

Climate Zonation Analysis for Bighorn Canyon National Recreation Area

Michael Tercek, Ph.D.

Walking Shadow Ecology

PO Box 1085

Gardiner, Montana 59030 USA

www.YellowstoneEcology.com

March, 2010

EXECUTIVE SUMMARY

The Greater Yellowstone Network (GRYN), which is a division of the National Park Service (NPS) Inventory and Monitoring (I&M) Program, seeks to understand climate status, variability, and trends in its three park units. Averaging weather observations across an entire park would be of little value because it would obscure much of the spatial (place to place) variability among regions. The analysis presented here defines four climate zones with consistent seasonal and temporal dynamics. These weather station groups will be treated as reporting units in annual climate status summaries for Bighorn Canyon National Recreation Area, northeast Wyoming, USA. Three statistical techniques were used: (1) Cluster analysis of seasonal weather station data (temperature and precipitation), (2) Principal Components Analysis of long-term monthly variability, and (3) Analysis of snow cover to define elevation-based stratification of stations. The map shown in Figure 5 indicates the final grouping of weather stations.

INTRODUCTION

The Greater Yellowstone Network (GRYN), which is a division of the National Park Service (NPS) Inventory and Monitoring (I&M) Program, seeks to understand climate status, variability, and trends in its three park units (Frakes et al. 2009, Kittell et al. 2010). Climate is a key driver of variability in the other ecological “vital signs” that are monitored by the network (Frakes et al. 2009), and human industrial activity is very likely to cause rapid climate changes in the next century (Solomon et al. 2007). In this context, the goals of the GRYN I&M program are (1) to evaluate variability and trends in key climate parameters and (2) to provide climate data for analyses of other vital signs (Frakes et al. 2009).

In topographically complex areas, averaging weather observations across an entire park would be of little value because it would obscure much of the spatial (place to place) variability among regions (Gray 2008). Furthermore, there are likely to be differing temporal (year to year) changes observed among groups of weather stations, and when these changes are in conflict they should be reported separately (Gray 2008). Therefore, the goal of this report is to define groups of weather stations in and near Bighorn Canyon National Recreation Area (northeastern Wyoming, USA) that have (1) similar average seasonal (month to month) patterns of temperature and precipitation, and (2) similar long-term patterns of variability, over the period 1895-2008. An additional analysis of snow cover timing is used to determine whether there is stratification among weather stations according to elevation.

The groups of weather stations defined by this analysis will function as climate zones, or strata, that will be used as reporting units in annual climate status reports and in periodic climate trend reports produced by GRYN. The statistical methods in this report are being used to define climate zones in eight other NPS units, including Yellowstone and Grand Teton National Parks (Tercek et al. In prep).

METHODS

Weather Station Selection

Weather stations in and near Bighorn Canyon National Recreation Area (BICA) were identified by Davey et al. (2006). The analyses in this report include all weather stations within 40 km of BICA that are operated by the National Weather Service Cooperative Observer Program (COOP) and by the

Natural Resources Conservation Service SNOTEL network. Weather stations maintained by other agencies were not included either because they report observations for only a portion of each year, have data that are not commensurate with each other, or have been in operation for a time period that is not sufficient to establish long-term averages (Davey et al. 2006). Nine weather stations were selected for the analyses. Station names appear in the figures, see Results.

Cluster Analysis of 1971 – 2000 monthly normals.

Monthly average maximum temperature (Tmax), minimum temperature (Tmin) and precipitation data were downloaded for the period 1971 – 2000 from the websites of the Western Regional Climate Center (WRCC) (www.wrcc.dri.edu), the Natural Resources Conservation Service (NRCS) (www.wcc.nrcs.usda.gov), and the Parameter-elevation Regressions on Independent Slopes Model (PRISM) climate group (<http://www.prism.oregonstate.edu/>). For each weather station, Diurnal Temperature Range (DTR) and mean temperature (Tmean) were calculated for each month from Tmax and Tmin. Tmax and Tmin often vary at different rates, making them both important to the analysis, but DTR and Tmean usually show lower correlation with each other than Tmax and Tmin, which gives greater power to cluster analysis (Easterling et al. 1997, Vose et al. 2005, Kittell et al. 2010).

For COOP stations, NCDC monthly normal data were incorporated into the analyses directly from the WRCC web site. In contrast, only precipitation data from the NRCS web site were used for the SNOTEL stations. This was done because SNOTEL stations often have poor quality temperature data or missing temperature values that make it impossible to calculate accurate 1971-2000 normals (Kittell et al. 2010). In order to replace the missing temperature data, 1971 – 2000 normals were extracted from 800m resolution, gridded PRISM data. PRISM data do not suffer from the same quality problems as individual SNOTEL temperature data because the values for each grid cell are interpolated from a network of surrounding stations (Daly et al. 1994, 2001). Monthly average 1971 – 2000 Tmax and Tmin values were extracted from the PRISM grid cells occupied by each SNOTEL station.

Since the goal of the cluster analysis was to group weather stations according to their similarity in seasonal patterns, regardless of the absolute magnitude of the observations, the data were standardized as z-scores. For each weather station separately, the annual averages for precipitation, DTR, and Tmean were subtracted from their respective monthly values and the result was divided by the annual standard deviations for each variable. This produced a data matrix of 36 values for each

weather station, with 12 monthly standardized values each for precipitation, DTR, and Tmean.

Hierarchical Agglomerative Cluster analysis was performed following the methods of Fovell and Fovell (1993) and Unal et al. (2003). Euclidean distance was the distance metric. Both Ward's and Average clustering algorithms were used and the results were compared (Unal et al. 2003). If a cluster of weather stations appeared unchanged in dendrograms produced by both Ward's and Average clustering, it was considered robust to changes in method and retained. If a cluster contained different weather stations in Ward's vs. Average clustering, it was treated as a polytomy (unstructured group) and a larger cluster of stations containing this polytomy was examined for similarity between the two clustering algorithms. Successively larger clusters of stations were examined until groups with the same membership in both Ward's vs. Average clustering were found throughout the entire dendrogram.

Confidence levels were assigned to each cluster of stations with bootstrapping (Suzuki and Shimodaira 2006). Ten thousand pseudo-replicates of the data set were produced and the nodes in the final dendrogram were labeled with the percentage of pseudo-replicate dendrograms that contained each cluster.

The cluster analysis was performed with the R scripting language (R development core team 2009).

Principal Components Analysis of 1895 – 2008 monthly temperature and precipitation

Data for the Principal Components Analysis (PCA) were extracted from 1895-2008 monthly, 800m resolution, gridded data produced by the PRISM climate group (Daly et al. 2000, <http://www.prism.oregonstate.edu/>). Monthly time series for Tmax, Tmin, and precipitation were extracted from the grid cells occupied by each weather station. Tmean for each station was calculated from Tmax and Tmin. Each precipitation and Tmean time series was split into separate summer (June, July, August) and winter (December, January, February) time series for the period 1895 – 2008. This resulted in four separate data sets for analysis: winter precipitation, summer precipitation, winter Tmean, and summer Tmean.

PRISM data have two advantages over raw weather station data. First, since PRISM data for each grid cell are interpolated by using a network of surrounding stations, including the station occupying the grid cell, it is possible to extract time series that extend further back in time than the individual weather stations themselves. Second, PRISM data have been quality checked and peer-

reviewed. As a result, they are less prone to biases such as urban heat island effects, station moves, instrument changes, and incompleteness of record that make long-term analyses of raw weather station data difficult (Daly et al. 2000, 2001; <http://www.prism.oregonstate.edu/>).

S-mode PCA was performed using methods adapted from Serrano et al. (1999) and Comrie and Glenn (1998). S-mode PCA is based on a data matrix in which weather stations occupy the columns, and months for a single variable, e.g. precipitation, make up the rows. A standard PCA scatter plot of this data would show the separation of months along the first two principal components (PCs), so differences among weather stations instead must be interpreted from the loadings. Data for each weather station were natural-log transformed, scaled (performed on the correlation rather than the covariance matrix), and detrended with linear regression prior to analysis. Varimax rotation was used to prevent the rectangular shape of the geographic area being analyzed from affecting the results (Buell 1975, Serrano et al. 1999). The data were preprocessed with a script written in Scientific Python (www.scipy.org), and the PCA was performed with the R scripting language (R development core team 2009).

Raster mapping to determine the geographic area with high correlation to weather stations in each zone.

A correlation map was created as an estimate of the geographical area represented by the weather stations in each climate zone. The monthly normals from the cluster analysis data set were averaged across all the weather stations in each newly defined climate zone, to produce 36 monthly values (12 months each for DTR, Tmean, and precipitation). Pearson's correlation coefficients were then calculated between the 36 monthly zone averages and the 36 corresponding values associated with each 800m grid cell in the PRISM 1971 – 2000 monthly normal data set. The correlation values for each climate zone were converted into graphical raster files and a mapped. Pixels were color-coded to indicate the climate zones with which they had 0.9 or greater correlation. Several alternative maps were produced with greater or smaller correlation thresholds. The 0.9 correlation threshold was chosen because it provided minimal geographic overlap among zones. Lower correlation thresholds created geographically larger zones and greater overlap.

Analysis of snow cover timing to determine elevation-based strata within climate zones

Snow cover timing was used to estimate the stratification of the weather stations according to elevation. For each weather station, 1971 – 2000 daily snow cover data were downloaded from the NCDC and NRCS web sites (cited above). Since COOP and SNOTEL stations collect snow data incommensurately, with COOP stations reporting snow depth and SNOTEL reporting snow water equivalent (SWE), winter length was estimated as the number of days with snow cover greater than zero. Station data files were organized according to water year, which runs from October 1st – September 30th of the following year and analyzed by a script written in Scientific Python which determined the start and end dates of the winter season. This script contained the following rule set:

1. To correct for the fact that SNOTEL stations record SWE in tenths of inches, but COOP stations do not record a measurement until snow depth exceeds one inch, SNOTEL data were not considered greater than zero until they exceeded 0.5 inches SWE. All non-zero COOP values were counted.
2. To correct for the fact that there are often several isolated snow events in early fall and late spring, winter was not deemed as “started” until the seventh day of non-zero snow cover was encountered in the data file. Similarly, winter was not over until the fourteenth consecutive snow free day was encountered.
3. In the COOP station files, data flagged with “2” (invalid data element), “T” (failed internal consistency check), “U” (failed area consistency check) were replaced with missing values. If any month had more than seven missing values after this replacement, the entire water year was excluded from the analysis.

For each year at each weather station, the first day of winter, last day of winter and length of winter in days were estimated. Mean winter length was then calculated across years. Estimates from stations with fewer than five years of valid data were discarded.

RESULTS AND DISCUSSION

The cluster analysis of 1971 – 2000 monthly normals produced station groupings with strong (78 - 100%) bootstrap support. A comparison of Ward's vs. Average clustering dendrograms defined four climate zones (Fig. 1).

Tercek – Climate Zonation Analysis for Bighorn Canyon NRA

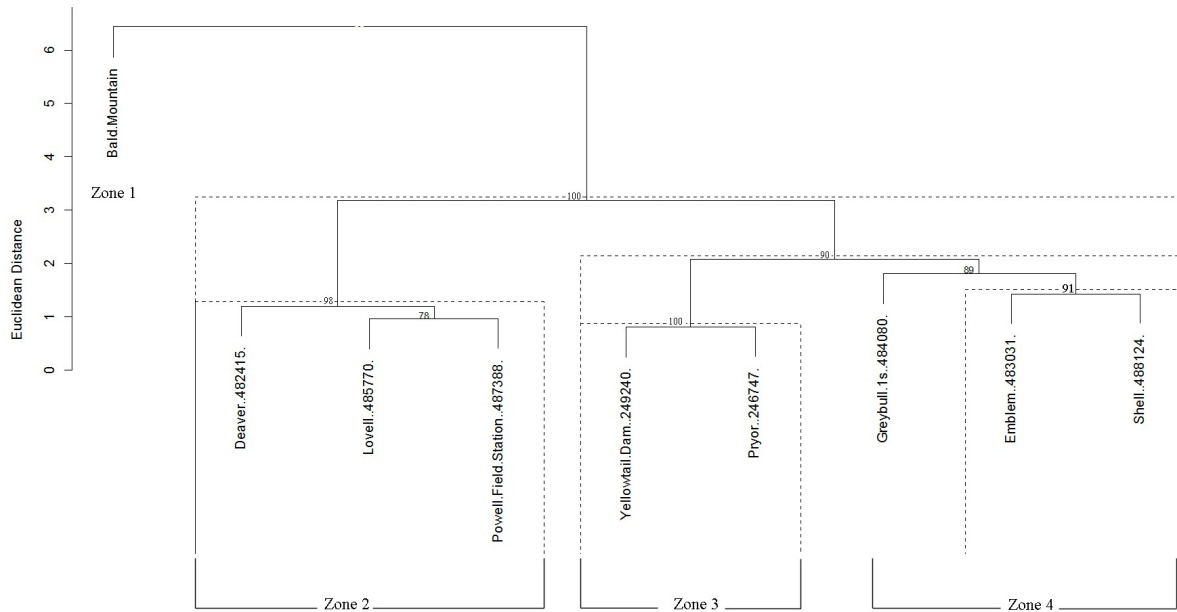


Figure 1. Ward's cluster analysis for weather stations in and near Bighorn Canyon National Recreation Area. Numbers near each node indicate bootstrap confidence, based on 10,000 pseudoreplicates. Dotted lines delineate clusters that have 90% or more bootstrap support. Climate zones, which were selected according to criteria described in the text, have been labeled on the bottom. The dendrogram based on average clustering contained the same four zones.

A comparison of average precipitation, Diurnal Temperature Range (DTR), and mean temperature (Tmean) values shows that Zone 1, which contains the sole SNOTEL station in the analysis, differs markedly from the other three zones (Fig. 2). The cluster analysis was based on

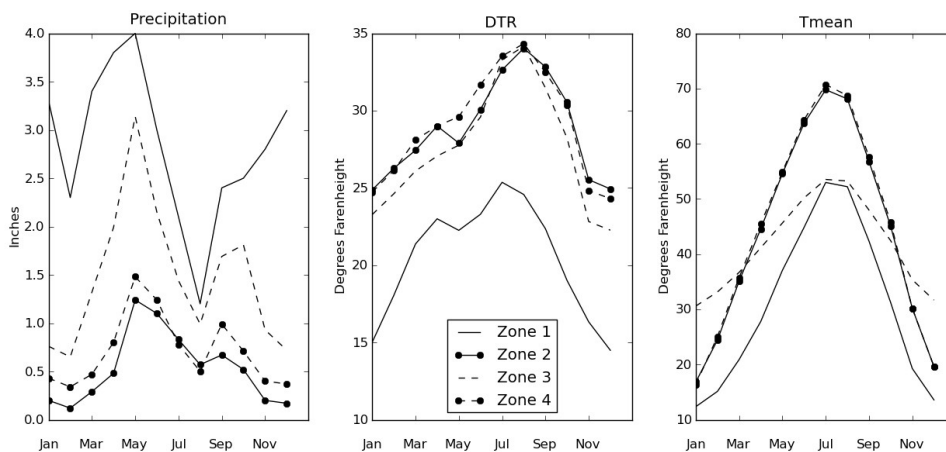


Figure 2. Monthly mean precipitation, diurnal temperature range (DTR), and mean temperature (Tmean) for each climate zone. Means were calculated from 1971-2000 normals from all weather stations in each zone.

standardized (z-score) values, but the data in Fig. 2 are in their original units for better readability.

Because the data in the cluster analysis were standardized, the station groupings in Fig. 1 were based only on month-to-month variability and not the absolute magnitude of the observations. The fact that Zone 1 had greater precipitation than the others did not affect the analysis. Zone 1 is unique in this analysis, and had the greatest Euclidean distance from other stations (Fig. 1), because it had a relatively greater proportion of its precipitation in November – January (Fig. 2). At the same time, all four zones shared relative peaks in precipitation during May and September. Zones 1 and 3 differ from Zones 2 and 4 because they have declines in DTR during May – June. Tmean seasonal patterns are very similar for all zones, but Zone 3 has lower amplitude change over the course of the year (Fig. 2).

The Principal Components Analysis (PCA) of 1895 – 2008 monthly precipitation and Tmean showed better separation for Zones 2 – 4 than the cluster analysis. Loadings for all stations were very similar on principal component 1 (PC1) in all four PCAs, so the distinction among stations was found in PCs 2 and 3, which explained 1 – 7% of the variance (Table 1).

Table 1. Loadings for Principal Components Analysis of 1895 – 2008 monthly precipitation and mean temperature. Weather Station names appear in the rows. Principal Components (PC) appear in columns.

Summer Precipitation	PC1	PC2	PC3	Summer Tmean	PC1	PC2	PC3
Bald Mountain	-0.34	0.07	-0.37	Bald Mountain	0.33	-0.4	0.51
Deaver	-0.34	-0.12	0.46	Deaver	0.34	0.14	-0.27
Emblem	-0.34	-0.26	-0.01	Emblem	0.34	0.07	-0.27
Greybull	-0.33	-0.36	-0.45	Greybull	0.34	0.21	0.26
Lovell	-0.34	-0.1	0.37	Lovell	0.34	0.16	-0.09
Powell Field Stn	-0.33	-0.2	0.4	Powell Field Stn	0.34	0.16	-0.44
Pryor	-0.32	0.6	0.11	Pryor	0.34	-0.03	-0.13
Shell	-0.34	-0.17	-0.34	Shell	0.33	0.4	0.55
Yellowtail	-0.32	0.59	-0.17	Yellowtail	0.32	-0.75	-0.09
Proportion of Variance Explained	0.84	0.05	0.04	Proportion of Variance Explained	0.95	0.02	0.01
Winter Precipitation	PC1	PC2	PC3	Winter Tmean	PC1	PC2	PC3
Bald Mountain	-0.35	0.03	-0.16	Bald Mountain	0.31	-0.42	0.8
Deaver	-0.33	-0.33	0.41	Deaver	0.34	0.14	0.03
Emblem	-0.36	-0.16	-0.12	Emblem	0.34	0.19	0.06
Greybull	-0.34	-0.03	-0.49	Greybull	0.33	0.4	-0.09
Lovell	-0.35	-0.19	0.06	Lovell	0.34	0.18	0.01
Powell Field Stn	-0.32	-0.36	0.45	Powell Field Stn	0.34	0.07	0.02
Pryor	-0.29	0.6	0.35	Pryor	0.33	-0.33	-0.16
Shell	-0.35	-0.04	-0.46	Shell	0.33	0.3	-0.13
Yellowtail	-0.31	0.59	0.08	Yellowtail	0.31	-0.61	-0.55
Proportion of Variance Explained	0.69	0.1	0.07	Proportion of Variance Explained	0.93	0.04	0.02

Loading plots on PCs 2 and 3 show good separation for the all the climate zones with respect to summer and winter precipitation (left side of Figs. 3, 4). In contrast, the Tmean plots (right side of Figs 3, 4) show poor separation for zones 2 and 4 (indicated by green and red text).

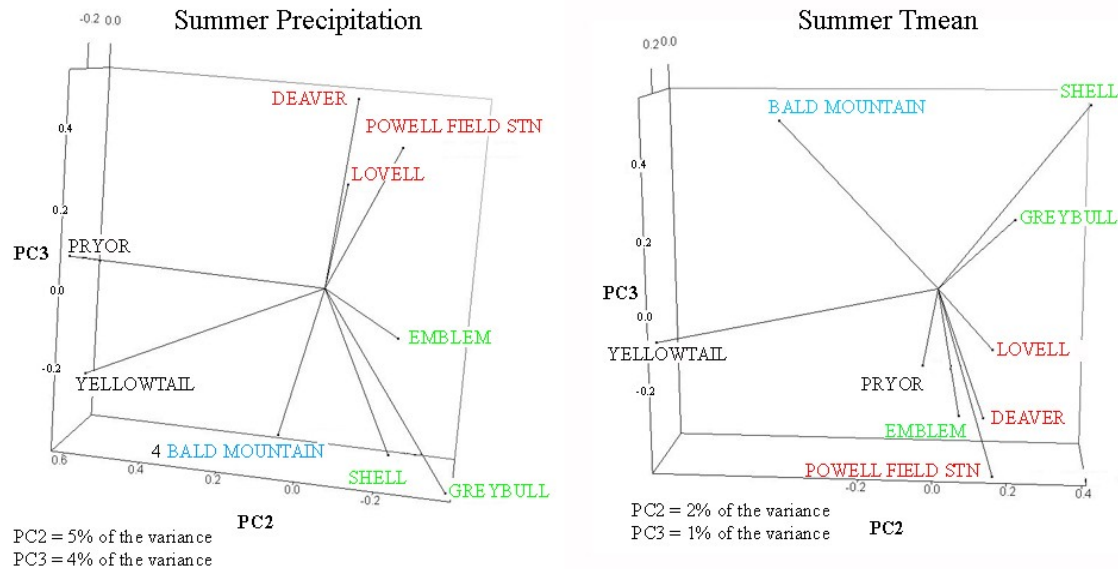


Figure 3. Loading plots of S-mode Principal Components Analysis of 1895 – 2008 summer (June, July, August) precipitation and mean temperature for weather stations in and near Bighorn Canyon National Recreation Area. Colored text indicates zone membership for each station as determined by the cluster analysis of monthly normals. Blue = Zone 1, Red = Zone 2, Black = Zone 3, Green = Zone 4.

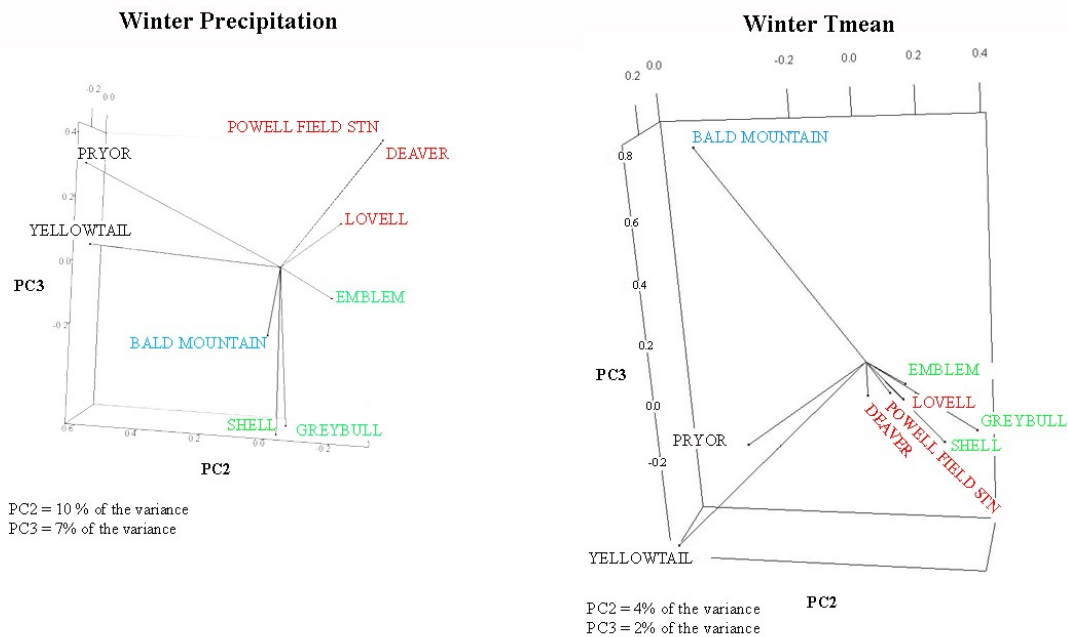


Figure 4. Loading plots of S-mode Principal Components Analysis of 1895 – 2008 winter (December, January, February) precipitation and mean temperature for weather stations in and near Bighorn Canyon National Recreation Area. Colored text indicates zone membership for each station. Blue = Zone 1, Red = Zone 2, Black = Zone 3, Green = Zone 4.

Despite their separation into distinct clusters, the correlation map showed a high degree of overlap in the geographic area represented by the zones. Zone 1, which contains the sole SNOTEL station on Bald Mountain, was the only zone that mapped distinctly (Fig. 5). Despite this overlap, zones 2 – 4 have been maintained as separate groups of stations because they show differing long-term dynamics in the PCAs (Figs 3 and 4).

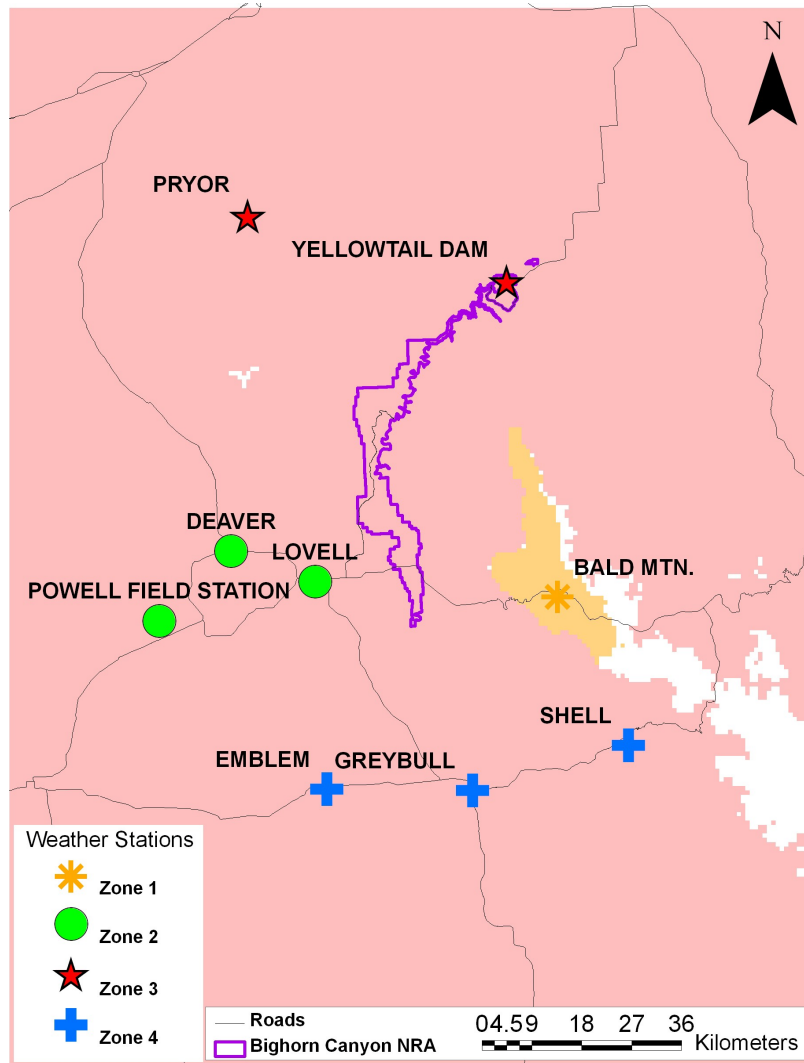


Figure 5. Map showing the location of weather stations that were classified by cluster analysis (Fig. 1). Background colors indicate the areas that have 0.9 or greater correlation with weather stations in each zone. Because of the high degree of overlap, zones 2 – 4 have been plotted with the same color (pink). The boundaries of Zone 1 are indicated in orange.

The snow cover analysis ranked the weather stations according to mean length of winter season,

but because the number of stations involved was small, formal elevation-based strata were not defined (Table 2).

Table 2. Estimated winter length (days) during 1971 – 2000 for weather stations in and near Bighorn Canyon. NRA. Stations are arranged in ascending order of winter length. N/A indicates that there were insufficient data to obtain an accurate estimate. Stations with insufficient data are arranged in ascending order of elevation.

Station Name	Climate Zone	Elevation (m)	Mean Winter Length (Days)	# Winters Calculated	Station Type
Yellowtail	3	1007	N/A	0	COOP
Lovell	2	1170	49.14	7	COOP
Greybull	4	1155	66.11	9	COOP
Deaver	2	1251	70.5	12	COOP
Shell 1ne	4	1279	74.27	11	COOP
Pryor	3	1240	96.22	18	COOP
Powell Field Station	2	1332	N/A	0	COOP
Emblem	4	1356	N/A	0	COOP
Bald Mountain	1	2859	254.87	23	SNOTEL

CONCLUSIONS

Bighorn Canyon National Recreation Area straddles a precipitation gradient that acts as a key ecological driver for many park resources, and it is arguably the most important feature of the NRA's climate. For example, grass-dominated vegetation communities in the north, which receive approximately 18 inches of precipitation annually, differ greatly from the dry semi-desert in the south, which receives an average of only 6.5 inches per year (Knight et al. 1987).

The primary goal of this analysis was to define climate zones with internally consistent seasonal and long-term dynamics. As a result, the absolute magnitude of the weather observations were ignored in favor of their month-to-month or year-to-year patterns of variation. However, despite this study's methodological blindness to the precipitation gradient, the zones presented here can be arranged along it in a recognizable way. Zone 1 (Bald Mountain), which is situated at the highest elevation, reports the highest annual precipitation (Table 2, Fig. 2). Zone 3, which contains the region's most northerly weather stations, reports the second-highest annual precipitation, and the southerly Zones 2 and 4 report the lowest (Fig. 2).

LITERATURE CITED

- Buell, C. 1975. The topography of empirical orthogonal functions. 4th Conference on Probability and Statistics in Atmospheric Science, Tallahassee, FL. American Meteorological Society, Boston, p 188–193.
- Comrie, A.C. and E.C. Glenn. 1998. Principal components-based regionalization of precipitation regimes across the southwest United States and northern Mexico, with an application to monsoon precipitation variability. *Climate Research* 10: 201-215.
- Daly, C., R.P. Neilson, and D.L. Phillips. 1994. A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *Journal of Applied Meteorology* 33: 140 – 158.
- Daly, C., T. Kittell, A. McNab, W. Gibson, J.A. Royle, D. Nychka, T. Parzybok, N. Rosenbloom, and G. Taylor. 2000. Development of a 103-year high-resolution climate data set for the coterminous United States. In: *Proceedings of the 12th American Meteorological Society Conference on Applied Climatology* 249-252.
- Daly, C., G. Taylor, W. Gibson, T. Parzybok, G. Johnson, and P. Pasteris. 2001. High quality spatial climate data sets for the United States and beyond. *Transactions of the American Society of Agricultural Engineers* 43: 1957-1962.
- Davey, C. A., K. T. Redmond, and D. B. Simeral. 2006. Weather and Climate Inventory, National Park Service, Greater Yellowstone Network. Natural Resource Technical Report NPS/GRYN/NRTR 2006/001. National Park Service, Fort Collins, Colorado.
- Easterling, D.R., B. Horton, P. Jones, T. Peterson, T. Karl, D. Parker, M. Salinger, V. Razuvayev, N. Plummer, P. Jamason, C. Folland. 1997. Maximum and minimum temperature trends for the globe. *Science* 277: 364-367.
- Fovell, R.G., and M.C. Fovell. 1993. Climate zones of the conterminous United States defined using cluster analysis. *Journal of Climate* 6: 2103 – 2135.
- Frakes, B. , S. Ostermann-Kelm, I. Ashton, J. Burke, R. Daley, M. Tercek, D. Pillmore, C. Jean, M. Britten, S. Gray, and T. Kittel. 2009. Rocky Mountain Climate Protocol: Climate monitoring in the Greater Yellowstone and Rocky Mountain Inventory and Monitoring Networks.
- Gray, S. T. 2008. Framework for linking climate, resource inventories and ecosystem monitoring. Natural Resource Technical Report NPS/GRYN/NRTR-2008/110. National Park Service, Fort Collins, Colorado.
- Kittell, T., S. Ostermann-Kelm, B. Frakes, M. Tercek, S. Gray, C. Daly. 2010. A Framework for Climate Analysis and Reporting for Greater Yellowstone (GRYN) and Rocky Mountain (ROMN) Networks. A Report from the GRYN/ROMN Climate Data Analysis Workshop Bozeman, MT — 7-

8 April 2009.

Knight, D. H., G. P. Jones, Y. Akashi, and R. W. Myers. 1987. Vegetation ecology in the Bighorn Canyon National Recreation Area. Laramie, Wyo.: University of Wyoming, Department of Botany. R Development Core Team (2009).

R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R.org>, project.org.

Serrano, A, J. Garcia, V. Mateos, M. Cancillo, and J. Garrido. 1999. Monthly modes of variation of precipitation over the Iberian Peninsula. *Journal of Climate* 12: 2894-2919.

Solomon, S., and many co-authors. 2007. Climate Change 2007: The Physical Science Basis. Technical Summary, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC). (Available in PDF format online at <http://www.ipcc.ch>).

Suzuki, R. and H. Shimodaira. 2006. Pvcust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics* 22: 1540 – 1542.

Tercek, M., S. Ostermann-Kelm, and S. Gray. In prep. Climate zones for Yellowstone and Grand Teton National Parks.

Unal, Y., T. Kindap and M. Karaca. 2003. Redefining the climate zones of Turkey using cluster analysis. *International Journal of Climatology* 23: 1045–1055.

Vose, R.S., D. Easterling, B. Gleason. 2005. Maximum and minimum temperature trends for the globe: an update through 2004. *Geophysical Research Letters* 32: L23822.