week of the month (a–d), and YY are the last two digits of the year. For example: July, Week 1, 2008 = Jula08

[b] The mean of all 17 residual values within each paired annual comparison

water temperature were the main drivers of multivariate water quality regime patterns, and (3) moving the monitoring station location in 2014 did not appear to drastically alter water quality regime observations. While data continue to be collected in this system to allow for a rigorous water quality trend analysis, the analyses presented in this case study allowed us to address basic monitoring program concerns and begin to determine the variables driving water quality regime patterns across multiple years. These results can be used to inform managers or other interested parties that the long-term monitoring program appears to be functioning as intended.

Case study C

Objectives

The Southwest Alaska Network contains two of the largest lakes in the National Park system, Naknek Lake

(58,824 ha) and Lake Clark (31,117 ha), as well as hundreds of smaller named and unnamed lakes. The network has monitored core parameters in these large lakes since 2008 and in smaller lakes since 2009. At each lake, measurements are taken along a vertical profile at fixed depth increments of 0, 1, 2, 3, 4, and 5 m, and then every 5 m down to 50 m or the lake bottom, whichever is reached first. Because of the difficulty involved in accessing the smaller lakes, measurements are made once per lake per year, near the mid-point on a single day between late July and late August, when water temperatures typically exhibit the greatest vertical stratification (Shearer et al. 2015).

The water quality data from these smaller lakes are temporally limited but spatially complex. The spatial complexity arises from the number of lakes and depths sampled, multiplied by the number of variables of interest (up to 37 lakes × 15 depths × 4 variables). The only other analysis of these data focused on one variable (pH) measured at one depth (5 m) in two lakes (Wilson and Moore 2013). Here, we used multivariate techniques to explore additional dimensions of the data. Using PCA and Procrustes analysis, we asked, do multivariate ordinations group lakes intuitively on the basis of their water quality data? If so, (1) do the groupings differ among years; (2) do the groupings differ among depths; and (3) do any environmental variables correlate with the groupings?

*Approach*

(1)   The water quality data for smaller lakes span 6 years and 15 depths, but few lakes were monitored in all years at all depths. Therefore, we selected years and depths so as to maximize the number of lakes in the analysis. Years 2013 and 2014 were chosen because the number of lakes sampled exceeded previous years ($n = 32$ in 2013 and 37 in 2014 versus $n = 1$–16 in 2009–2012). Depths of 1 and 30 m were chosen to represent the epilimnetic and hypolimnetic waters (above and below the thermocline), after visual assessments of graphed and tabular data suggested that (1) temperatures consistently stabilized below 25 m, and (2) sample sizes decreased with each subsequent depth interval, because many smaller lakes have depths <50 m. Including only two depths simplified the case study, but a PCA data matrix could easily be augmented with additional depths.

We trimmed the water quality data to include only smaller lakes with data in both 2013 and 2014, down to at least 30-m depth ($n = 21$ lakes). The resulting four data matrix dimensions were $21 \times 4$ (rows of lakes × columns of core parameter values) for each year-depth combination: 2013 = 1 m, 2013 = 30 m, 2014 = 1 m, and 2014 = 30 m (Online Resource 7).

In addition to water quality data, several environmental variables were quantified for each lake using ArcGIS software (v10.2, Environmental Systems Research Institute, Redlands, CA, USA) or field measurements (Online Resource 8). Specifically, lake centroid coordinates (latitude and longitude) and total watershed area ($m^2$) were estimated from lake and watershed boundary data layers (https://irma.nps.gov/App/Reference/Profile/2216048 and https://irma.nps.gov/App/Reference/Profile/2216046 [both accessed 28 September 2015]). Average watershed elevation (m) and slope (°) were estimated from a digital elevation model (NPS unpublished data), and maximum depth (m) was measured on site with a standard recreational depth finder.

(2)   The water quality data matrices were evaluated for normality using the "check.norm" function in R (Online Resource 1). Shapiro-Wilk test results indicated that, in all four matrices, DO was distributed normally ($W > 0.933$; $p > 0.158$), SC was distributed non-normally ($W < 0.532$; $p < 0.001$), and temperature and pH distributions were matrix-dependent (i.e., normal in some matrices, but not in others). Common data transformations (e.g., ln, square root, Box-Cox) produced mixed results. For example, natural log transformation significantly improved the normality of temperature, but other variables' normality improved only marginally (e.g., SC) or worsened (DO and pH). Therefore, as in case studies A and B, no transformations were applied in the final analysis. Data were standardized to 0 mean and 1 SD.

(3)   PCA was implemented for each data matrix via the "princomp" function in R (Online Resource 9). For each PCA, the first two PCs explained 74–84 % of the observed variation in each data matrix (Table 2).

(4)   In comparison to case studies A and B, broken-stick model results were more challenging to interpret (Online Resource 9; Table 2). For the 2013 at 1

m PCA, only PC2 was considered interpretable, while PCs 2–4 were considered interpretable for 2013 at 30 m (Table 2). For the 2014 at 1 m PCA, PCs 2 and 4 were considered interpretable, while PCs 1, 3, and 4 were considered interpretable for 2014 at 30 m (Table 2). We also employed scree plots to determine which PCs to interpret. To increase the reproducibility of the scree plot assessment, eigenvalues generated by the four PCAs were compared with eigenvalues averaged from 1000 random, uncorrelated datasets (Horn 1965). Results indicated that interpretation should focus on PCs 1 and 2 in all PCAs except 2014—30 m, for which only PC1 should be interpreted (Fig. 6; Online Resource 9).

(5) The loadings followed two general patterns, depending on depth. For the 1-m analyses, DO and temperature had the strongest loadings on PC1, and SC and pH had the strongest loadings on PC2 (Table 3). In contrast, PC1 at 30-m depth was most strongly associated with SC and pH (Table 3). Temperature was the strongest driver of PC2, followed by pH in 2013 and DO in 2014.

(6) Ordination results for each of the four data matrices were visualized using biplots. As with the loadings, different patterns emerged for 1 and 30 m depths. At 1-m depth, larger values of PC1 corresponded to lower DO and higher temperature, while larger values of PC2 corresponded to higher pH and SC. In contrast, at 30-m depth in 2013, larger values of PC1 corresponded to lower pH and SC, and larger PC2 values indicated higher temperatures and lower DO. At 30-m depth in 2014, temperature and DO vectors were perpendicular (uncorrelated), while DO and SC vectors were positively correlated and radiating in similar directions. These initial biplot results can be recreated using R code from Online Resource 9.

(7) The patterns evident in the loadings and biplots suggested that sampled lakes grouped in a meaningful way. Thus, we used two separate Procrustes analyses to examine whether water quality differed among years at a given depth (2013 versus 2014 at 1 m; 2013 versus 2014 at 30 m). Both analyses indicated significant similarity among years (at 1 m: $m^2 = 0.170$, $p = 0.001$; at 30 m: $m^2 = 0.557$, $p = 0.003$). Two additional Procrustes analyses were performed to assess differences among depths within a given year: (1 vs. 30 m in 2013 data; 1 vs. 30 m in 2014). Despite the apparent
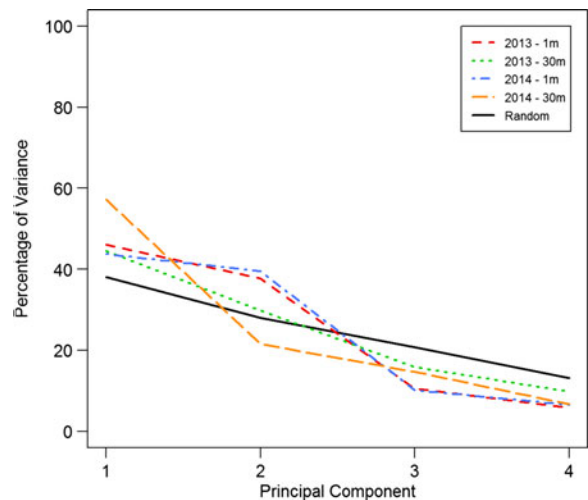


**Fig. 6** Modified scree plot for case study C (see step 4). *Colored lines* represent eigenvalues from four PCAs. The *black line* signifies mean eigenvalues from 1000 PCAs generated using random data. The *point where the black line crosses each colored line* indicates the maximum limit where components are considered interpretable (Horne 1965)

depth-related pattern in the loadings and biplots, Procrustes analyses also indicated significant similarity among depths (in 2013: $m^2 = 0.738$, $p = 0.024$; in 2014: $m^2 = 0.657$, $p = 0.005$).

Biplots created with the "plotrix" package "plot" function (R Core Team 2014) enabled color-coded, simultaneous visualization of PC scores and external environmental gradients that were not included in the analysis (Fig. 5). The choice of which environmental variable to overlay on a given biplot was made by selecting the variable that correlated most strongly with PCs 1 and 2 (i.e., $\sqrt{(r_{PC1}^2 + r_{PC2}^2)}$ was maximized ($r_{max}$)). Specifically, lake latitude was selected for two biplots: 2013 at 1 m ($r_{max} = 0.754$) and 2014 at 30 m (($r_{max} = 0.636$). Watershed elevation and slope were overlayed on the other two biplots (2013 at 30 m ($r_{max} = 0.434$) and 2014 at 1 m ($r_{max} = 0.768$), respectively).

*Summary*

This case study demonstrated the utility of PCA and Procrustes analyses to efficiently explore lake water quality patterns in a spatially complex dataset. Although the loadings and biplots suggested that patterns in water quality differed among epilimnetic (1 m) and

hypolimnetic (30 m) waters, Procrustes analyses indicated that differences were not significant. Likewise, we found no significant differences among the years included in our analyses. This case study also showcased difficulties encountered when applying PCAs to real-world data; for example, data transformations sometimes created less normal data distributions, and broken-stick models were difficult to interpret in the final case study. Here, we offered acceptable workarounds that enabled the exploratory analyses to proceed. Finally, this case study presented R code for overlaying gradients in environmental variables and automating processes when running multiple multivariate analyses.

## Conclusions

Across diverse river and lake systems, our three case studies from national parklands demonstrated that PCA and Procrustes analysis are useful tools for exploring patterns in large water quality data sets and can be efficiently implemented by water quality monitoring practitioners to address basic questions of program design and effectiveness. Case study A illuminated differences in water quality regimes based on known geological differences within the watershed. Case study B validated that the water quality regime for a site appeared unchanged despite moving the monitoring location downstream 150 m. Procrustes residuals in case study B clearly determined that June and July were highly variable across years relative to other monitored months and may require higher resolution or longer-term sampling to adequately capture water quality dynamics. Case study C demonstrated the utility of PCA and Procrustes analysis in describing spatially complex water quality data sets for diverse lakes and the utility of incorporating informative environmental covariates.

Although PCA is a widely used statistical technique, modern studies still include avoidable analytical errors and oversights. In a recent comprehensive review of 49 water quality studies using PCA, Olsen et al. (2012) found that a large proportion of articles did not discuss how missing data were imputed, used redundant variables in the final data matrix, and did not describe how non-normal data values were transformed. The generalized steps presented earlier and R code provided in the online resources accompanying this article promote a standardized approach to water quality analysis and

visualization that may help overcome some of the analytical errors and oversights observed in earlier literature (Olsen et al. 2012). Our hope is that this content creates the impetus necessary for water quality monitoring programs to more widely integrate rigorous multivariate analyses into reporting frameworks and share complex results with other stakeholders.

## References

Bowles, D. E., Peitz, D. G., & Cribbs, J. T. (2013). Aquatic invertebrate community structure in the Niobrara River, Agate Fossil Beds National Monument, Nebraska, 1996-2009. *Great Plains Research, 23*, 1–10.

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: a practical information-theoretic approach*. New York, NY: Springer.

Fancy, S. G., Gross, J. E., & Carter, S. L. (2009). Monitoring the condition of natural resources in US National Parks. *Environmental Monitoring and Assessment, 151*, 161–174.

Gutentag, E. D., Heimes, F. J., Krothe, N. C., Luckey, R. R., & Weeks, J. B. (1984). Geohydrology of the high plains aquifer in parts of Colorado, Kansas, Nebraska, New Mexico, Oklahoma, South Dakota, Texas, and Wyoming. USGS Professional Paper 1400-B.

Hair, J. F., Jr., Anderson, R. E., & Tatham, R. L. (1987). *Multivariate data analysis*. New York, NY: MacMillan.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30*, 179–185.

Jackson, D. A. (1993). Stopping rules in principal components analysis: a comparison of heuristical and statistical approaches. *Ecology, 74*, 2204–2214.

Legendre, P., & Legendre, L. (1998). *Numerical ecology*. Amsterdam, The Netherlands: Elsevier Science B. V.

Linting, M., Meulman, J. J., Groenen, P. J. F., & Van der Kooij, J. J. (2007). Nonlinear principal components analysis: introduction and application. *Psychological Methods, 12*, 336–358.

McGarigal, K., Cushman, S., & Stafford, S. (2000). *Multivariate statistics for wildlife and ecology research*. New York, NY: Springer Science + Business Media.

Muangthong, S., & Shrestha, S. (2015). Assessment of surface water quality using multivariate statistical techniques: case study of the Nampong River and Songkhram River, Thailand. *Environmental Monitoring and Assessment*. doi: 10.1007/s10661-015-4774-1.

NPS. (2002). Recommendations for core water quality monitoring parameters and other key elements of the NPS vital signs program water quality monitoring component. National Park Service white paper, Fort Collins, Colorado. http://www.nature.nps.gov/water/vitalsigns/assets/docs/COREparamFINwSIGpg.pdf. Accessed 18 September 2015].

Olden, J. D., Jackson, D. A., & Peres-Neto, P. R. (2001). Spatial isolation and fish communities in drainage lakes. *Oecologia, 127*, 575–585.

Olsen, R. L., Chappell, R. W., & Loftis, J. C. (2012). Water quality sample collection, data treatment and results presentation for principal components analysis—literature review and Illinois River watershed case study. *Water Research, 46*, 3110–3122.

Ouyang, Y. O. (2005). Evaluation of river water quality monitoring stations by principal components analysis. *Water Research, 39*, 2621–2635.

Peres-Neto, P. R., & Jackson, D. A. (2001). How well do multivariate data sets match? The advantages of a procrustean superimposition approach over the Mantel test. *Oecologia, 129*, 169–178.

Peres-Neto, P. R., Jackson, D. A., & Somers, K. M. (2003). Giving meaningful interpretation to ordination axes: assessing loading significance in principal component analysis. *Ecology, 84*, 2347–2363.

R Core Team. (2014). *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org/. Accessed 18 September 2015.

Sergeant, C. J., & Nagorski, S. A. (2014). The implications of monitoring frequency for describing riverine water quality regimes. *River Research and Applications, 31*, 602–610.

Shearer, J., Moore, C., Bartz, K. K., Booher, E. C. J., & Nelson, J. (2015). *Monitoring freshwater systems in the Southwest Alaska Network: Standard operating procedures. Natural Resource Report NPS/SWAN/NRR—2015/925.1*. Fort Collins, Colorado: National Park Service.

Spurgeon, J. J., Stasiak, R. H., Cunningham, G. R., Pope, K. L., & Pegg, M. A. (2014). Status of native stream fishes within selected protected areas of the Niobrara River in Western Nebraska. *Great Plains Research, 24*, 71–78.

Starkey, E. N. (2012). *Upper Columbia Basin Network integrated water quality annual report 2011: Nez Perce National Historical Park (NEPE). Natural Resource Technical Report NPS/UCBN/NRTR—2012/571*. Fort Collins: National Park Service.

Starkey, E. N., Garrett, L. K., Rodhouse, T. J., Dicus, G. H., & Steinhorst, R. K. (2008). *Upper Columbia Basin Network integrated water quality monitoring protocol: narrative version 1.0. Natural Resource Report NPS/UCBN/NRR—2008/026*. National Park Service: Fort Collins, CO.

Wagner, R. J., Boulger Jr., R. J., Oblinger, C. J., & Smith, B. A. (2006). Guidelines and Standard procedures for continuous water-quality monitors: station operation, record computation, and data reporting: U.S. Geological Survey Techniques and Methods 1–D3, 51.

Wilson, T. L., & Moore, C. (2013). *A review of lake vertical profile monitoring in the Southwest Alaska Network: recommendations for future efforts. Natural Resource Technical Report NPS/SWAN/NRTR—2013/689*. Fort Collins, Colorado: National Park Service.

Wilson, M. H., & Wilson, S. K. (2014). *Water quality monitoring protocol for wadeable streams and rivers in the Northern Great Plains Network: Standard operating procedures version 1.0. Natural Resource Report NPS/NGPN/NRR—2014/868.1*. Fort Collins, Colorado: National Park Service.

Zar, J. H. (2010). *Biostatistical analysis*. Upper Saddle River, NJ: Prentice-Hall.